Linear and Nonlinear Causal Analysis with Sparse Modeling

by

Nived Madhusoodanan

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Process Control

Department of Chemical and Materials Engineering
University of Alberta

Abstract

Data has become an integral part of analysis and decision-making across various disciplines including health, economy, biology, process systems engineering, and sports. Despite the abundance of big data, the identification of the fundamental source of disruption in process systems engineering continues to pose a formidable challenge. Recognizing the significance of the rapidly evolving industrial landscape, in this thesis, we focus on various ways of using Causal Inference to determine the relationship among process variables to identify the root source of process abnormalities.

The first contribution of the thesis considers the causal inference among non-linear systems. The research investigates the applicability of the Gaussian Process in the Causality analysis among non-linear systems. Employing Gaussian processes facilitates the identification of causal relationships, thereby enabling a better understanding of complex non-linear systems. Since the traditional Gaussian Process-based causal inference tends to provide spurious causations, the study focuses on deriving sparse solutions to enhance the interpretability of the causal links identified. Sensitivity analysis is conducted to assess the robustness of the findings, providing insights into the reliability of the identified causal relationships within the intricate system dynamics.

Given that conventional causal approaches rely solely on data, the integrity of the results is heavily contingent upon the quality of the data. Owing to the susceptibility of real-world industrial data to inaccuracies stemming from inadequate sensor maintenance, the effectiveness of complete data-based approaches is notably compromised. Hence, the second contribution of the thesis explores the possibility of amalgamat-

ing expert human knowledge with process data to reduce the over-dependence on data and produce more reliable and accurate modeling. In this case, it is achieved by using a constraint optimization technique where human knowledge is provided as constraints in the optimization. To verify the existence of the derived causal relationships and eliminate the possibility of chance findings, a novel surrogate generation algorithm specifically designed for oscillating data sets is also proposed.

The third part of the thesis applies the previously developed methodology, integrating physics-based information with process data, to diagnose the challenge of flooding in separation columns using real industrial process data. Moreover, an advanced predictive model for forecasting instances of flooding is formulated and its results are analyzed. A user-friendly Graphic User Interface (GUI) toolbox is also developed to automate the process of combining process data and domain knowledge.

The efficacy of all the contributions is verified through the numerical case study or industrial case study.

Preface

This thesis is an original work conducted by Nived Madhusoodanan under the supervision of Dr.Biao Huang and is funded in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

1. A research article is being prepared based on Chapter 2 of this thesis - "Nonlinear Sparse Causal Inference based on Gaussian Process". **N.Madhusoodanan**, R.Chiplunkar, V.K.Puli, B.Huang

2. Chapter 3 of this thesis has been published as **N. Madhusoodanan**, R. Chiplunkar, VK Puli, B. Huang, "Physics-informed sparse causal inference for source detection of plant-wide oscillations". AIChE J. 2024; e18362. doi:10.1002/aic.18362

3. A research article is being prepared based on Chapter 4 of this thesis which is an application of the research methodology developed in Chapter 3 for diagnosis and advanced prediction of flooding in separation columns of real process industries

Nived Madhusoodanan is responsible for the development of ideas, algorithm derivation, simulation, and manuscript preparation for all the research articles listed above. Dr.Biao Huang has provided indispensable guidance, and critical feedback on the technical content, and meticulously reviewed the manuscript for all the aforementioned research articles. Dr.Ranjith Chiplunkar and Dr.Vamsi Krishna Puli have guided the simulations and manuscript preparation through discussions. The data for the application presented in Chapter 4 was provided by Quincy Hu from Shell Canada. For proprietary reasons, all data have been normalized in the thesis

Dedicated to

My loving parents Thanooja and Madhusoodanan, my wife Anjitha, and all my teachers

Acknowledgements

This thesis would not have been possible without the efforts and contributions of many, the foremost of whom is my supervisor Dr.Biao Huang. I express my sincerest gratitude to Dr.Biao Huang for allowing me to work under his supervision. Working in his group has been a greatly rewarding experience for me both in terms of technical and personal aspects. I appreciate his patience and support throughout my MSc. His insightful suggestions have greatly helped me in resolving many challenges I faced in research as a graduate research student. I am also grateful to him for exposing me to various real-world process systems engineering problems through industrial projects and discussions.

I am deeply grateful for the enriching experience of working in an expansive research group, which has exposed me to diverse areas of research. I extend my heartfelt gratitude to Dr. Ranjith Chipunkar for his invaluable encouragement, unwavering support, insightful feedback, and dedicated time spent reviewing my research and manuscripts. Additionally, I am indebted to Dr. Vamsi Krishna Puli for his enlightening discussions that have significantly contributed to the advancement of my research. I extend my sincere appreciation to my esteemed colleagues in the research group - Amir, Jainish, Jansen, Alireza, Amirezza, Yousef, and others - for their valuable camaraderie and collaboration. A special acknowledgment goes to Mrs. Terry Runyon for her support throughout my research journey.

I would like to express my sincere gratitude to the Natural Sciences and Engineering Research Council (NSERC) of Canada for their generous financial support. My heartfelt thanks also go to the University of Alberta and the Department of Chemi-

cal and Materials Engineering for their invaluable financial support and provision of necessary resources, enabling the smooth execution of my research endeavors.

I am immensely grateful to my friends Rahul, Abhishek, and Gaurav, whose unwavering companionship has not only enriched my graduate studies but also made this journey truly memorable. My deepest gratitude is reserved for my beloved wife Anjitha, whose unwavering support and understanding have been my pillars of strength throughout this challenging endeavor. I am forever indebted to her for her love and sacrifices, which have been the cornerstone of my success. I owe a debt of gratitude to my parents, Thanooja and Madhusoodanan, for their endless love, unwavering encouragement, and unyielding sacrifices. This journey would not have been possible without their relentless support and guidance. Their unwavering faith in my abilities has been the driving force behind my achievements.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The application of data analytics and data-driven modeling is finding increased prominence with ever-improving computational capabilities. Data has become an integral part of analysis and decision-making across a wide variety of disciplines including health, economy, biology, process systems engineering, and even sports. The primary objective in the domain of process systems engineering is always related to safe and reliable plant operation apart from economic benefits, and this is achieved by early detection and prediction of the root cause of abnormalities. Data-driven modeling plays an important role in the identification of the root cause of abnormalities as first principle-based modeling has its limitations due to limited understanding of the processes, presence of feedback, and complex interconnections [1]. Causality analysis is a widely accepted data-based technique for root source identification and hence has been extensively studied across multiple disciplines [2–4]. Since the analytics methods are sensitive to data quality, results obtained are subject to variations depending on the quality and reliability of the data.

In process industries, the importance of data quality cannot be overstated, particularly when performing tasks such as Root Cause Analysis through causal inference, as highlighted in this thesis. Data quality serves as the cornerstone for accurate and reliable analyses, providing the foundation upon which meaningful insights and informed

decisions can be built. In process systems, where intricate networks of interconnected variables and parameters influence outcomes, the precision and reliability of data are paramount. Flawed or inaccurate data can lead to misguided causal inferences, potentially resulting in ineffective or even counterproductive interventions [5, 6]. Reliable and accurate data contributes to the overall efficiency, safety, and sustainability of process industries. Since engineers and technicians working in process industries possess years of experience, combining them with the data helps in mitigating the issues associated with reduced quality data as the heavy reliance on data is overcome.

As all the process systems are not linear and many involve nonlinear relationships among their variables, nonlinear causal inference is explored in the thesis. Nonlinear causal inference holds significant importance in various fields, including process systems engineering, as it allows for a more nuanced understanding of complex relationships that may not follow linear patterns. Nonlinear causal inference techniques, including advanced machine learning algorithms, empower researchers and practitioners to uncover hidden patterns, identify complex dependencies, and discern the nonlinear dynamics that underlie intricate processes. Due to the non-parametric nature, Gaussian Process-based nonlinear causal inference attracts a lot of interest. As these methodologies may involve producing spurious causations in addition to true underlying causations, sparse-based nonlinear causal inference is explored in this thesis.

Therefore, in this thesis, novel methods are proposed to address the problems associated with complete reliance on data for causal inference and the issue of spurious causations. While over-dependence on data is overcome by proposing a Physics-Informed causal inference, the issue of spurious causations in nonlinear causal inference is addressed by performing sensitivity analysis. In physics-informed causal analysis, the problem is formulated as a constraint optimization to achieve more accurate and reliable inference. The developed framework has been then applied to the diagnosis of flooding in the separation column where process data is obtained through

industrial collaboration. In the forthcoming sections, contributions and an outline of the thesis are detailed.

1.2 Thesis Contributions

The thesis contributes to the development of methodologies for root cause analysis in Process Systems Engineering using causal inference. Contributions of the thesis are listed in detail as follows.

1. A nonlinear sparse causal inference using variational Gaussian process is proposed focusing on nonlinear time-invariant systems.

2. A method of amalgamating human knowledge with data to improve the accuracy and reliability and overcome the shortcomings of purely data-based approaches for developing causal maps for systems involving oscillations is proposed.

3. A surrogate data-based hypothesis testing for systems involving oscillations is proposed to verify that the causal-effect relationships obtained are not by chance.

4. Physics-Informed sparse causal inference approach is applied for diagnosis of flooding in separation column and a Principal Component Analysis-based Hotelling's T^2 statistics method is developed for advanced prediction of flooding in separation column.

1.3 Thesis Outline

With the presented motivations and the contributions reviewed in Chapter 1, the thesis proceeds to detail each of the contributions in the forthcoming chapters. The rest of the thesis is organized as follows.

Chapter 2 presents the first contribution of the thesis which is nonlinear sparse causal inference using variational Gaussian process. This chapter focuses on nonlinear

time-invariant systems. Traditional Granger-based approaches fail to perform well for nonlinear systems and the regular Gaussian process approach results in a significant number of spurious causations in addition to the true causal relations. The proposed methodology significantly mitigates spurious causation by incorporating a sensitivity term. The effectiveness of the method is demonstrated through case studies.

Chapter 3 discusses the second contribution of the thesis in detail which is the Physics-Informed sparse causal inference for source detection of systems involving oscillations. The shortcomings of traditional data-based approaches for root source identification in systems involving oscillations are discussed and causal inference methods like Granger causality and Sparse Granger causality are detailed in the chapter. To verify the cause-effect relationships identified, a novel surrogate data-based hypothesis testing for systems involving oscillations is proposed. The efficacy of the proposed approach is demonstrated through simulation and industrial case studies.

The third contribution of the thesis is discussed in Chapter 4 which applies the developed Physic-Informed sparse causal inference approach for diagnosis and advanced prediction of flooding in separation column for a real industrial data set. After performing a diagnosis using the proposed methodology, an algorithm combining dimensionality reduction and anomaly detection techniques is developed for advanced prediction of flooding. To streamline the application of the proposed methodology, a Graphic User Interface (GUI) toolbox is developed. The chapter thoroughly explores the features and implementation of this toolbox.

Chapter 2

Non-linear Sparse Causal Inference using Gaussian Process

This Chapter presents the first contribution of the thesis which is developing a non-linear sparse causal inference approach using variational Gaussian Process models.

2.1 Introduction

The chapter focuses on exploring the intricacies of causal inference within non-linear time-invariant dynamical systems. In systems involving complex relationships and interactions, traditional causal inference approaches like Granger causality do not perform well especially when confronted with nonlinear dynamics. After the introduction of Granger Causality, researchers have proposed various formulations that are either extensions or modifications of the Granger framework. However, the framework of Granger causality and such extensions are primarily based on the autoregressive-based modeling and hence can be applied only to linear models requiring the condition of separability to be satisfied which assumes that the driven information from causative factors can be removed from the effects [2]. It has been observed that Granger causality, in its traditional sense, may not be well-suited for detecting casual relationships in nonlinear deterministic dynamical systems [7]. Researchers have developed and extensively studied various methods to address this limitation and quantify causal relationships among variables in nonlinear systems. These methods

include transfer entropy[8], non-linear cross mapping using state space reconstruction models [9, 10], conditional mutual information [11], and recurrence plots [12]. Faes et al. [13] proposed a causality estimator based on nonlinear exogenous auto-regressive (NARX) modeling but suffer from the fact that it is not appropriate for higher order non-linear systems. Chen et al. [14] proposed a nonlinear extension to Granger causality, termed an extended Granger causality, based on fitting locally linear models (AR) to randomly selected neighborhoods of the embedded time-series and estimating the neighborhood Granger causality. Ancona et al. [15] proposed a bi-variate nonlinear causality estimator that replaces auto-regressive-based modeling with the Radial Basis Function. Marinazzo et al. [16] proposed a Kernel-Granger causality-based estimator, initially for bi-variate systems, and then later extended it to multivariate estimations [17]. Even though kernel-based approaches provide much better performance compared to other existing approaches, the author pointed out the fact that the performance of this approach will be substantially reduced in the presence of significant noise [18]. Schiff et al. [19] presented another approach using mutual prediction which is not explicitly a causal measure but more a measure of the direction of information flow.

Majority of the approaches discussed above suffer significantly from the problem of over-fitting and many of them result in spurious causations [20, 21]. By specifying particular conditions on the mapping mechanism and the distributions of the cause and noise variables, the causal direction becomes identifiable [22]. For instance, Hoyer et al. [23] assume that the effect is a non-linear transformation of the cause plus some independent additive noise. A potential drawback of these methods is that the assumptions made by the particular model considered could be unrealistic for the data under study. Pearl and Judea [24] in their work proposed a non-parametric approach for non-linear causal inference but was based on Markov factorization because of which temporal and causal inference from samples further in the past are ignored in modeling. Hence in this work, we propose a Gaussian Process-based nonlinear

sparse causal inference approach. Whether we try to find the underlying behavior of the data, estimate the parameters of the model, or identify the interactions and relationship between the input and output, it requires us to assume a certain function for these operations. More often than not, the specific algebraic form of such an underlying function is unknown to us. In these situations, Gaussian process regression can serve as a useful tool for performing inference both passively (for example, describing a given data set as best as possible, allowing one also to predict future data) as well as actively (for example, learning while choosing input points to produce the highest possible outputs) [25]. Gaussian process regression is a powerful, non-parametric Bayesian approach to regression problems that can be utilized in exploration and exploitation scenarios [26]. In their investigation, Amblard et al. [27] utilized Gaussian Processes for causal inference. Nevertheless, the susceptibility to spurious causations renders their approach impractical for higher-dimensional data. Feng et al. [28] used the Bayesian version of Convergent Cross Mapping (CCM) using deep Gaussian processes (DGPs) for causal discovery in coupled time series, overcoming limitations of traditional CCM but again the approach is computationally expensive and suffers from spurious causations.

The proposed Gaussian process-based non-linear causal inference addresses the limitations of existing approaches. Non-parametric estimation is driven by the data itself rather than being constrained by a predetermined parametric model. This flexibility enables adapting to various data distributions and structures. It can deal with variables having nonlinear structure, and there is no restriction to the order of nonlinearity. As the modeling in this work involves the concepts of variational inference and inducing points, the model can deal with higher dimensional data sets. Variational inference offers scalability and computational efficiency, making it advantageous for approximating complex posterior distributions in Bayesian models, especially in scenarios with large datasets or intricate models [29]. Its flexibility in choosing variational families allows adaptation to diverse problem characteristics, and applications

extend to various machine learning tasks [30]. The sensitivity analysis carried out in this work is an extension of LASSO Granger causality and it helps in overcoming the problem of spurious causations. However, the extension is not straightforward for Gaussian process-based approaches. Thus the main contribution of this chapter is outlined below:

1. A novel framework for sparse non-linear causal inference using variational inference and Gaussian process

2. A novel sensitivity analysis to overcome spurious causations

The remainder of this chapter is organized as follows. Section 2.2 provides an overview of the concepts of GC and Section 2.3 provides an overview of sparse Granger causality (SGC) which is a variant of GC. Section 2.4 provides an overview of the Gaussian Process Regression, and testing causality using the Gaussian Process is presented. Section 2.5 provides an overview of the Sparse Variational Gaussian Process. In Section 2.6, a novel framework of nonlinear sparse causal inference using variational Gaussian Process is presented. In Section 2.7, simulation and industrial case studies are presented that demonstrate the effectiveness of the proposed methodology. The concluding remarks are presented in Section 2.8.

2.2 Granger Causality

Granger [31] proposed an approach to identify causal relations among the variables of linear stationary processes which is popularly referred to as Granger causality. The two important axioms of GC are

- Cause happens prior to the effect

- Cause has unique information about the future of its effect

Let X and Y be two sets of time series data. If the past values of X help in improving the prediction of Y, then we say X Granger causes Y. To understand the concept

of GC from a mathematical point of view, let the prediction equations for two time series X and Y be :

$$y(t) = a_1 y(t-1) + a_2 y(t-2) + ... + a_d y(t-d) + e_1(t) \tag{2.1}$$

$$\begin{aligned} y(t) = {} & a_1 y(t-1) + a_2 y(t-2) + ... + a_d y(t-d) \\ & + b_1 x(t-1) + b_2 x(t-2) + ... + b_d x(t-d) + e_2(t) \end{aligned} \tag{2.2}$$

where d is the order of the vector auto-regressive (VAR) model, and $e_1(t)$ and $e_2(t)$ are white noises for the above two processes.

If it is observed that $\text{var}(e_2) < \text{var}(e_1)$, it can be inferred that with the addition of the past values of X to predict the future of Y, the prediction accuracy has increased. In this case, one can confirm that X Granger causes Y.

If one writes it in a Vector autoregressive (VAR) model in terms of instantaneous (X_t) and delayed variables (X_{td}), which is the most widely used approach for detecting Granger Causality, we have

$$X_t = M \; X_{td} + e_t \tag{2.3}$$

where,

$$X_t = \begin{bmatrix} x_1(t) \\ \vdots \\ x_2(t) \end{bmatrix}; \quad M = \begin{bmatrix} M_{11} & \dots & M_{1(p\times d)} \\ \vdots & \vdots & \vdots \\ M_{1p} & \dots & M_{(p\times d)(p\times d)} \end{bmatrix};$$

$$X_{td} = \begin{bmatrix} x_1(t-1) \\ \vdots \\ x_1(t-d) \\ \vdots \\ x_p(t-1) \\ \vdots \\ x_p(t-d) \end{bmatrix}. \tag{2.4}$$

As an alternative to analyzing causal relations based on the variance of the residual e_t, one may use the coefficient matrix M to achieve the same objective. The test on the variance of the residuals translates into a test on the elements of the coefficient matrix [32]. Thus, the GC inference procedure involves the identification of M followed by a test for the significance of the elements of M. The first step of identifying M is formulated as an optimization expressed as follows:

$$\min_{M} \quad ||X_t - MX_{td}||_2 \tag{2.5}$$

Here, $||\cdot||_2$ represents the l^2-norm.

Despite the successful application of Granger Causality (GC) in inferring causal relationships from stochastic data, it suffers from a notable limitation. Due to the nature of Granger Causality, which involves conducting cross-correlation between the source and effect across various time lags, its application to oscillatory data often leads to spurious causations. To illustrate this limitation, we present a simulated example that highlights the challenges of using GC with oscillatory signals. Consider two signals $x(t)$ and $y(t)$ that are generated using the equations 2.6-2.7, as shown in Figure 2.1.

$$x(t) = \cos(0.01\, t) + e_x(t); \tag{2.6}$$

$$\begin{aligned} y(t) = {} & 0.6x(t-1) - 0.7y(t-1) - 0.5y(t-2) \\ & + 0.1y(t-3) + e_y(t); \end{aligned} \tag{2.7}$$

where $e_x(t)$ and $e_y(t)$ are Gaussian white noise with variances of 0.05 and 0.1 respectively.

Figure 2.2 displays the cross-correlation between $x(t)$ and $y(t-\tau)$ for all $0 \leq \tau \leq N$. The red horizontal lines represent the approximate upper and lower 95% confidence bounds. Notably, the cross-correlation exceeds the confidence region for a substantial number of positive lags τ. As a result, it is possible to mistakenly infer a causal

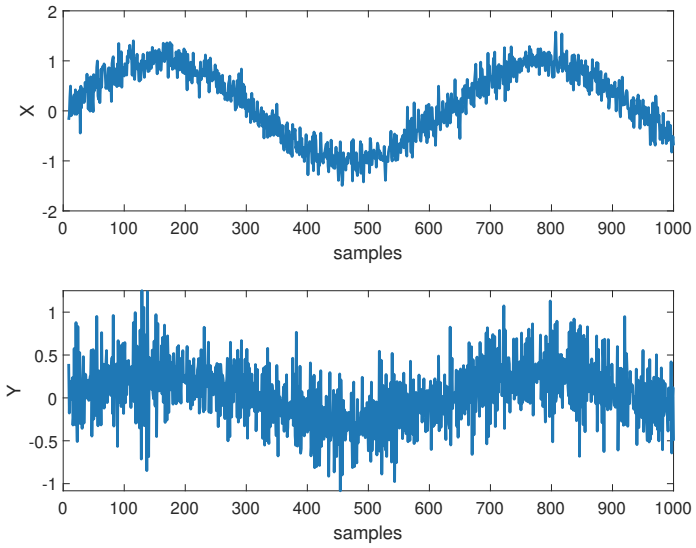


Figure 2.1: X and Y series

relationship from Y to X, despite this not being the case in the true relation. This observation is further supported by fitting a vector auto-regressive model to X, which reveals significant coefficients corresponding to the past regressors of Y.

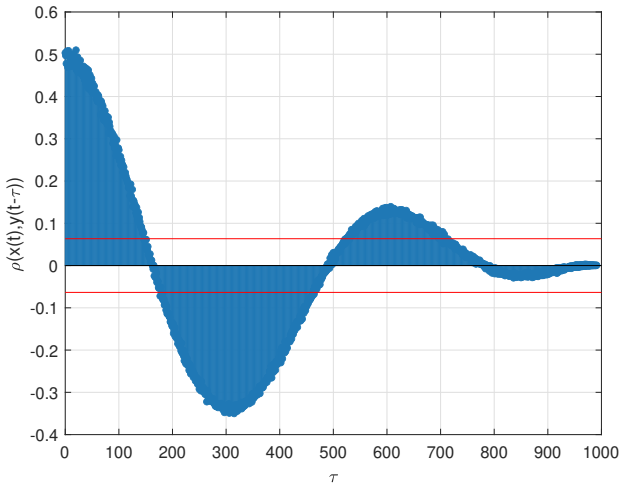


Figure 2.2: Test for Y causing X

2.3 Sparse Granger Causality

SGC is an extension of GC, and as the name suggests, is aimed at obtaining a sparse causal network. Valdes-Sosa et al [33] proposed to enforce the estimation of sparse AR coefficients using l^1-norm regularized models such as the LASSO [34]. As the dimension of the data set increases, GC becomes computationally expensive and tends to provide a lot of cause-effect relationships, which might not be of significant causal relation in the true network. Apart from that, when the system under consideration involves periodic oscillations, GC always tends to produce spurious causation, which causes misleading results. The sparse Granger approach overcomes this shortcoming of GC by introducing a sparsity term or a penalty term in the objective function of GC. One of the best-known types of penalty in statistical learning literature is the LASSO [35]. The objective function of SGC therefore is:

$$\min_{M} \quad ||X_t - MX_{td}||_2 + \lambda||M||_1 \tag{2.8}$$

The penalty term is the sum of the absolute values of the coefficients multiplied by a hyperparameter λ. The level of sparsity to be achieved is decided by λ. The higher the value of λ, the sparser the resulting VAR model. Thus, SGC mitigates the issue of spurious causations observed in the GC method. But on the flip side, it may also remove weak causal relations present in the network leading to a failure in identifying existing causations. To deal with this issue, additional modifications are needed to the VAR identification problem, which leads to the proposed physics-informed SGC framework.

2.4 Gaussian Process Regression

Regression analysis is a fundamental tool for modeling relationships between variables. Traditional regression methods often rely on assuming a specific functional form, limiting their flexibility in capturing complex patterns. A Gaussian Process

is a collection of random variables, any finite number of which have (consistent) joint Gaussian distributions [36]. Gaussian Process Regression (GPR) offers a non-parametric Bayesian alternative by modeling distributions over functions, allowing for adaptive and data-driven predictions without imposing rigid assumptions. At the core of GPR is the notion of Gaussian processes, representing distributions over functions. The choice of kernel functions, determining the similarity between data points, and hyperparameters governing the characteristics of these functions are pivotal in shaping the model's behavior [37]. By leveraging Bayesian principles, GPR provides a principled way to incorporate prior beliefs and update them with observed data.

In the context of a regression task, let (X, Y) represent the training dataset, where X and Y are float vectors of length n, denoting the input features and corresponding output values, respectively, with n being the number of training data points. The objective is to determine a regression function that maps the input vectors X to the target output values Y. The model has two parts - the prior as shown in equation 2.9 and the likelihood as shown in equation 2.10. The likelihood is the same as the Gaussian likelihood in equation 2.10.

$$\begin{bmatrix} f(x_*) \\ f(x) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} k(x_*, x_*) & k(x_*, x) \\ k(x_*, x)^T & k(x, x) \end{bmatrix}\right) \tag{2.9}$$

$$y(x) \sim \mathcal{N}(f(x), \eta^2 I_n) \tag{2.10}$$

where k is the kernel function, which encapsulates a prior understanding of the smoothness of the underlying function.

The Gaussian Process prior (GP prior) is a multivariate Gaussian distribution over random variable vectors $f(X)$ and $f(X_*)$. $f(X)$ is a random variable vector of length n which represents possible values of the underlying regression function at training locations X and $f(X_*)$ is also a random variable vector that represents possible values

of the underlying regression function at testing locations X_*.

In the likelihood, $y(x)$ is a random variable vector of length n. It comes from a multivariate Gaussian distribution with mean $f(X)$, and covariance $\eta^2 I_n$, where η^2 is a scalar model parameter called noise variance, and I_n is an $n * n$ identity matrix. The marginal likelihood then can be written as:

$$p(f) = \frac{1}{(2\pi)^{n/2} det(K)^{1/2}} \exp(\frac{-1}{2} y^T K^{-1} y) \tag{2.11}$$

Prediction in GPR is performed by incorporating a new test point X_* and conditioning on the observed data along with hyperparameters. To make predictions for f_* at testing locations X_*,

$$p(f_*|y) = \mathcal{N}(k_{*x}(K + \eta^2 I_n)^{-1} y,\ k_{**} - k_{*x}(K + \eta^2 I_n)^{-1} k_{*x}^T) \tag{2.12}$$

The Gaussian Process (GP) stands as a non-parametric model, necessitating the presence of training data explicitly during test time to formulate the predictive distribution, as evident from the expression mentioned above. However, Gaussian Processes become computationally prohibitive for extensive datasets, demanding $O(N^3)$ time for training due to the inversion of the covariance matrix. Subsequent predictions, once the inversion is completed, incur $O(N)$ complexity for the predictive mean and $O(N^2)$ for the predictive variance per new test case.

2.4.1 Testing Causality using Gaussian Process

Consider two scalar time series x_t and y_t and two structural models as shown below:

Model 1:

$$\begin{aligned} x_t &= f_x^1(x^{t-1}) + \epsilon_{x,t}^1 \\ y_t &= f_y^1(y^{t-1}) + \epsilon_{y,t}^1 \end{aligned} \tag{2.13}$$

Model 2:

$$x_t = f_x^2(x^{t-1}, y^{t-1}) + \epsilon_{x,t}^2$$
$$y_t = f_y^2(y^{t-1}, x^{t-1}) + \epsilon_{y,t}^2 \tag{2.14}$$

where f is distributed as a Gaussian process and the dynamical noises $\epsilon_{x,t}^{1,2}$ and $\epsilon_{x,t}^{1,2}$ are i.i.d.

The marginal likelihood of the model also called the evidence of the model, is used to test causality. Consider a causal relationship $x- > y$ and to test whether such a causal relationship exists, a statistic $d_{x->y}$ that compares the log-evidences of Model 1 and Model 2 is defined as follows.

$$d_{x->y} = \max_{\theta_2} \ logP_2(f_y|X, Y) - \max_{\theta_1} \ logP_1(f_y|Y) \tag{2.15}$$

Based on the computed value of d, the inference is that a causal relationship from x to y is established when $d_{x->y} > 0$. This inference is grounded in the understanding that a positive d signifies a higher likelihood of the model incorporating the historical information of x for the prediction of y compared to the scenario where y's past is not considered. Conversely, when $d_{x->y} < 0$ it is deduced that x does not cause y to adhere to the same reasoning.

2.5 Sparse Variational Gaussian Process

Sparse Variational Gaussian Processes (SVGP) represent a powerful extension of standard Gaussian Processes, addressing scalability concerns by efficiently handling large datasets. In SVGP, the key idea is to introduce a set of inducing points, which are a smaller, carefully chosen subset of the original data points. These inducing points act as a representative sample, enabling the approximation of the full posterior distribution. By optimizing the positions of these inducing points and the associated hyperparameters, SVGP strikes a balance between computational efficiency and maintaining the expressive power of Gaussian Processes.

2.5.1 Sparse Formulation

In the Gaussian Process regression model, both the objective function for parameter learning and the predictive distribution require the matrix inversion$(K_N + \sigma^2 I)^{-1}$ as discussed in 2.4. For instance, a GPR model for a dataset with 10,000 data points needs 10^{12} operations to invert its covariance matrix. Since the size of data n, controls the shape of the covariant matrix K, there are different ways to approach the problem. One such approach is down-sampling by selecting a subset of the training data. Even though it is easy to implement, it is difficult to decide which part of the training data to remove. The most widely accepted approach is the introduction of inducing points which are commonly referred to as pseudo points.

Consider a new set of random variables $f(X_s)$, shortened as f_s, referred to as inducing random variables at some locations X_s, referred to as inducing locations. X_s is a vector of scalars of length n_s. where the subscript s stands for sparse. The size of the inducing points n_s is chosen such that it is much smaller than n. A multivariate Gaussian distribution is used to establish the relationship between f_s and f and hence the new sparse GP prior is:

$$\begin{bmatrix} f(X) \\ f(X_s) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} k(X,X) & k(X,X_s) \\ k(X,X_s)^T & k(X_s,X_s) \end{bmatrix}\right) \tag{2.16}$$

The likelihood is the same as the Gaussian likelihood in equation 2.10. Here the marginal likelihood serves as the quantity that measures how well the inducing variables summarize the training data. The derivation of the marginal likelihood for the Sparse Variational Gaussian Process (SVGP) is presented below.

$$p(y) = \int\int p(y|f, f_s)\ p(f, f_s)\ df\ df_s \tag{2.17}$$

Dropping f_s in the likelihood and re-organizing the terms gives

$$p(y) = \int p(y|f)(\int p(f, f_s) df_s)\ df \tag{2.18}$$

Integrating f_s out and substituting with the corresponding distribution gives

$$p(y) = \int \mathcal{N}(y; f, \eta^2 I_n)\ \mathcal{N}(f; 0, K)\ df \tag{2.19}$$

By using the Gaussian linear transformation theorem and further simplifying the obtained equation results in :

$$p(y) = \mathcal{N}(y; 0, K + \eta^2 I_n) \tag{2.20}$$

The marginal likelihood obtained is the same as the GPR model. This is because when f_s is integrated out from the joint distribution of $p(f, f_s)$, it results in the marginal distribution $p(f)$ and also removes all references of X_s. This makes the computation of the posterior distribution difficult as shown below.

$$p(f, f_s|y) = \frac{p(y|f, f_s)\ p(f, f_s)}{p(y)} \tag{2.21}$$

When applying Bayes' rule for posterior inference, the computation involves determining the marginal likelihood, $p(y)$. However, calculating $p(y)$ often encounters challenges. To address this issue, an alternative approach is to leverage variational inference techniques. Unlike traditional Bayesian methods that rely on the explicit computation of $p(y)$, variational inference directly approximates the posterior distribution $p(f, f_s|y)$. By bypassing the explicit computation of $p(y)$, variational inference provides a practical and computationally efficient means of estimating the posterior without the constraints associated with marginal likelihood computation.

2.5.2 Variational Inference

The posterior distribution, denoted as $p(f, f_\mathrm{s}|y)$, represents a joint distribution over the random variable vectors f and f_s. In the context of variational inference, the objective is to approximate this true posterior using a variational distribution, denoted as $q(f, f_\mathrm{s})$. This variational distribution is defined over the same set of random

variables, and the goal is to ensure that $q(f, f_s)$ closely aligns with the characteristics of the true posterior. For mathematical trackability, $q(f, f_s)$, which is a joint distribution, is defined by the following factorization:

$$q(f, f_s) = p(f|f_s)\ q(f_s) \tag{2.22}$$

Here $p(f|f_s)$ is derived by applying the multivariate Gaussian conditional rule on the sparse GP prior formulation and $q(f_s)$ is defined as a multivariate Gaussian distribution. The expressions for these terms are shown in equations 2.23 and 2.24 respectively.

$$\begin{aligned} p(f|f_s) &= \mathcal{N}(K_{XS}\ K_{SS}^{-1}\ f_s,\ K - K_{XS}\ K_{SS}^{-1}\ K_{XS}^T) \\ &= \mathcal{N}(f; Af_s,\ B) \end{aligned} \tag{2.23}$$

where $A = K_{XS}\ K_{SS}^{-1}$ and $B = K -\ K_{XS}\ K_{SS}^{-1}\ K_{XS}^T$

$$q(f_s) = \mathcal{N}(f_s; \mu,\ \Sigma) \tag{2.24}$$

Since the marginal likelihood, $p(y)$, serves as a metric for assessing how effectively the SVGP model explains the training data, it has to be maximized with respect to the model parameters. For computational convenience, the most widely used approach is to maximize the logarithm of the marginal likelihood, $\log p(y)$. Since directly computing $\log p(y)$ is not straightforward due to computational complexities, variational inference employs the Evidence Lower Bound (ELBO) as an alternative maximization objective. The Evidence Lower Bound (ELBO) is a fundamental concept in variational inference, particularly in the context of probabilistic modeling. It serves as a lower bound on the log marginal likelihood and is crucial for approximating complex posterior distributions. By maximizing the ELBO with respect to the model parameters, practitioners iteratively refine a variational distribution to closely match the

true posterior, enabling efficient inference in models with latent variables, such as Gaussian processes and other Bayesian frameworks. The derivation of the ELBO is presented below.

$$\log(p(y)) = \log \int \int p(y|f, f_s)\ p(f, f_s)\ df\ df_s \tag{2.25}$$

By dropping f_s where the context is clear and introducing the distribution $q(f, f_s)$ results in the following expression.

$$\log(p(y)) = \log \int \int p(y|f)\ p(f, f_s)\ \frac{q(f, f_s)}{q(f, f_s)} df\ df_s \tag{2.26}$$

Re-writing the above expression in terms of Expectation form gives

$$\log(p(y)) = \log \mathbb{E}_{f, f_s \sim q(f, f_s)} \left[p(y|f) \frac{p(f, f_s)}{q(f, f_s)} \right] \tag{2.27}$$

Applying Jensen's inequality to the above expression results in the following form

$$\log(p(y)) >= \mathbb{E}_{f, f_s \sim q(f, f_s)} \left[\log \left(p(y|f) \frac{p(f, f_s)}{q(f, f_s)} \right) \right] \tag{2.28}$$

Re-writing it back to the integral form

$$\log(p(y)) = \int \int \log(p(y|f)) q(f, f_s) df df_s - \int \int \log \left(\frac{q(f, f_s)}{p(f, f_s)} \right) q(f, f_s) df df_s \tag{2.29}$$

The second term in the above expression is the Kullback–Leibler divergence (KL divergence) between the distributions $q(f, f_s)$ and $p(f, f_s)$, and the first term is referred to as the likelihood term. Hence the above expression can be slightly modified and expressed as:

$$\log(p(y)) = \int \int \log(p(y|f)) q(f, f_s) df df_s - KL(q(f, f_s)||p(f, f_s)) \tag{2.30}$$

The expressions for the KL divergence term and the likelihood term can be derived as follows. For the KL divergence term,

$$KL(q(f, f_s)||p(f, f_s)) = \int \int \log \left(\frac{q(f, f_s)}{p(f, f_s)} \right) q(f, f_s) df df_s \tag{2.31}$$

Using the definition of $q(f, f_s)$ from equation 2.22 , the above expression can be written as:

$$KL(q(f, f_s)||p(f, f_s)) = \int \int \log \left(\frac{p(f|f_s)\ q(f_s)}{p(f, f_s)} \right) q(f, f_s) df df_s \tag{2.32}$$

Using the reverse chain rule to split the GP prior term $p(f, f_s)$ and canceling out the common term in the numerator and denominator gives:

$$KL(q(f, f_s)||p(f, f_s)) = \int \int \log \left(\frac{q(f_s)}{p(f_s)} \right) q(f, f_s) df df_s \tag{2.33}$$

Further re-organizing the terms and computing the expression of marginal $q(f_s)$ and substituting it results in the following expression.

$$KL(q(f, f_s)||p(f, f_s)) = \int \log \left(\frac{q(f_s)}{p(f_s)} \right) \left(\int q(f, f_s) df \right) df_s \tag{2.34}$$

$$= \int \log \left(\frac{q(f_s)}{p(f_s)} \right)\ q(f_s)\ df_s$$

$$= KL(q(f_s)||p(f_s)$$

The final expression is new KL divergence between the marginal variational distribution $q(f_s)$ and the marginal GP prior $p(f_s)$. Since both $q(f_s)$ and $p(f_s)$ are multivariate Gaussian distributions, the KL divergence is analytical, and the final expression for the KL divergence term is :

$$\mathrm{KL}(q(f_s)||p(f_s)) = \frac{1}{2} \left[\log \left(\frac{\det(K_{ss})}{\det(\Sigma)} \right) - n_s + \mathrm{tr}(K_{ss}^{-1}\Sigma) + (0 - \mu)^T K_{ss}^{-1}(0 - \mu) \right] \tag{2.35}$$

where *det* is the matrix determinant operator, *tr* is the trace operator. n_s is the length of the random variable vector f_s, or equivalently, the number of inducing locations.

Now for the likelihood term,

$$\int \int \log(p(y|f)) q(f, f_s) df df_s = \int \log(p(y|f)) \left(\int q(f, f_s) df_s \right)\ df \tag{2.36}$$

Integrating out f_s by computing the integration in the parenthesis into the marginal $q(f)$ results in the following expression.

$$\int \log(p(y|f)) \left(\int q(f, f_s) df_s \right) \, df = \int \log(p(y|f)) q(f) \, df \tag{2.37}$$

To solve the above integral and obtain the final expression for the likelihood term, the expression for $q(f)$ has to be derived. The marginal distribution $q(f)$ is obtained by integrating out the variable f_s from the joint distribution $q(f, f_s)$. However, it should be noted that $q(f)$ is a distribution of the random variable vector f and not the marginal variational distribution $q(f_s)$. The multivariate Gaussian marginalization rule cannot be applied to simply read off the marginal distribution q(f) from the joint distribution $q(f, f_s)$. This is because the variational distribution $q(f, f_s)$ is not defined as a multivariate Gaussian distribution; it is defined as the factorization $p(f|f_s)\ q(f_s)$.

$$q(f) = \int q(f, f_s) \, df_s \tag{2.38}$$

Using the definition of $q(f, f_s)$ and then substituting the terms with corresponding distribution gives,

$$\begin{aligned} q(f) &= \int p(f|f_s)\ q(f_s)\ df_s \\ &= \int \mathcal{N}(f; Af_s, B)\ \cdot\ \mathcal{N}(f_s; \mu, \Sigma)\ df_s \end{aligned} \tag{2.39}$$

Applying the multivariate Gaussian linear transformation rule gives:

$$\begin{aligned} q(f) &= \int \mathcal{N}(f; A\mu, A\Sigma A^T + B) \cdot \mathcal{N}(f_s; \mu, \Sigma)\ df_s \\ &= \mathcal{N}(f; A\mu, A\Sigma A^T + B) \end{aligned} \tag{2.40}$$

Substituting this in the expression for likelihood in the equation 2.37 gives

$$\int \log(p(y|f)) q(f) \, df = \int \log p(y|f)) \cdot \mathcal{N}(f; A\mu, A\Sigma A^T + B) \, df \tag{2.41}$$

Substitute the Gaussian distribution $p(y|f)$:

$$\int \log\left(\frac{1}{\sqrt{(2\pi\eta^2)^k}}\exp\left(-\frac{1}{2\eta^2}\|y-f\|^2\right)\right)\cdot\mathcal{N}(f;A\mu,A\Sigma A^T+B)\,df$$

Now using the Gaussian quadrature rule, the final expression for the likelihood term is :

$$\begin{aligned}\textit{Likelihood term} = &-\frac{1}{2\eta^2}\left(y^Ty-2\left(A\mu\right)^Ty+\operatorname{tr}\left(A\Sigma A^T+B\right)\right) \\ &-\frac{k}{2}\log(2\pi\eta^2)+\frac{1}{4}(f-A\mu)^T(A\Sigma A^T+B)^{-1}(f-A\mu)+\frac{k}{4}\log(2\pi) \\ &+\frac{1}{4}\log(\det(A\Sigma A^T+B))\end{aligned} \tag{2.42}$$

The expressions for the likelihood (2.42) and KL divergence term (2.35) in the ELBO have been derived and hence ELBO can be maximized to obtain the optimal parameters.

2.6 Proposed Methodology

Nonlinear Causal inference using a regular Gaussian process suffers from the issue of spurious causations and computational complexities associated with the inversion of the covariance matrix. Hence in this work, a sparse causal inference using a spare variational Gaussian process is proposed. Sparse causal inference to overcome the spurious relations is carried out using a sensitivity analysis. For a linear system, this is straightforward by placing an L1 norm on the coefficient matrix obtained by the partial differentiation as shown below:

Letting X_t denote the measured variables at time t, and X_{td} represent the past of X with a time delay d, then

$$\begin{aligned}x_t &= M\cdot x_{t_d} \\ M &= \frac{\partial x_t}{\partial x_{t_d}} \\ \|M\|_1 &= \left\|\frac{\partial x_t}{\partial x_{t_d}}\right\|_1\end{aligned} \tag{2.43}$$

But for the nonlinear system under consideration, it is not straightforward as shown above since it involves dealing with probability distributions. The derivation for the sensitivity term that ensures sparse solution is detailed below:

Consider the system: $x_t = f(x_{td}) + e_t$ where X_t denotes the measured variables at time t, X_{td} represents the past of X with a time delay d, e is Gaussian white noise and f is modeled as a Gaussian process.

The probability density function (PDF) for multivariate Gaussian distribution is:

$$p(x_t) = \frac{1}{(2\pi)^{D/2}|K|^{1/2}} \quad exp(\frac{-1}{2}(x_t - \mu)K^{-1}(x_t - \mu)) \tag{2.44}$$

Here μ and K are the mean and covariance obtained for the prediction equation based on inducing points as discussed and derived in section 2.5. Let α be the pseudo input, β be the pseudo output and x^* be the new test input in this case. Then the expressions for μ and K become: $\mu = K_x K_M^{-1}\beta$ and $K = K_{xx} - K_x K_M^{-1} K_x^T + \sigma^2$ where $K_{xx} = K(x^*, x^*)$, $K_x = K(\alpha, x^*)$, $K_M = K(\alpha, \alpha)$.

Now taking *log* on both sides,

$$log\ p(x_t) = \frac{-D}{2}log(2\pi) - \frac{1}{2}log|K| - \frac{1}{2K}(x_t - \mu)^T(x_t - \mu) \tag{2.45}$$

Expanding the last term and rewriting the above expression gives:

$$log\ p(x_t) = \frac{-D}{2}log(2\pi) - \frac{1}{2}log|K| - \frac{1}{2K}(x_t^T x_t - x_t^T\mu - \mu^T x_t + \mu^T\mu) \tag{2.46}$$

Hence the sensitivity term that needs to be evaluated is

$$S = \frac{\partial \log p(x_t)}{\partial x_{td}} \tag{2.47}$$

The first term in equation 2.45 is a constant and hence the derivative becomes 0. The following part of the derivation shows the partial derivative of each term in equation 2.45 one by one. The kernel function used in this study is the Radial basis function (RBF) kernel.

For the second term $\frac{-1}{2}log|K|$:

$$\frac{\partial}{\partial x_{td}}(\frac{-1}{2}\ log\ |K|) = \frac{-1}{2K}\frac{\partial}{\partial x_{td}}(|K|) \tag{2.48}$$

We have to find $\frac{\partial K}{\partial x_{td}}$. Now this is a partial differentiation of a scalar with respect to a vector. But when we consider 1 sample at a time, it becomes a partial derivative of a scalar wrt a scalar.

$$\frac{\partial K}{\partial x^i_{td}} = \frac{\partial K_{xx}}{\partial x^i_{td}} - \frac{\partial K_x K_M^{-1} K_x^T}{\partial x^i_{td}} \tag{2.49}$$

$$\frac{\partial K}{\partial x^i_{td}} = -\frac{\partial K_x K_M^{-1} K_x^T}{\partial x^i_{td}} \tag{2.50}$$

Applying the Chain rule,

$$\frac{\partial K}{\partial x^i_{td}} = -\frac{\partial K_x^T K_M^{-1} K_x}{\partial x^i_{td}} = -\frac{\partial K_x}{\partial x^i_{td}}\ K_M^{-1} K_x^T - K_x K_M^{-1}\ \frac{\partial K_x^T}{\partial x^i_{td}} \tag{2.51}$$

In the above expression we need to find $\frac{\partial K_x}{\partial x^i_{td}}$

$$K_x = \sigma^2\ exp(\frac{-1}{2l^2}\ (x^i_{td} - \alpha^i)^T(x^i_{td} - \alpha^i)) \tag{2.52}$$

$$K_x = \sigma^2\ exp(\frac{-1}{2l^2}\ (x^{iT}_{td}x^i_{td} - 2x^i_{td}\alpha^{iT} + \alpha^{iT}\alpha^{iT})) \tag{2.53}$$

$$\frac{\partial K_x}{\partial x^i_{td}} = \frac{\partial}{\partial x^i_{td}}\ (\sigma^2\ exp(\frac{-1}{2l^2}\ (x^{iT}_{td}x^i_{td} - 2x^i_{td}\alpha^{iT} + \alpha^{iT}\alpha^{iT}))) \tag{2.54}$$

Here,

$$\begin{aligned}\frac{\partial x^{iT}_{td}x^i_{td}}{\partial x^i_{td}} &= \frac{\partial}{\partial x^i_{td}}\sum_{j=1}^{p}(x^i_{td})^2 \\ &= 2x^i_{td} \\ \frac{\partial x^i_{td}\alpha^i}{\partial x^i_{td}} &= \alpha^i\end{aligned} \tag{2.55}$$

$$\frac{\partial K_x}{\partial x^i_{td}} = \sigma^2\ exp(\frac{-1}{2l^2}\ (x^{iT}_{td}x^i_{td} - 2x^i_{td}\alpha^{iT} + \alpha^{iT}\alpha^{iT})) * \frac{-1}{2l^2}(2x^i_{td} - 2\alpha^i) \tag{2.56}$$

Hence,

$$\begin{aligned}\frac{\partial K}{\partial x^i_{td}} = [\sigma^2\ exp(\frac{-1}{2l^2}\ (x^{iT}_{td}x^i_{td} - 2x^i_{td}\alpha^{iT} + \alpha^{iT}\alpha^{iT})) * \frac{1}{2l^2}(2x^i_{td} - 2\alpha^i)\ K_M^{-1}K_x^T + \\ K_x K_M^{-1}\sigma^2\ exp(\frac{-1}{2l^2}\ (x^{iT}_{td}x^i_{td} - 2x^i_{td}\alpha^{iT} + \alpha^{iT}\alpha^{iT})) * \frac{1}{2l^2}(2x^i_{td} - 2\alpha^i)]\end{aligned} \tag{2.57}$$

Substituting the above expression in equation 2.51 gives the final expression for term 1.

$$\frac{\partial}{\partial x_{td}}(\frac{-1}{2}\ log\ |K|) = \frac{-1}{2|K|}[\sigma^2\ exp(\frac{-1}{2l^2}\ (x_{td}^{iT}x_{td}^{i}-2x_{td}^{i}\alpha^{iT}+\alpha^{iT}\alpha^{iT}))*\frac{1}{2l^2}(2x_{td}^{i}-2\alpha^{i})\ K_M^{-1}K_x^T+$$
$$K_xK_M^{-1}\sigma^2\ exp(\frac{-1}{2l^2}\ (x_{td}^{iT}x_{td}^{i}-2x_{td}^{i}\alpha^{iT}+\alpha^{iT}\alpha^{iT}))*\frac{1}{2l^2}(2x_{td}^{i}-2\alpha^{i})] \quad (2.58)$$

This is the final expression for term 2.

For the third term: $\frac{\partial}{\partial x_{td}}(\frac{-1}{2}x_t^TK^{-1}x_t)$

Applying the chain rule gives,

$$\frac{\partial}{\partial x_{td}^{i}}(\frac{-1}{2}x_t^{iT}K^{-1}x_t^{i}) = \frac{-1}{2}x_t^{iT}\ \frac{\partial K^{-1}}{\partial x_{td}^{i}}\ x_t^{i} \quad (2.59)$$

Here we have to find $\frac{\partial K^{-1}}{\partial x_{td}^{i}}$

Let $I = KK^{-1}$. Taking derivative on both sides gives,

$0 = (KK^{-1})^1$

$0 = K^1K^{-1} + K(K^{-1})^1$

$$\frac{\partial K^{-1}}{\partial x_{td}^{i}} = -K^{-1}\frac{\partial K}{\partial x_{td}^{i}}K^{-1} \quad (2.60)$$

The expression for $\frac{\partial K}{\partial x_{td}^{i}}$ has already been found in equation 2.57. Substituting equation 2.60 in equation 2.59 gives

$$\frac{\partial}{\partial x_{td}^{i}}(\frac{-1}{2}x_t^{iT}K^{-1}x_t^{i}) = \frac{1}{2}x_t^{iT}K^{-1}\frac{\partial K}{\partial x_{td}^{i}}K^{-1}x_t \quad (2.61)$$

Substituting equation 2.57 in equation 2.61 gives:

$$\frac{\partial}{\partial x_{td}^{i}}(\frac{-1}{2}x_t^{iT}K^{-1}x_t^{i}) = \frac{1}{2}x_t^{iT}K^{-1}[\sigma^2\ exp(\frac{-1}{2l^2}\ (x_{td}^{iT}x_{td}^{i}-2x_{td}^{i}\alpha^{iT}+\alpha^{iT}\alpha^{iT}))*$$
$$\frac{1}{2l^2}(2x_{td}^{i}-2\alpha^{i})\ K_M^{-1}K_x^T \quad +$$
$$K_xK_M^{-1}\sigma^2\ exp(\frac{-1}{2l^2}\ (x_{td}^{iT}x_{td}^{i}-2x_{td}^{i}\alpha^{iT}+\alpha^{iT}\alpha^{iT}))*\frac{1}{2l^2}(2x_{td}^{i}-2\alpha^{i})]K^{-1}x_t^{i} \quad (2.62)$$

This is the final expression for term 3.

For the fourth term: $\frac{\partial}{\partial x_{td}^i}(x_t^{iT}K^{-1}\mu)$

Applying the chain rule gives,

$$\frac{\partial}{\partial x_{td}^i}(x_t^{iT}K^{-1}\mu) = x_t^T\frac{\partial K^{-1}}{\partial x_{td}^i}\mu \quad + x_t^{iT}K^{-1}\frac{\partial \mu}{\partial x_{td}^i} \tag{2.63}$$

Here we need to calculate $\frac{\partial \mu}{\partial x_{td}^i}$. Applying the chain rule gives,

$$\frac{\partial \mu}{\partial x_{td}^i} = \frac{\partial}{\partial x_{td}^i}(K_x K_M^{-1}\beta) \tag{2.64}$$

$$\frac{\partial \mu}{\partial x_{td}^i} = \frac{\partial K_x}{\partial x_{td}^i}K_M^{-1}\beta \tag{2.65}$$

Therefore substituting equation 2.64 and equation 2.60 in equation 2.63 gives

$$\frac{\partial}{\partial x_{td}^i}(x_t^{iT}K^{-1}\mu) = -x_t^{iT}K^{-1}\frac{\partial K}{\partial x_{td}^i}K^{-1}\mu + x_t^{iT}K^{-1}\frac{\partial K_x}{\partial x_{td}^i}K_M^{-1}\beta \tag{2.66}$$

Substituting equation 2.57 and equation 2.56 in equation 2.66 gives

$$\begin{aligned}\frac{\partial}{\partial x_{td}^i}(x_t^{iT}K^{-1}\mu) = -x_t^{iT}K^{-1}[\sigma^2\, exp(\frac{-1}{2l^2}\,(x_{td}^{iT}x_{td}^i - 2x_{td}^i\alpha^{iT} + \alpha^{iT}\alpha^{iT}))*\frac{1}{2l^2}(2x_{td}^i - 2\alpha^i)\, K_M^{-1}K_x^T + \\ K_xK_M^{-1}\sigma^2\, exp(\frac{-1}{2l^2}\,(x_{td}^{iT}x_{td}^i - 2x_{td}^i\alpha^{iT} + \alpha^{iT}\alpha^{iT}))*\frac{1}{2l^2}(2x_{td}^i - 2\alpha^i)]K^{-1}\mu \quad + \\ x_t^{iT}K^{-1}[\sigma^2\, exp(\frac{-1}{2l^2}\,(x_{td}^{iT}x_{td}^i - 2x_{td}^i\alpha^{iT} + \alpha^{iT}\alpha^{iT}))*\frac{-1}{2l^2}(2x_{td}^i - 2\alpha^i)]K_M^{-1}\beta\end{aligned} \tag{2.67}$$

This is the final expression for the fourth term.

For the fifth term: $\frac{\partial}{\partial x_{td}^i}(\mu_t^TK^{-1}\mu)$

Here μ is a scalar for a single sample. Hence,

$$\begin{aligned}\frac{\partial}{\partial x_{td}^i}(\mu_t^TK^{-1}\mu) &= \frac{\partial}{\partial x_{td}^i}(\mu^i)^2K^{-1} \\ &= \frac{\partial}{\partial x_{td}^i}((K_xK_M^{-1}\beta)^2K^{-1})\end{aligned} \tag{2.68}$$

Applying the chain rule gives,

$$\frac{\partial}{\partial x_{td}^i}(\mu_t^TK^{-1}\mu) = (K_xK_M^{-1}\beta)^2\frac{\partial K^{-1}}{\partial x_{td}^i} + \quad K^{-1}*2(K_xK_M^{-1}\beta)\frac{\partial K_x}{\partial x_{td}^i}K_M^{-1}\beta \tag{2.69}$$

$$\frac{\partial}{\partial x_{td}^i}(\mu_t^TK^{-1}\mu) = -(K_xK_M^{-1}\beta)^2K^{-1}\frac{\partial K}{\partial x_{td}^i}K^{-1} + K^{-1}*2(K_xK_M^{-1}\beta)\frac{\partial K_x}{\partial x_{td}^i}K_M^{-1}\beta \tag{2.70}$$

Substituting equation 2.56 and equation 2.57 in equation 2.69 gives

$$\frac{\partial}{\partial x_{td}^i}(\mu_t^T K^{-1}\mu) = -(K_x K_M^{-1}\beta)^2 K^{-1}[\sigma^2\ exp(\frac{-1}{2l^2}\ (x_{td}^{iT}x_{td}^i - 2x_{td}^i\alpha^{iT} + \alpha^{iT}\alpha^{iT}))* \frac{1}{2l^2}(2x_{td}^i - 2\alpha^i)\ K_M^{-1}K_x^T \quad + K_x K_M^{-1}\sigma^2\ exp(\frac{-1}{2l^2}\ (x_{td}^{iT}x_{td}^i - 2x_{td}^i\alpha^{iT} + \alpha^{iT}\alpha^{iT})) * \frac{1}{2l^2}(2x_{td}^i - 2\alpha^i)]K^{-1} \quad + K^{-1} * 2(K_x K_M^{-1}\beta)[\sigma^2\ exp(\frac{-1}{2l^2}\ (x_{td}^{iT}x_{td}^i - 2x_{td}^i\alpha^{iT} + \alpha^{iT}\alpha^{iT})) * \frac{-1}{2l^2}(2x_{td}^i - 2\alpha^i)]K_M^{-1}\beta \tag{2.71}$$

This is the final expression for the fifth term.

Hence the expression of the Sensitivity term is :

$$S = \frac{\partial \log p(x_t)}{\partial x_{td}} = Term\ 1 + Term\ 2 + Term\ 3 + Term\ 4 + Term\ 5 \tag{2.72}$$

Now we have obtained the expressions for the ELBO term as well as the sensitivity term. ELBO being the lower bound on the log-likelihood should be maximized and the sensitivity should be minimized for achieving the desired results. Hence the objective function for the proposed sparse nonlinear causal inference using variational Gaussian Process is:

$$\max_{\theta}\ ELBO - \lambda\ |||S||_1 \tag{2.73}$$

where the parameters are l, σ^2 (hyper-parameters of the kernel function), η^2 (observational noise), α, β (the inducing points), and μ, Σ (The mean and variance of the variational q distribution.

Once the optimal parameters are determined, causal relationships can be obtained based on the d-statistic defined in section 2.4.1 using the determined optimal parameters.

2.7 Case study

In this section, the result obtained from the simulation case study is presented demonstrating the efficacy of the proposed approach compared to traditional Granger causality and existing Gaussian process-based causal inference.

2.7.1 Simulation Case study 1

A causal map consisting of 4 variables and 3 direct causal relations is considered for the analysis in this simulation case study. The cause-effect relationships are shown in Figure 2.3. It can be observed from the causal map that $X1$ and $X3$ are the source variables and $X2$ and $X4$ are the effect variables. The nonlinear equations used for the data generation of the variables are as follows:

$$
\begin{aligned}
x_1(t) &= log(1 + [2x_1(t-1) - x_1(t-2)]^2) + e_1(t) \\
x_2(t) &= \frac{1}{(x_1(t-1)+5)^2} + 2x_3(t-2)^2 + e_2(t) \\
x_3(t) &= 5exp(-0.1x_3(t-2)^2 - 4x_3(t-1)^2) + e_3(t) \\
x_4(t) &= \left(\frac{x_4(t-1}{1 + x_2(t-1)^2} + 5\right) sin(x_2(t-2)) + e_4(t)
\end{aligned}
\tag{2.74}
$$

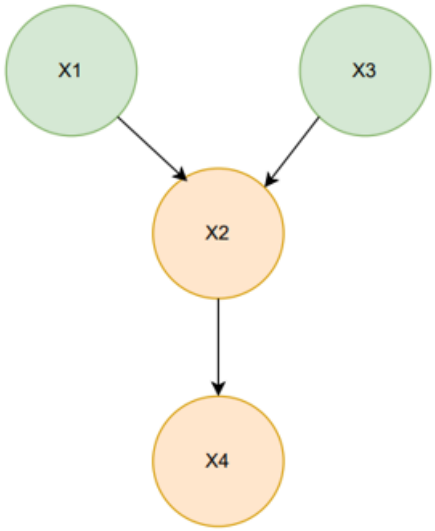
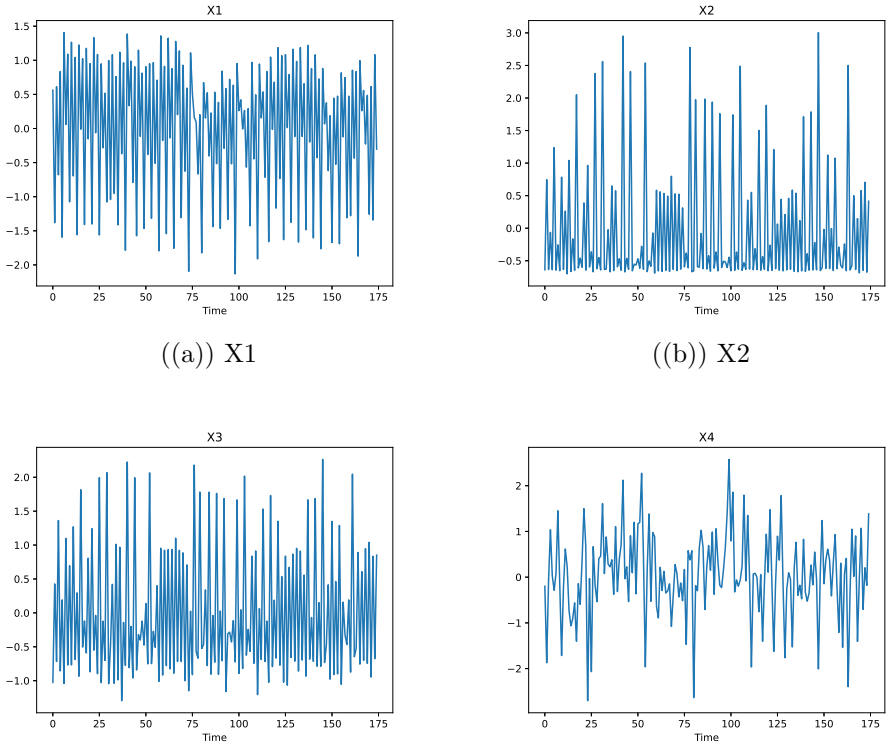


Figure 2.3: True Causal map

The covariance for noise $e_{1:4}(t)$ is taken as $\text{diag}(0.05, 0.05, 0.05, 0.05)$. A total of 1000 samples were generated using the above equation. The corresponding plots of all the variables are shown in Figure 2.4. For better visualization, only the first 175 data points are shown in the plots.



((a)) X1 ((b)) X2

((c)) X3 ((d)) X4

Figure 2.4: Plots of the variables of simulation case study

First, the GC method is applied to generate the cause-effect relationship. The obtained causal map is shown in Figure 2.5

From the figure, it is evident that Granger Causality fails to identify the underlying true causal relations in a nonlinear system, and in addition to that, GC gave four spurious causations (indicated by dashed arrows). These are as follows: $X3$ causing $X4$, $X1$ causing $X4$, $X1$ causing $X3$ and $X3$ causing $X1$. This illustrates that Granger causality cannot accurately reconstruct the true causal map when we have a nonlinear relationship between the variables in the system. SGC was then applied to the data and the obtained causal relations were exactly as similar to the results of Granger causality. The spurious causations were so strong that they were not removed. This further verifies the fact that Granger causality does not perform well in a nonlinear framework as it fails to identify the true causal relations and also results in spurious causal links.

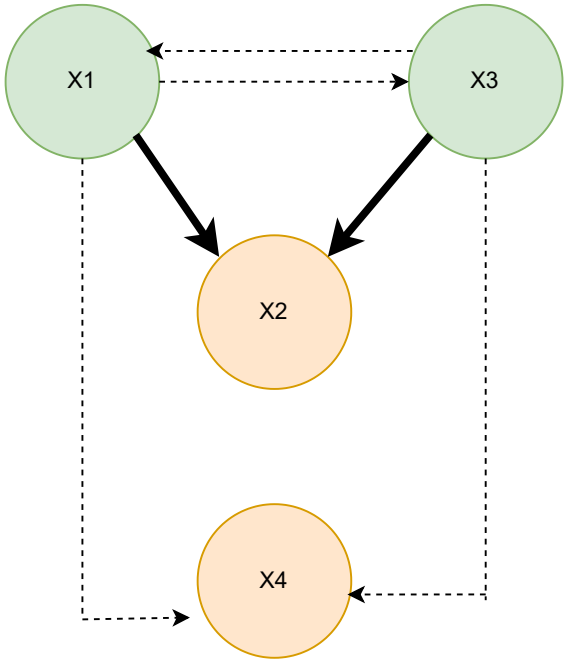


Figure 2.5: Causal map obtained by Granger Causality

Gaussian process-based causal inference was then performed on the data and the obtained causal map is shown in Figure 2.6. It can be observed from the figure that the GP-based approach was able to identify all the true underlying causal relationships within the data, unlike the Granger-based approaches. But as discussed before, GP-based approaches tend to produce a lot of spurious causations and it is evident from the causal map. The GP-based approach gave 6 spurious causation. These are as follows: $X2$ causing $X1$, $X2$ causing $X3$, $X4$ causing $X2$ and $X4$ causing $X1$, $X4$ causing $X3$, and $X3$ causing $X4$.

The proposed nonlinear sparse causal inference based on the variational Gaussian process was then implemented on the data set and the obtained causal map is shown in figure 2.7. It is evident from the causal map that, the proposed methodology correctly identifies all the underlying true causal relations and gives only 1 spurious causation which is a significant improvement considering the existing Gaussian process-based methodology had 6 spurious causations. The Confusion Matrix in Figure 2.8 summarizes the results of this case study.

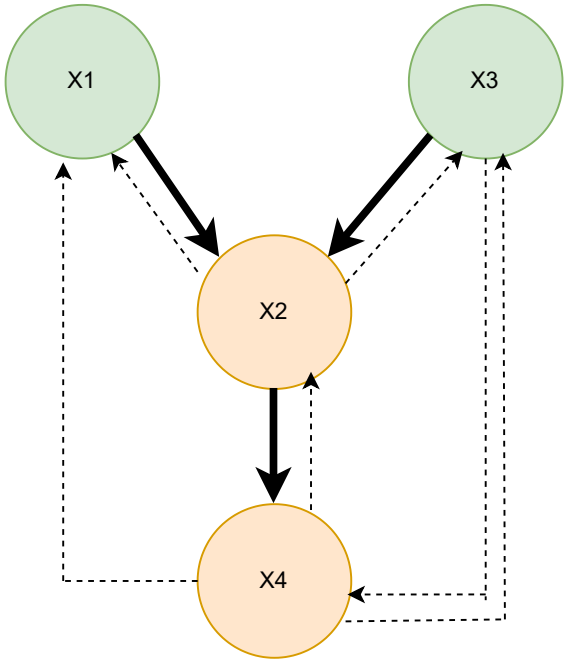


Figure 2.6: Gaussian process-based causal map

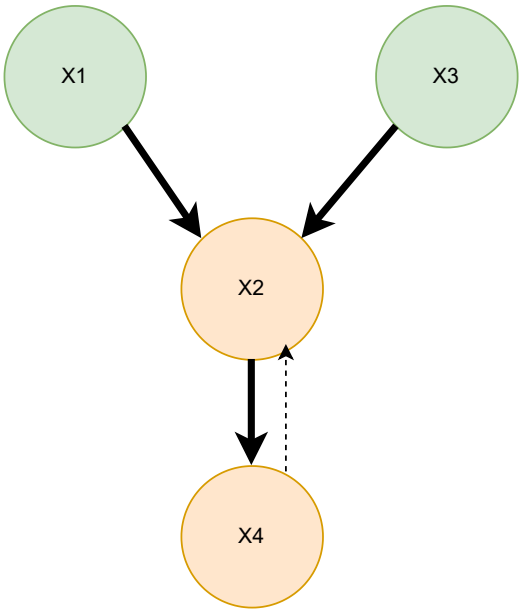
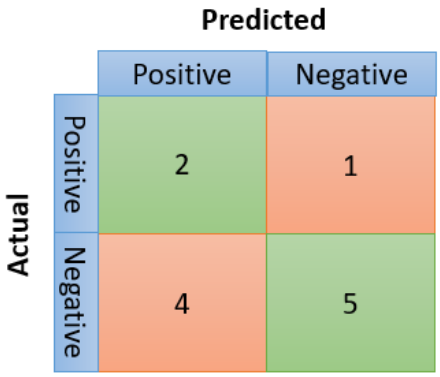


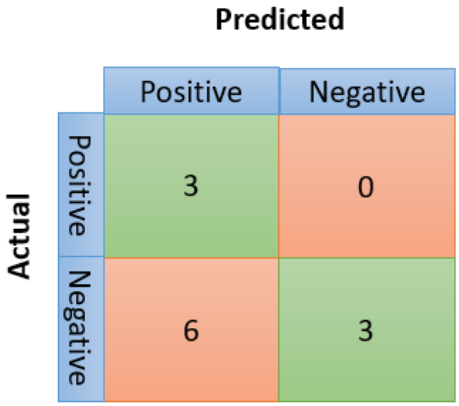
Figure 2.7: Proposed sparse nonlinear variational Gaussian process-based causal map

2.7.2 Simulation Case study 2

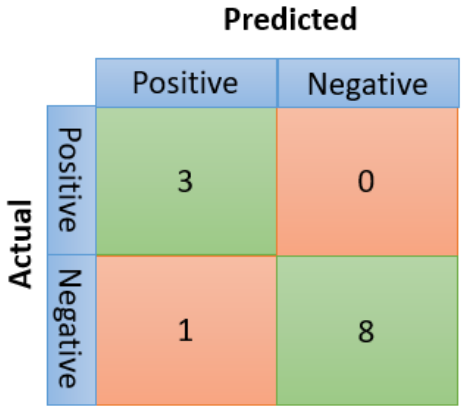
A causal map consisting of 5 variables and 6 direct causal relations is considered for the analysis in this simulation case study. The data generation for this case study is



((a)) Granger Causality



((b)) Gaussian process-based causality



((c)) Proposed method

Figure 2.8: Confusion matrix obtained for the simulation case study 1. In each figure, each block represents true positives, false negatives, true negatives, and false positives (clockwise starting from the top left). Positive and negative imply the existence and non-existence of causal relations respectively.

based on the study done by Wismüller et al. in their work towards nonlinear causal inference [38]. The data generation process presented in their work has been modified here to make the causal map more complex. The cause-effect relationships are shown in Figure 2.9. It can be observed from the causal map that $X1$ is the source variable. The nonlinear equations used for the data generation of the variables are as follows:

$$x_1(t) = 0.95\sqrt{10 + 2x_1(t-1)} - 0.9025x_1(t-2) + e_1(t) \tag{2.75}$$

$$x_2(t) = 0.5x1(t-2)^2 + e_2(t)$$

$$x_3(t) = -0.4x1(t-3) + e_3(t)$$

$$x_4(t) = -0.5x1(t-2)^2 + 0.5\sqrt{10 + 2x_4(t-1)} + 0.250.95\sqrt{10 + 2x_5(t-1)} + e_4(t)$$

$$x_5(t) = -0.5\sqrt{20 + 2x_4(t-1)} + 0.5\sqrt{20 + 2x_5(t-1)} + e_5(t)$$

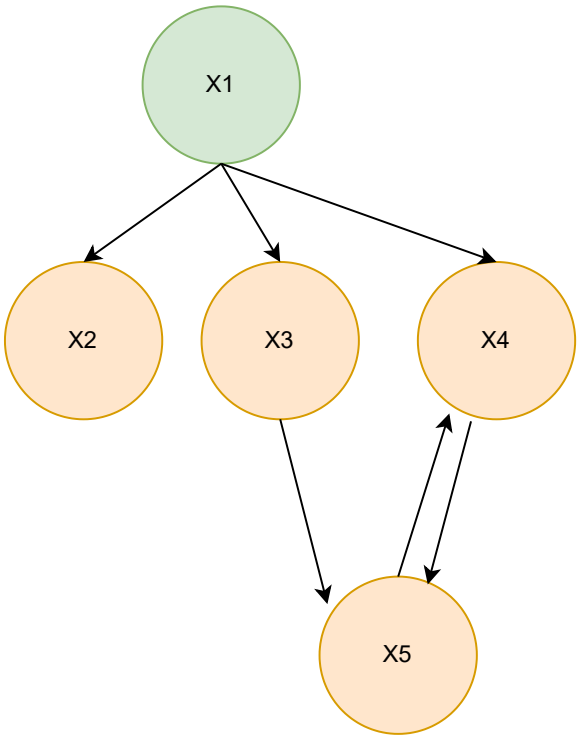
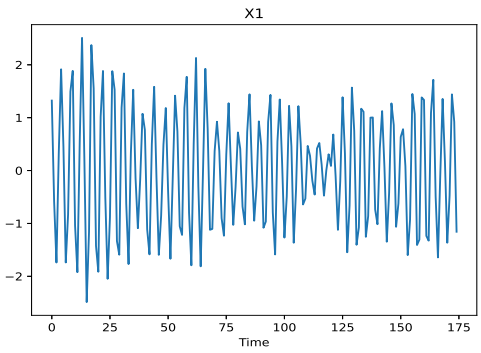


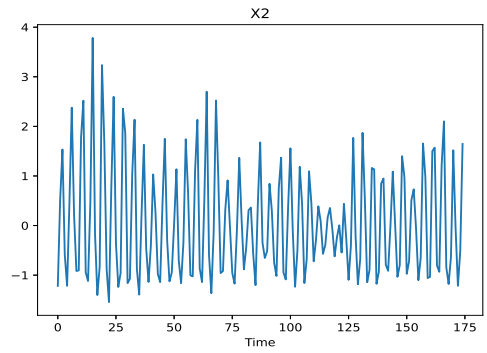
Figure 2.9: Actual causal map for simulation case study 2

The covariance for noise $e_{1:4}(t)$ is taken as $\text{diag}(0.05, 0.05, 0.05, 0.05, 0.05)$. A total of 1000 samples were generated using the above equation. The corresponding plots of all the variables are shown in Figure 2.10. For better visualization, only the first 175 data points are shown in the plots.

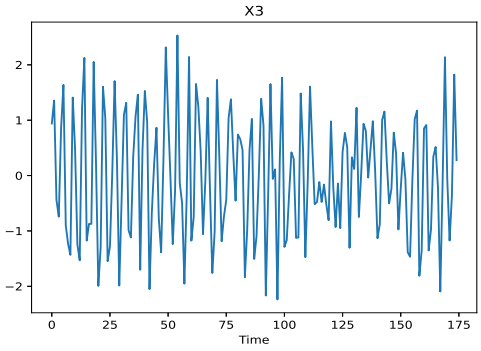
First, the GC method is applied to generate the cause-effect relationship. The obtained causal map is shown in Figure 2.11



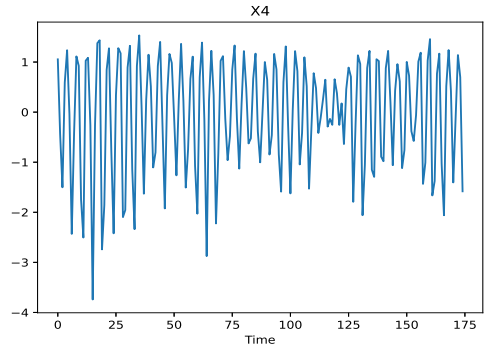
((a)) X1



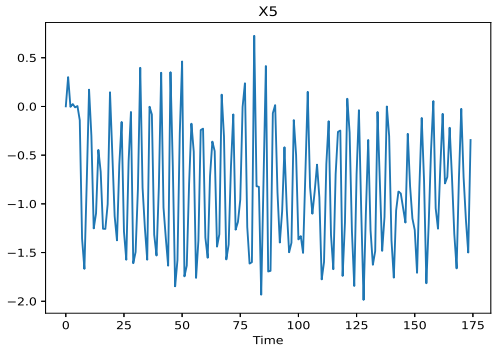
((b)) X2



((c)) X3



((d)) X4



((e)) X5

Figure 2.10: Plots of variables of simulation case study 2

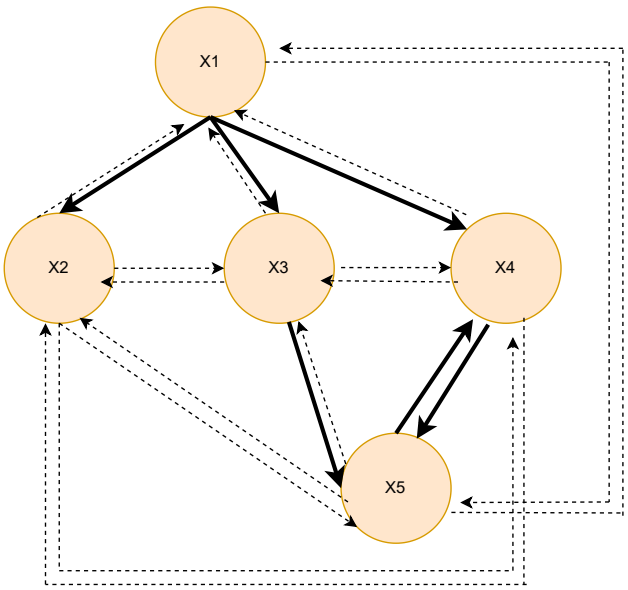


Figure 2.11: Causal map obtained by Granger Causality in Simulation 2

From the illustration, it becomes apparent that Granger Causality (GC) falls short in uncovering the genuine causal connections within a nonlinear system. Furthermore, GC erroneously establishes cause-effect relationships for all variables among themselves, leading to the emergence of spurious relations, as depicted by dashed arrows.

Gaussian process-based causal inference was subsequently applied to the dataset, and the resulting causal map is illustrated in Figure 2.12. Examination of the figure reveals that the GP-based approach successfully identified all genuine underlying causal relationships in the data. However, as previously discussed, it is noteworthy that GP-based methods often generate numerous spurious causations, as evident in the causal map. The GP-based approach gave 8 spurious causation. These are as follows: $X2$ causing $X1$, $X2$ causing $X5$, $X3$ causing $X1$, $X3$ causing $X2$, $X3$ causing $X4$, and $X4$ causing $X1$, $X4$ causing $X3$, $X5$ causing $X3$.

The nonlinear sparse causal inference method based on variational Gaussian processes was subsequently applied to the dataset, and the resulting causal map is depicted in Figure 2.13. The causal map illustrates that the proposed methodology

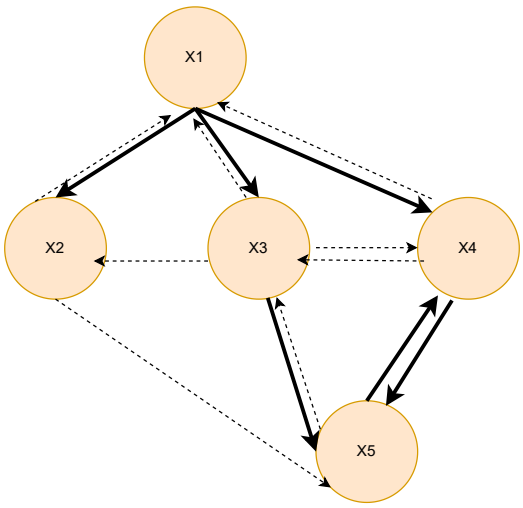


Figure 2.12: Causal map obtained by GP-based Causality in Simulation 2

accurately identifies all the true underlying causal relations, without any spurious causation. This represents a significant improvement compared to the existing Gaussian process-based methodology, which yielded eight spurious causations. The Confusion Matrix in Figure 2.14 summarizes the results of this case study.

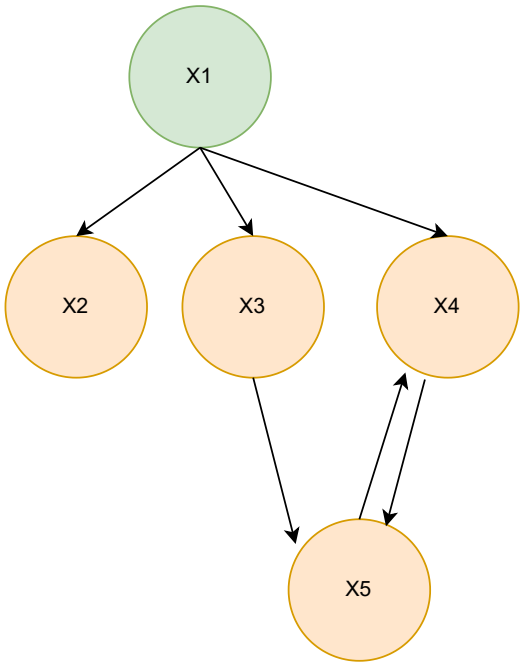
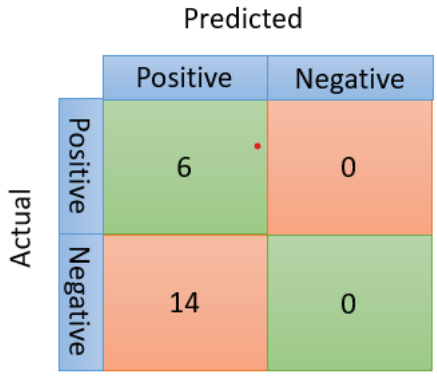
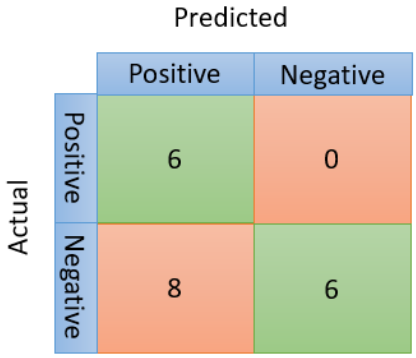


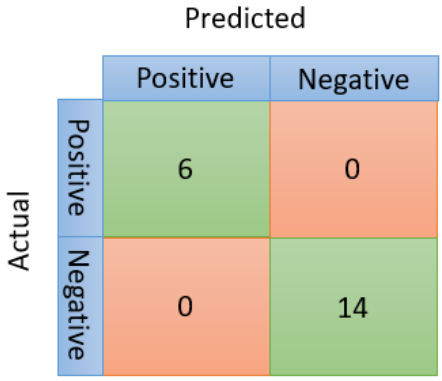
Figure 2.13: Proposed sparse nonlinear variational Gaussian process-based causal map for simulation 2



((a)) Granger Causality



((b)) Gaussian Process-based causality



((c)) Proposed Methodology

Figure 2.14: Confusion matrix obtained for the simulation case study 2. In each figure, each block represents true positives, false negatives, true negatives, and false positives (clockwise starting from the top left). Positive and negative imply the existence and non-existence of causal relations respectively.

2.7.3 Industrial Case Study

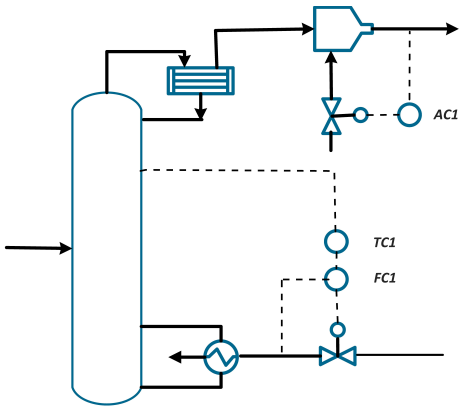


Figure 2.15: Schematic representation of the Australian refinery separation unit

The proposed method is applied to an industrial data set containing oscillations to identify the root source of these oscillations through the construction of the causal map of the system. We use the Australian refinery process data provided by Thornhill (2007) [39]. Figure 2.15 depicts the separation process considered and it contains five control loops, namely Analyzer (AC1), steam flow (FC1), temperature control (TC1), upstream pressure control (PC1), and downstream pressure control (PC2). The data for this case study has been taken during the time range when the set point remained constant and because of this, we used only the process variables (PV) and controller output variables (OP) of all the control loops and did not use the set point (SP) and controller error(SP-PV) variables. Further, the whole data is divided into the training data set and cross-validation data set to identify the optimal value of the hyper-parameter λ where the objective function is minimized for Sparse Granger causality. For this case, the number of samples in the training set and the cross-validation are taken to be 500 and 200 respectively. The dataset contains 10 measured variables. Some variables have non-stationarity in them and hence first order difference was applied to remove the non-stationary. Plots of the variables after removing non-stationary features are shown in Figure 2.16. It was identified in Thornhill (2005) [40] that a faulty steam sensor in the steam flow loop $FC1$ was the root cause of the

oscillation.

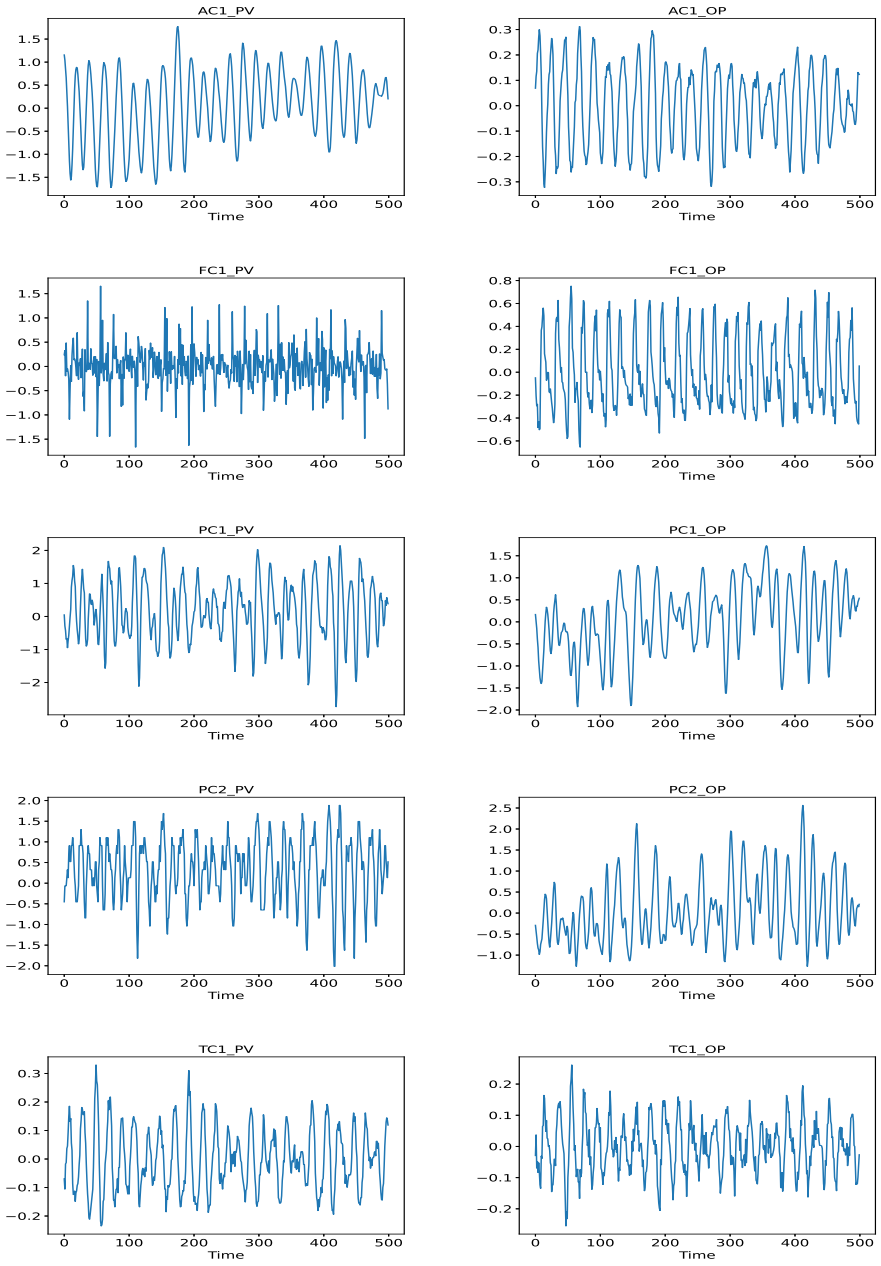


Figure 2.16: Plots of variables in the industrial case study after first order differencing

The GC approach was then implemented on the data set and the causal network was reconstructed. The obtained cause-effect network in the case of GC is shown in Figure 2.17 where it can be observed that upstream pressure (PC1) is caused by column temperature (TC1) which is an unlikely causal link. Additionally, GC also

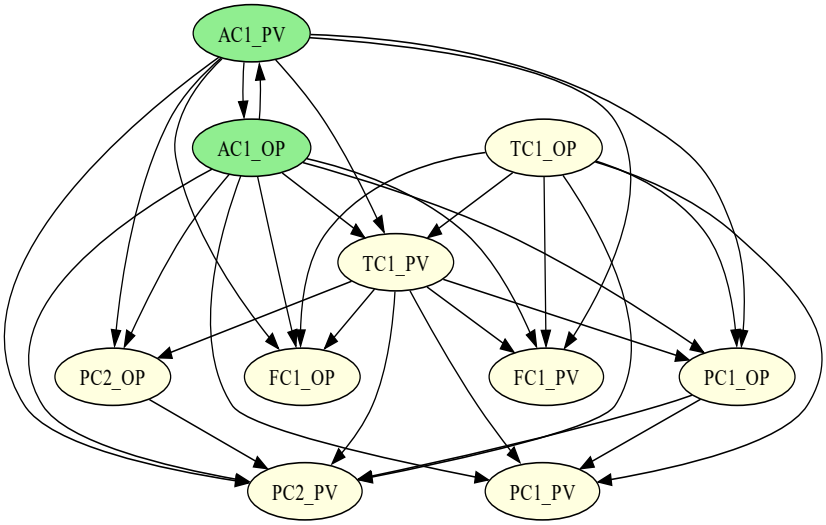


Figure 2.17: Granger causal map for industrial case study

failed to identify the true source of plant-wide oscillation. The cause-effect network obtained by GC shows no causal links originating at the $FC1$ loop and primarily identifies it as an effect rather than a cause. The source candidates identified by GC are $AC1\_OP$ and $AC1\_PV$ where both of them are not the true source.

Sparse Granger Causality was then applied to the dataset and the determined cause-effect network is shown in Figure 2.18. The hyper-parameter λ for this case study was tuned using a cross-validation dataset. This is achieved by evaluating the objective function of the optimization problem on the validation dataset and then finding the value of λ for which the objective function value is minimum. The optimal value of λ determined is 0.5. Figure 2.19 illustrates the variation of the objective function value with varying values of λ. Based on the determined cause-effect network for SGC, it is evident that the number of cause-effect relations is reduced as expected compared to GC, but it still failed to recognize the true source of plant-wide oscillation. Further, by analyzing the cause-effect network systematically,

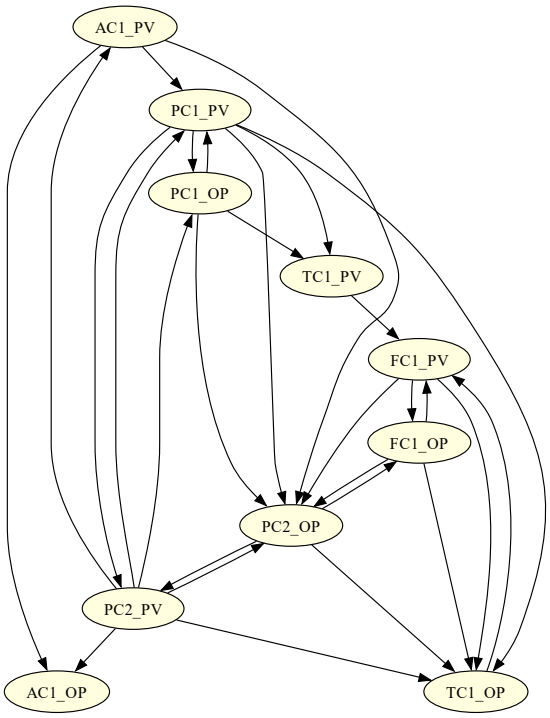


Figure 2.18: Sparse Granger causal map for industrial case study

it is observed that SGC fails to generate any distinct source in the reconstructed network. All the variables of the network have a significant number of causal links originating and departing from them. To establish a variable as a true source, a significant number of causal links need to originate from the variables compared to the number of causal links entering them. So, it can be concluded that the SGC approach failed to identify the true source of plant-wide oscillation.

Thus the linear-based approach completely fails to identify the source variable since Granger identifies the true source as the effect while Sparse Granger fails to identify any true source. Subsequently, the Gaussian Process-based nonlinear causal inference approach was applied to the data set, and the obtained causal map is shown in figure 2.20. The GP-based approach is an improvement on the linear-based approaches since it does not identify the $FC1$ loop as the effect since a significant number of arrows are originating from the $FC1$ variable but still has its limitations as it fails to identify

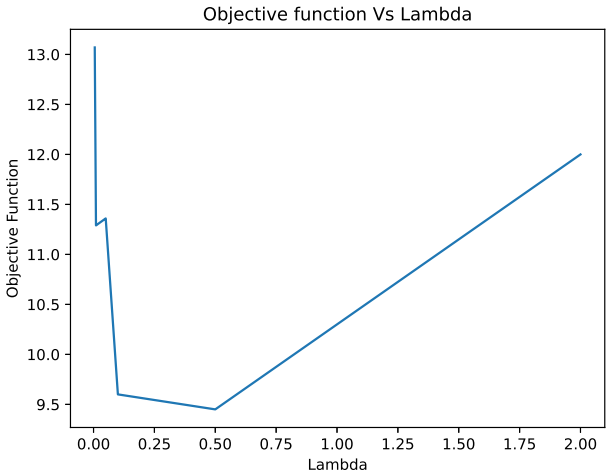


Figure 2.19: Objective function Vs Lambda for Sparse Granger Causality

any variable as the true source.

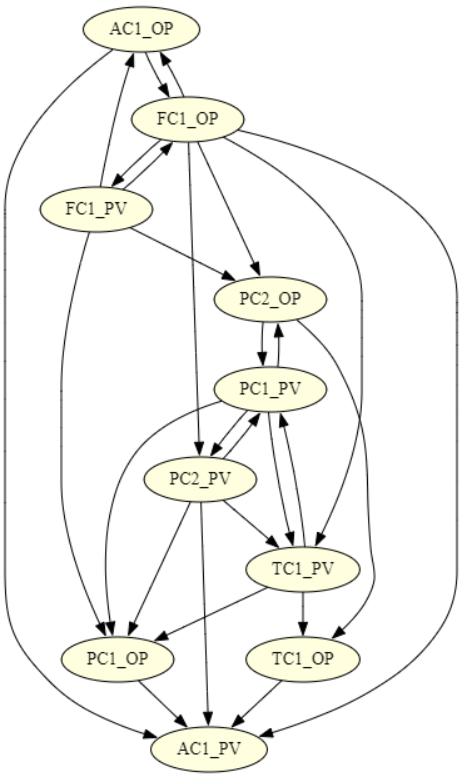


Figure 2.20: GP-based causal map for industrial case study

The nonlinear sparse causal inference method was then applied to the dataset, yielding a causal map depicted in Figure 2.21. Notably, the proposed approach re-

frains from designating any variable as an unequivocal source but highlights $AC1$ and $FC1$ as potential sources of causation. Considering the results obtained by other approaches this is a significant improvement on the existing approaches. This verifies the efficacy of the proposed approach in real industrial data. The omission of $FC1$ as the exclusive source may be attributed to the inherent challenges posed by the quality of observed industrial data, which is susceptible to errors. Additionally, analyzing the variable plots reveals inherent periodic patterns, potentially influencing the causal map. The periodic nature of certain features, recurring at regular intervals, introduces a layer of complexity to causal inference, as causative relationships may manifest intermittently. Despite these challenges, our proposed method produces significantly better results when compared to the existing approaches. The table 2.1 summarizes the results of all the approaches in this industrial case study.

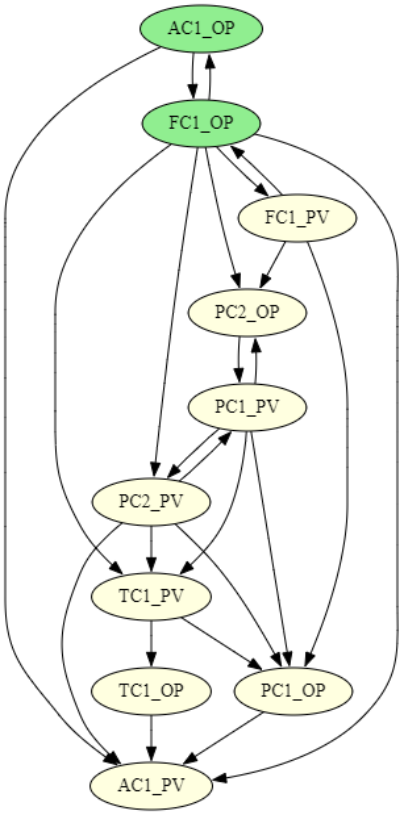


Figure 2.21: Proposed nonlinear Spare GP causal map for industrial case study

Method	Identified possible source variables
Granger Causality	$AC1\_OP$, $AC1\_PV$
Sparse Granger Causality	No distinct source variable
Gaussian-Process based Causality	No distinct source variables
Proposed Methodology	$AC1\_OP$, $FC1\_OP$

Table 2.1: Method and Identified source variables

2.8 Conclusion

This chapter proposes a novel nonlinear sparse causal inference method based on the variational Gaussian process. The proposed method addresses the shortcomings of existing approaches like Granger causality, sparse Granger causality, and Gaussian process-based causality to identify and produce the true cause-effect relations among nonlinear variables with minimum spurious relations. The simulation and industrial case study results demonstrate the proposed method's effectiveness. These promising results pave the way for further research and exploration of causal inference in nonlinear time-invariant dynamical systems.

Chapter 3

Physics-Informed Sparse Causal Inference for Source Detection of Plantwide Oscillations[1]

This Chapter presents the second contribution of the thesis which is amalgamating human knowledge with process data in causal inference for root cause analysis.

3.1 Introduction

Studies about the detection and root cause analysis of plant-wide oscillations have been extensively performed as this is a very prevalent issue in process industries. Studies carried out by Ender and Bialkoswki point to the fact that satisfactory performance was provided only by 30 percent of the operational controllers [41, 42]. This is because oscillations generated at any control loop propagate throughout the plant feedback and interconnections, resulting in a substantial reduction in controller performance in multiple loops [43].

A cross-correlation-based approach for identifying the propagation path of plant-wide disturbances was proposed [44], but correlation does not always ensure causality. Duan et al. [45] in their work discuss identifying the root cause of oscillations using data analysis in temporal and spectral domains. The most popular and well-

[1]This chapter has been published as: **N. Madhusoodanan**, R. Chiplunkar, VK Puli, B. Huang, "Physics-informed sparse causal inference for source detection of plant-wide oscillations". AIChE J. 2024; e18362. doi:10.1002/aic.18362

distinguished approach of causal inference given a set of time series data, is the prediction-based Granger Causality (GC) [31]. GC-based approaches for root cause diagnosis of plant-wide oscillations have been widely studied. Yuan et al. [46] coupled latent variable modeling for capturing oscillatory features and then applied GC for the diagnosis of the oscillatory source. Chen et al. [47] propose an approach that combines clustering oscillatory variables into various groups, and then the multivariate GC is applied to each group to identify the root cause of multiple plant-wide oscillations. Sundaramoorthy et al. [48] proposed a two-step approach where sparse inverse covariance estimation is coupled with likelihood score to address the problem of causal analysis. To deal with the time-varying nature of process systems, Raveendran et al. [49] introduced a causal modeling approach that relies on time-varying parameter models estimated under the variational Bayesian expectation-maximization framework. Raveendran and Huang [50], proposed a hybrid model by combining the factor analysis model and vector auto-regressive exogenous model to simultaneously mine causal connections and features responsible for contemporaneous correlations in a multivariate process. Deng et al. [51] proposed a kernel-based Granger approach for dealing with non-linear data and it was then further modified by Xiangzhi et al. [52] where the authors proposed a Fuzzy kernel Granger causality approach for root cause diagnosis. To deal with non-linearity and multivariate cases, Chen et al. [53] discuss integrating Gaussian Process regression (GPR) into the framework of multivariate GC. Bauer et al. [54] proposed a novel application of transfer entropy, a method based on conditional probability density functions that measures the directionality of variation for identifying the direction of disturbance propagation. Lindner et al.[55] provides a comparative study of Granger causality and Transfer entropy to present a decision flow for the application of oscillation diagnosis. Studies have also been carried out that discuss a comparative study of Granger-based causal analysis and dynamic Bayesian network inference in computational biology which specifically focuses on method selection when both of them give contradictory results [56]. Yu and

Rashid [57] proposed a dynamic Bayesian network-based networked process monitoring approach for fault detection, propagation identification, and root cause diagnosis. Reconstruction of causal maps based on the aforementioned formulations is completely data-driven and hence is contingent on quality data. High-quality data are the precondition for analyzing and guaranteeing the reliability of the results [58]. In causality analysis, such data quality issues result in two main errors: spurious causations and failure to identify existing causations. Although causal maps can be built based on the available expert knowledge of engineers and process flow diagrams (PFDs) as studied in [59, 60], it may not always be the apt option, particularly for complex and tightly-integrated processes.

Physics-informed modeling [61] is an area of research that has been gaining prominence recently and has been applied across a wide variety of areas such as climatology [62], rheology [63], and process systems engineering [64]. Fusing the physics information with the observed data for the reconstruction of causal maps for source detection of plant-wide oscillations is a scarcely studied problem. This is particularly important because the issue of spurious causations is common in datasets characterized by periodicity. Additionally, poor data quality may often lead to failure in identifying existing causal relations, particularly if the causality is weak and incorporating the physics information is useful in preserving such relations. With this motivation, this work presents a novel framework where one can reconstruct the causal maps for linear time-invariant dynamical systems by combining observed data and physics information thereby aiding in finding the source of plant-wide oscillations more reliably.

The cause-effect relationships identified based on statistical approaches have to be verified by hypothesis testing to establish the fact that the relationship truly exists and is not obtained by chance. Surrogate data-based hypothesis testing has been the main approach following the work of Theiler et al. [65] where it was used for testing non-linearity. Apart from testing nonlinearity, surrogate-based hypothesis testing has been extended to a wide range of applications such as tests for oscillations in noisy

signals [66], for chaos [67], for synchronization between two or more systems [68], and for causality analysis [69]. As will be discussed later, the existing approaches for generating surrogate datasets to test causal relationships are not effective for oscillatory datasets. Thus, in this chapter, a novel framework for generating surrogates for data sets involving periodic oscillations is also proposed. The main contributions of this chapter are outlined as follows:

1. A novel framework for amalgamating observed data with Physics-Information for the reconstruction of cause-effect networks.

2. A novel algorithm for the generation of surrogates for data sets with periodic oscillations.

The remainder of this chapter is organized as follows. In section 3.2, a novel framework of fusing the physics information with observed data, referred to as Physics-Informed Sparse Causal Inference, is introduced and demonstrated with an example. In section 3.3, a novel algorithm for the generation of surrogates to test causality for data sets involving periodic oscillations is introduced. Section 3.4.1 and 3.4.2 present the results from two case studies that demonstrate the effectiveness of the proposed method. In section 3.5, the concluding remarks are presented.

3.2 Proposed Methodology

Granger-based causality approaches face challenges with oscillatory data sets. Although sparse Granger causality removes spurious causations, it may also eliminate weak causal relationships. To address these limitations, this work presents a novel formulation that integrates physics information and observed data, resulting in an effective and dependable establishment of causal relationships.

3.2.1 Physics-Informed Sparse Causal Inference

The proposed model introduces expert knowledge as a constraint in the original sparse Granger causal problem and is formally defined as follows.

$$\begin{aligned} \min_{M} \quad & ||X_t - MX_{td}||_2 + \lambda||M||_1 \\ \text{s.t.} \quad & ||M_{ij}||_1 \text{ is } \begin{cases} = 0, & \text{for } x_i \not\rightarrow x_j \\ > \epsilon, & \text{for } x_i \rightarrow x_j \end{cases} \end{aligned} \tag{3.1}$$

where M represents the coefficient matrix. $||M_{ij}||_1$ represents the L_1 norm of the coefficient matrix elements corresponding to the effect of the i^{th} variable on the j^{th} variable. X_t denotes the measured variables at time t, X_{td} represents the past of X with a time delay d, and λ is the hyper-parameter that dictates the sparsity level. To illustrate the optimization problem shown in 4.1, let us examine a linear time-invariant dynamical process with three variables and pre-specified time lags up to two. The explicit definitions of relevant notations are presented below.

$$X_t = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \quad X_{td} = \begin{bmatrix} x_1(t-1) \\ x_1(t-2) \\ x_2(t-1) \\ x_2(t-2) \\ x_3(t-1) \\ x_3(t-2) \end{bmatrix} \tag{3.2}$$

$$M = \begin{array}{c} \begin{matrix} \text{effect of } x_1 & & \text{effect of } x_2 & & \text{effect of } x_3 & \end{matrix} \\ \begin{bmatrix} m_{11}^1 & m_{11}^2 & m_{21}^1 & m_{21}^2 & m_{31}^1 & m_{31}^2 \\ m_{12}^1 & m_{12}^2 & m_{22}^1 & m_{22}^2 & m_{32}^1 & m_{32}^2 \\ m_{13}^1 & m_{13}^2 & m_{23}^1 & m_{23}^2 & m_{33}^1 & m_{33}^2 \end{bmatrix} \end{array} \tag{3.3}$$

The time-delayed variables are vertically stacked so that the corresponding coefficient matrix M can be presented so that each of its columns corresponds to the effect of one variable on the other. In the illustration shown in Eq. 3.3, the yellow-colored

columns represent the effect of the past of X_1 on all the observed variables X_1, X_2 and X_3 (denoted as M_1). Similarly, the blue and green columns represent the effects of X_2 and X_3, denoted as M_2 and M_3, respectively.

$$M_1 = \begin{bmatrix} m_{11}^1 & m_{11}^2 \\ m_{12}^1 & m_{12}^2 \\ m_{13}^1 & m_{13}^2 \end{bmatrix} \quad M_2 = \begin{bmatrix} m_{11}^2 & m_{21}^1 \\ m_{22}^1 & m_{22}^2 \\ m_{23}^1 & m_{23}^2 \end{bmatrix}$$

$$M_3 = \begin{bmatrix} m_{31}^1 & m_{31}^2 \\ m_{32}^1 & m_{32}^2 \\ m_{33}^1 & m_{33}^2 \end{bmatrix} \tag{3.4}$$

Each element of the coefficient matrix m_{ij}^{τ} represents the influence of $x_i(t-\tau)$ on $x_j(t)$. Finally, the total effect of x_i on x_j is captured by the j^{th} row of M_i matrix, denoted as M_{ij}. For example, the effect of x_3 on x_2 is shown in Eq. 3.5.

$$M_{32} = \begin{bmatrix} m_{32}^1 & m_{32}^2 \end{bmatrix} \tag{3.5}$$

The magnitude of elements in matrix M_{ij} determines the presence or absence of a cause-effect relationship. Expert information can be incorporated into the optimization problem through constraints on M_{ij}, establishing reliable causal relationships. In the system being described, as an illustration, we consider two scenarios involving the following expert information.

1. $x_1 \rightarrow x_2$**:** Since x_1 causes x_2, it is necessary for at least one element in the vector M_{12} to have a statistically significant value. This condition can be incorporated using the following equation.

$$||M_{12}||_1 = |m_{12}^1| + |m_{12}^2| > \epsilon \tag{3.6}$$

where ϵ is a small positive number which ensures that the magnitude of the coefficient is not too low so that it does not appear as generated by noise.

2. $x_3 \not\rightarrow x_1$**:** Since x_3 does not cause x_1, then all elements in the vector M_{31} should be statistically insignificant. This condition can be incorporated using the Eq. 3.7.

$$||M_{31}||_1 = |m_{31}^1| + |m_{31}^2| = 0 \tag{3.7}$$

Thus, it is possible to derive more reliable cause-and-effect relationships from measured data while aligning with the insights and expertise of specialists in the field.

It is worth emphasizing that the requisite physics-related information essential for conducting constraint optimization need not necessarily come from first principles-based knowledge. It can be any rudimentary knowledge about the gains and interactions in the system. In real-world scenarios which involve complex interconnections, obtaining first-principle relations is difficult. This feature thus represents a noteworthy advantage of the proposed methodology.

3.3 Surrogate Data Analysis

Hypothesis testing serves as a statistical procedure to ascertain that the observed results are not merely coincidental but represent the true characteristics of the underlying system. Romano [70] et al. in his work summarizes various developments in the field of hypothesis testing. In the context of surrogate data techniques, a specific attribute of a data set is compared with the corresponding attribute calculated from a set of surrogate data. These surrogate data possess similarities to the original data set but lack the property being tested. Such comparative analysis enables us to determine the extent to which the observed property deviates from what would be expected by chance.

3.3.1 Proposed Surrogate Data Analysis

In the context of this research objective, the null hypothesis posits that there is no causal relationship. The primary aim is to produce a large number of surrogate data

sets using the observed data under this null hypothesis. Previous methodologies, such as Fourier transform-based surrogates and cycle phase permutation [71], have been ineffective in eliminating the property under investigation in surrogates for oscillatory data. This inadequacy can be attributed to periodicity in the data, as similar patterns repeat themselves at regular intervals. Consequently, a novel approach employing time-shifted surrogates [72] is proposed to overcome this limitation and assess the significance of the identified cause-effect relationship in oscillatory data sets. The algorithm is pictorially demonstrated in Figure 3.1.

In the case of VAR-based causality analysis, ensuring no causation essentially translates into ensuring no significant correlation between $x_i(t)$ and $x_j(t-1:t-d)$. Thus, the regressors and the predicted variables must be shifted in time such that the cross-correlations at the shifted lags are insignificant. This notion forms the basis for the time-shifted surrogates approach [72] where signals are shifted by random lags to generate the surrogates. In the case of periodic signals, such a random shift may not achieve a zero correlation among the variables due to the periodic nature of the cross-correlation function. Thus, this work proposes a modification to the time-shifted surrogates approach which is outlined in the following points.

The proposed surrogate data generation approach is schematically depicted in Figure 3.1. Upon generating N surrogates, a given causality inference method is applied to each of the surrogates which would result in a distribution for each of the coefficients in M. The upper and lower confidence limits, denoted as $m_{ij}^{\tau}|^{U}$ and $m_{ij}^{\tau}|^{L}$, respectively, are calculated based on the 95% confidence interval. The final decisions about the existence of a relationship between x_i and x_j are contingent upon the fulfillment of the following condition.

$$\text{Decision}: \begin{cases} x_i \not\rightarrow x_j, & \text{If } \; m_{ij}^{\tau}|^{L} \leq m_{ij}^{\tau} \leq m_{ij}^{\tau}|^{U} \;\; \forall\, 1 \leq \tau \leq d \\ x_i \rightarrow x_j & \text{Otherwise} \end{cases} \tag{3.8}$$

Algorithm 1 Proposed Surrogate Generation algorithm

1. The two signals x_i and x_j are shifted by a lag δ_{ij} where the cross-correlation between the two signals is zero. With this time shift, if $x_i(t)$ is regressed on the time-shifted $x_j(t - \delta_{ij} - 1 : t - \delta_{ij} - d)$, the corresponding coefficients would be zero.

2. To obtain multiple surrogates, multiple subsets from the time-shifted dataset are selected in a moving window fashion. Such a surrogate dataset generation ensures that each surrogate is at the same lag δ_{ij} where the cross-correlation is essentially zero.

3. A similar procedure is followed for all $j = 1 : p$ $(j \neq i)$ which removes causal relations of all the variables with x_i. It can be noted that only x_j's are shifted in time. Thus, this approach preserves the auto-correlation structure of x_i while removing cross-correlations between x_i and all the x_j's.

4. The same procedure is repeated for all $i = 1 : p$.

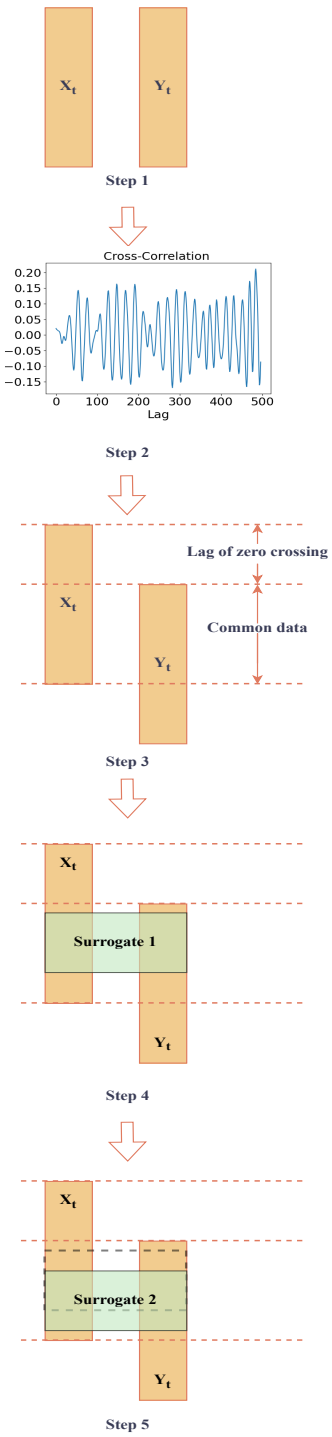


Figure 3.1: Pictorial representation of surrogate generation

3.4 Case studies

In this section, the results obtained from two case studies are presented that demonstrate the efficacy of the proposed approach compared to purely data-based approaches. The first one is a simulation case study in which the data is generated from an assumed VAR model. The second one is an industrial case study where the proposed approach is applied to a refinery separation unit data set characterized by plant-wide oscillation.

3.4.1 Simulation Case study

A causal map consisting of 5 variables and 5 direct causal relations is considered for the analysis in this simulation case study. The cause-effect relationships are shown in Figure 3.2. It can be observed from the causal map that $X1$ and $X5$ are the source variables and $X2$, $X3$, and $X4$ are the effect variables. To establish oscillations in the system, the source variables $X1$ and $X5$ are generated using sinusoidal signals as shown below where $\omega_1 = 0.01$ and $\omega_2 = 0.05$ and e_1 and e_2 are Gaussian white noise.

$$x_1(t) = \sin(\omega_1 t) + e_1(t) \tag{3.9}$$

$$x_5(t) = \sin(\omega_2 t) + \sin(\pi/2 - \omega_2 t) + e_5(t) \tag{3.10}$$

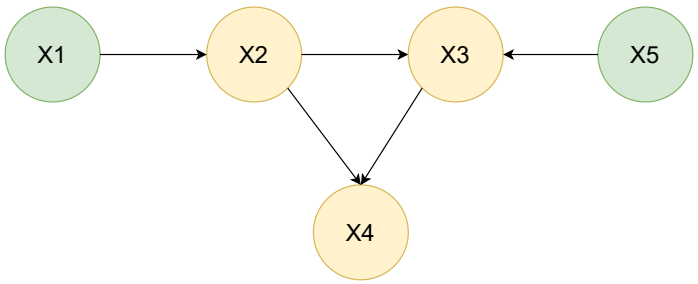


Figure 3.2: True Causal map

The variances of noises e_1 and e_2 are taken as 0.05 and 0.1 respectively.

Each cause-effect relationship in a causal map varies from each other based on the

strength of their causation and therefore, is generally classified into two categories as weak and strong causation. To consider this aspect of causation, we have introduced two weak links in the considered causal map. Here the causal link $X2$ causing $X3$ and $X3$ causing $X4$ are held as weak causation by giving a lower magnitude of coefficient 0.2 to these two links and a higher magnitude to the remaining causal links. The coefficient matrix M used for generating the data is given below.

$$M_{data} = \begin{bmatrix} 0.6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -0.7 & 0.2 & 0 \\ -0.5 & 0 & 0.8 \\ 0.1 & 0 & -0.3 \\ 0 & -0.4 & 0 \\ 0 & 0.5 & 0.2 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.5 \\ 0 & 0 & -0.3 \\ 0 & 0 & 0.4 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^T$$

It can be observed that apart from the causal relations, each variable is also generated by its past. The remaining variables are generated as follows:

$$\begin{bmatrix} x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} = M_{data}\ X_{td} + e_{2:4}(t) \tag{3.11}$$

The covariance for noise $e_{2:4}(t)$ is taken as $\mathrm{diag}(0.05, 0.05, 0.035)$. A total of 1000 samples were generated using the above equation. The corresponding plots of all the variables are shown in Figure 3.3.

First, the GC method is applied to obtain the coefficient matrix and the significance of each causation was verified by the proposed surrogate data analysis approach based on 100 generated surrogates. The causal map obtained by GC is shown in Figure 3.4. From the figure, it is evident that in addition to identifying actual cause-effect links

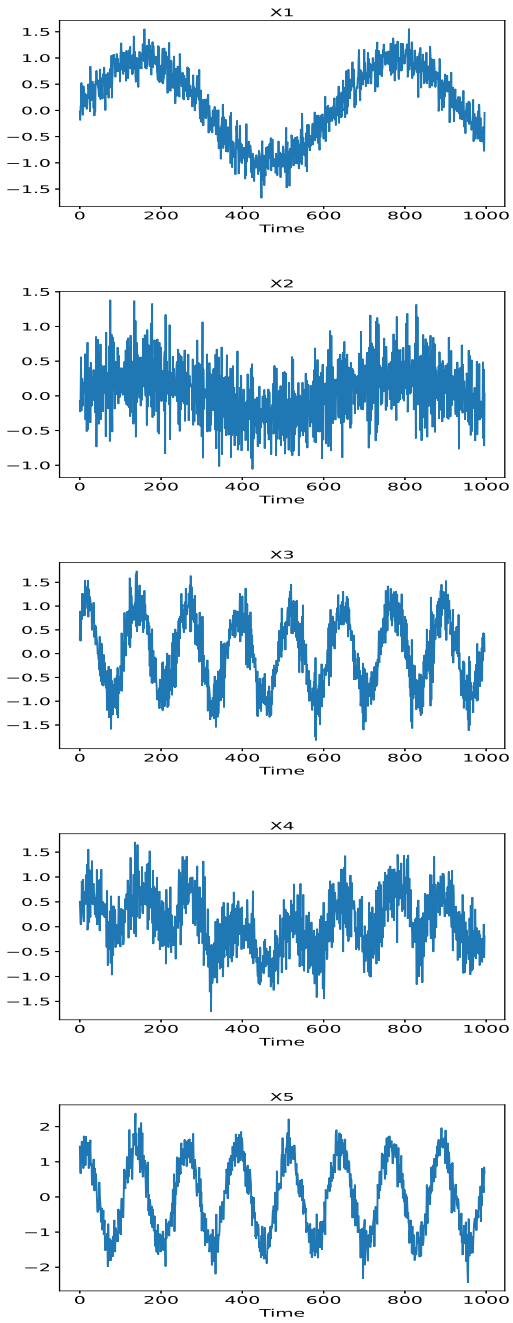


Figure 3.3: Plots of variables of simulation case study

(indicated by solid arrows), GC gave seven spurious causations (indicated by dashed arrows). These are as follows: $X3$ causing $X1$ and 5, $X4$ causing $X1$, $X2$, $X3$ and $X5$, and $X5$ causing $X1$. This illustrates that Granger causality cannot accurately reconstruct the true causal map when we have periodic oscillations in the system.

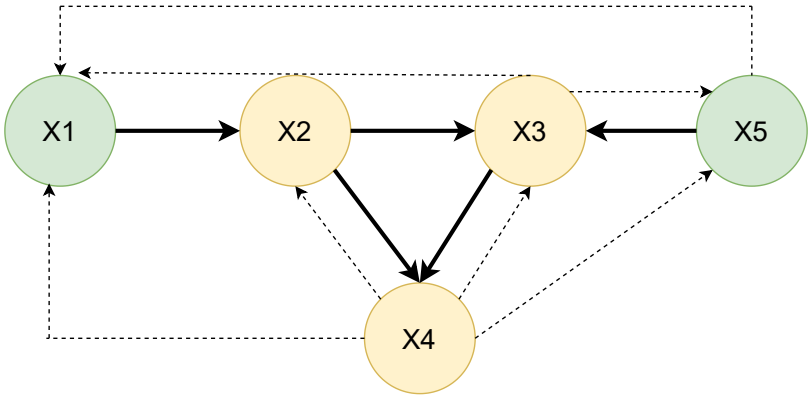


Figure 3.4: Causal map obtained by Granger Causality. The dotted arrows represent the spurious causal links while the boldened arrows represent the true existing causal links

SGC was then applied to the data, and the coefficient matrix was determined and causal relations were verified by the proposed surrogate data analysis approach. The obtained causal map for SGC is shown in Figure 3.5. As expected, SGC reduced the number of cause-effect links in the causal map compared to Granger, but it removed the true causal links while still retaining spurious ones. At a lower value of the hyper-parameter λ, it retains all the true causal links but results in many spurious causal links. When the value of λ is increased, the weak causal links disappear while spurious causal links remain. As the value of λ determines the level of sparsity that needs to be achieved, the existence of weak causal links limits our ability to increase its value since the weak links get removed at higher values of λ. It can hence be concluded that whenever the system has periodic oscillations coupled with the existence of weak causal links, the SGC approach fails to identify the true underlying cause-effect network.

As the purely data-based GC and SGC approaches failed to reconstruct the true

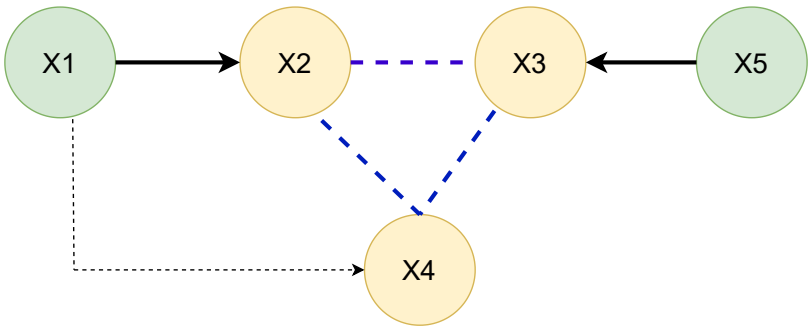


Figure 3.5: Causal map obtained by Sparse Granger Causality. The blue dotted lines represent the true causal links that were not identified.

cause-effect network, physics information about the true network was coupled with the observed data based on the proposed Physics-Informed Sparse Causal approach. After that, the determined coefficient matrix was analyzed by the proposed Surrogate data analysis approach. This is demonstrated by 2 scenarios where a different set of physics information is used as constraints, respectively. The physics information given as constraints in the first scenario are:

- $X2$ causes $X3$
- $X5$ does not cause $X4$

The obtained causal map for this scenario is shown in Figure 3.6. With the addition of the above two physics information, the proposed approach was able to reconstruct the true underlying cause-effect network. On analyzing the constraints in detail, it is evident that having some information about one of the causal links among the effect variables in the network helps preserve the weak causal links among the related effect variables. To demonstrate this effect further, a second scenario is considered in which a different set of physics information is considered. The physics information given as constraints in this scenario are:

- $X2$ causes $X4$
- $X3$ does not cause $X1$

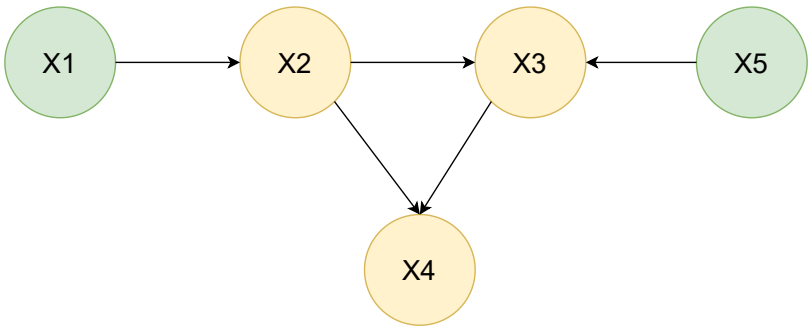
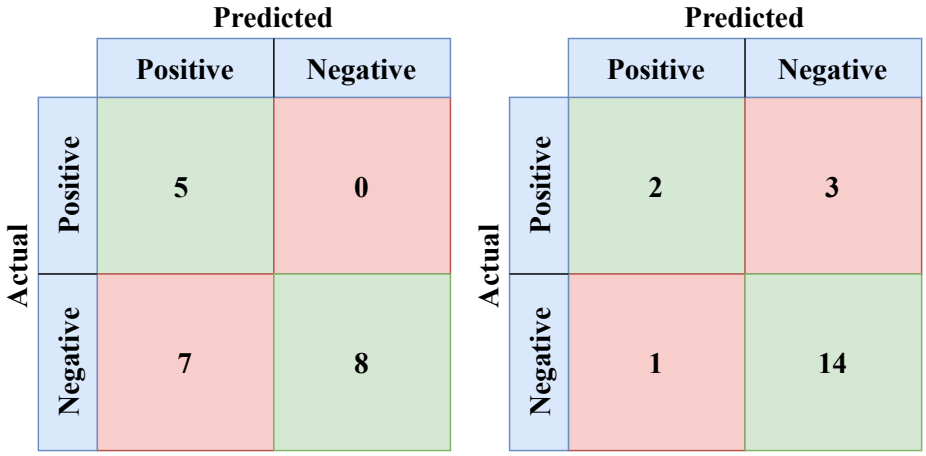


Figure 3.6: Causal Map obtained by the proposed physics-informed sparse causal inference approach

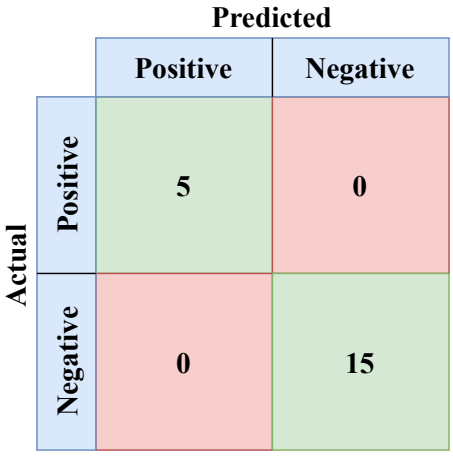
The obtained causal map for this scenario is similar to the one obtained in the first scenario and is shown in Figure 3.6. It is evident that, with the addition of the above physics information, the proposed method was again able to reconstruct the true underlying cause-effect network, which reinforces the conclusion drawn from the first scenario. It can thus be concluded that having some physics information about the system under consideration significantly improves the reconstruction of the underlying cause-effect network. In addition, the proposed approach also removes the spurious links which both GC and SGC approaches failed to do. This illustrates that, in the presence of periodic oscillating signals in the system coupled with weak and strong causal links, the proposed Physics-Informed Sparse Causal Inference identifies the true sources of the system and reconstructs the underlying cause-effect network. The Confusion Matrix in Figure 3.7 summarizes the results of this case study.

To further demonstrate the efficacy and robustness of the proposed approach under different oscillatory configurations, a different scenario involving time-varying frequencies and phase shifts is considered. In this case study, the data is structured such that the initial 1000 samples mirror the previous simulation, while an additional 1000 samples are deliberately designed to exhibit doubled frequencies ($\omega_1 = 0.02$ and $\omega_2 = 0.1$). Additionally, phase shifts of ϕ_1=45^0 and $\phi_2 = 60^0$ are introduced for X_1 and X_5 respectively, thereby creating a comprehensive data set that encapsulates these variations. The variable plots for this case are shown in 3.8. The data gen-



((a)) Granger causality

((b)) Sparse Granger causality



((c)) Proposed physics-informed sparse causal inference approach

Figure 3.7: Confusion matrix obtained for the simulation case study. In each figure, each block represents true positives, false negatives, true negatives, and false positives (clockwise starting from the top left). Positive and negative imply the existence and non-existence of causal relations respectively.

eration equations for the additional 1000 samples for the source variables are shown below.

$$x_1(t) = \sin(\omega_1 t + \phi_1) + e_1(t) \tag{3.12}$$

$$x_5(t) = \sin(\omega_2 t + \phi_1) + \sin(\pi/2 - \omega_2 t + \phi_2) + e_5(t) \tag{3.13}$$

While maintaining consistency in all other aspects of the first simulation setup, including the physics information, the optimization process successfully produces an accurate causal map that reflects the true underlying relationships as shown in 3.6. This outcome underscores the robustness of our framework while dealing with diverse oscillatory configurations.

A key aspect of VAR models is that the addition of delayed values of variables increases the number of parameters to be identified. This aspect thus increases the number of samples required to reliably identify the model parameters. Noting that identifying a VAR model is a multivariate regression problem, the sample size requirement for such problems depends on various factors such as the variability of data, the degree of dependency among the regressors, and the degree of noise. Although there are no definite methods to select the sample size particularly when many regressors exhibit linear correlations, there are certain thumb rules, one of which is that there must be at least 10 samples per independent regressor (or parameter). With 15 parameters per AR model in this case study, this ratio suggests a minimum sample size of 150 samples. It can be noted that the number of samples in this case study sufficiently exceeds this number thus resulting in a reliable estimation of the parameters.

3.4.2 Industrial Case study

The proposed method is applied to the industrial case study of the Australian Refinery process discussed in Chapter 2.

Being a refinery separation unit, the system under consideration imparts certain

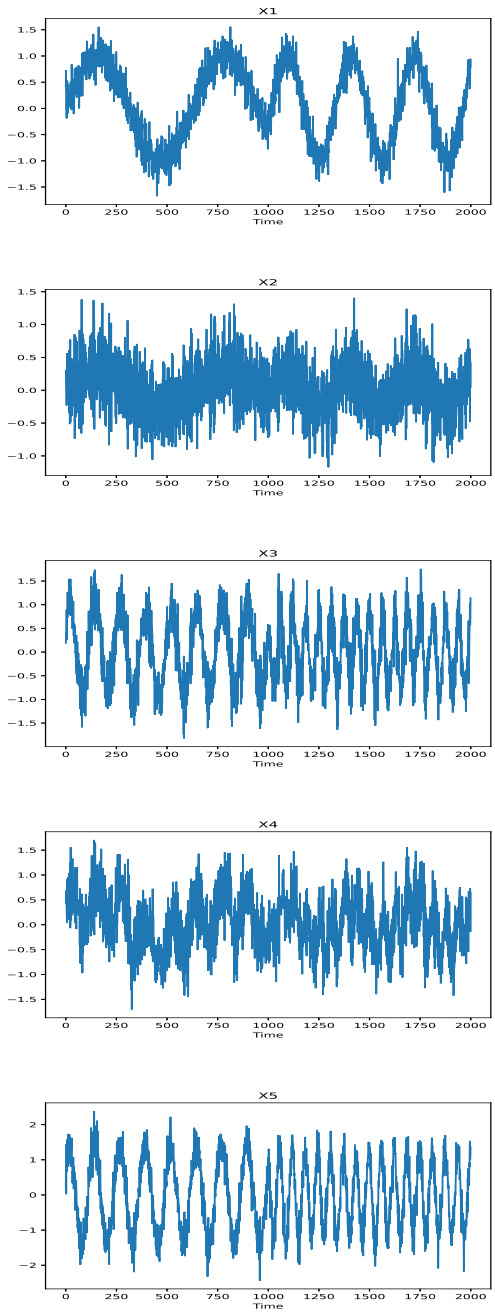


Figure 3.8: Plots of variables of additional simulation case study

common physics information that can be interpreted based on the common knowledge of the process. Here in this case study, the physics information used for constructing the cause-effect network is:

i $FC1\_OP$ causes $TC1\_PV$

ii $TC1\_OP$ causes $AC1\_PV$

iii $AC1\_OP$ does not cause $FC1\_PV$

iv $TC1\_OP$ does not cause $PC1\_PV$

v $PC2\_OP$ does not cause $PC1\_OP$

The rationale for these constraints is given in the following points.

i $FC1\_OP$, being the controller output of the steam flow into the separation unit, directly affects the temperature of the column/tray. Hence the first constraint is theoretically appropriate and can be given as a physics-information.

ii $AC1\_PV$ is the analyzer process variable for the column and is directly affected by the changes in temperature of the column $TC1\_OP$.

iii The amount of feed entering the column, $FC1\_PV$, is estimated based on the amount of product that needs to be withdrawn from the column and hence the analyzer output does not have any direct cause-effect relationship with the feed entering into the column.

iv $TC1\_OP$, being the temperature of the separation column under consideration, and $PC1\_PV$, being the pressure upstream to the column, are not directly related since the temperature of the column has a direct influence on the pressure within the column and not the pressure upstream of the column.

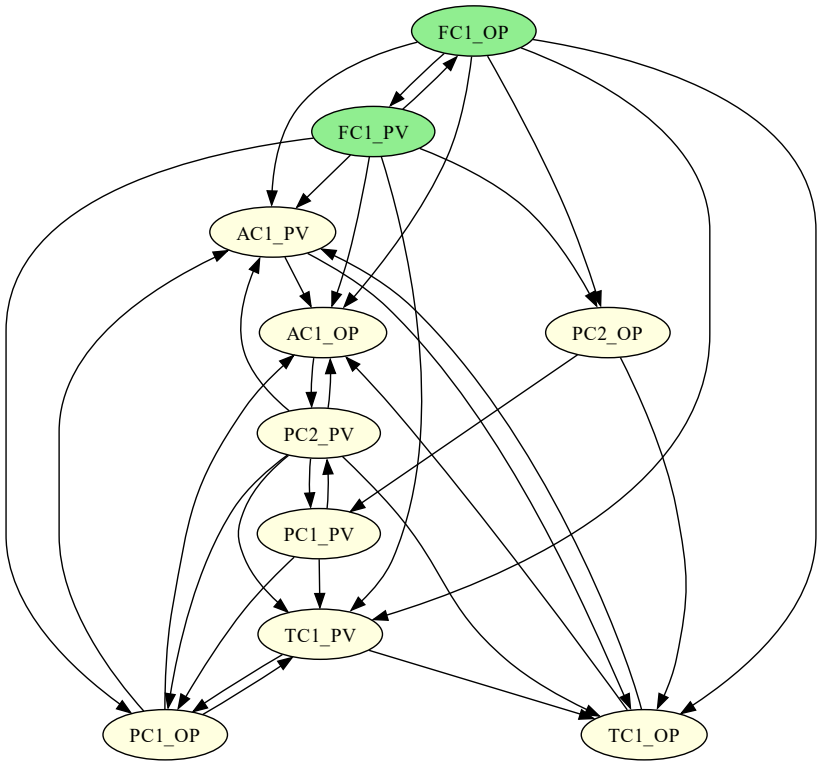


Figure 3.9: Physics-Informed causal map for Industrial case study

v $PC2\_OP$, being the pressure downstream to the separation column, and $PC1\_OP$, being the pressure upstream to the column, do not have any direct cause-effect relationship.

Based on the above constraints, the proposed method was applied to the dataset, and the obtained cause-effect network is shown in Figure 3.9.

The optimal value of hyper-parameter lambda for this case has been determined as explained previously and the optimal value of λ is evaluated as 0.05. Figure 3.10 depicts the variation of the objective function value with the values of lambda for the proposed method. The obtained cause-effect network using the proposed method is shown in Figure 3.9. The proposed method was able to identify the true source for plant-wide oscillation which is the $FC1$ loop. Based on this industrial case study, it can be concluded that purely data-based approaches like GC and SGC cannot

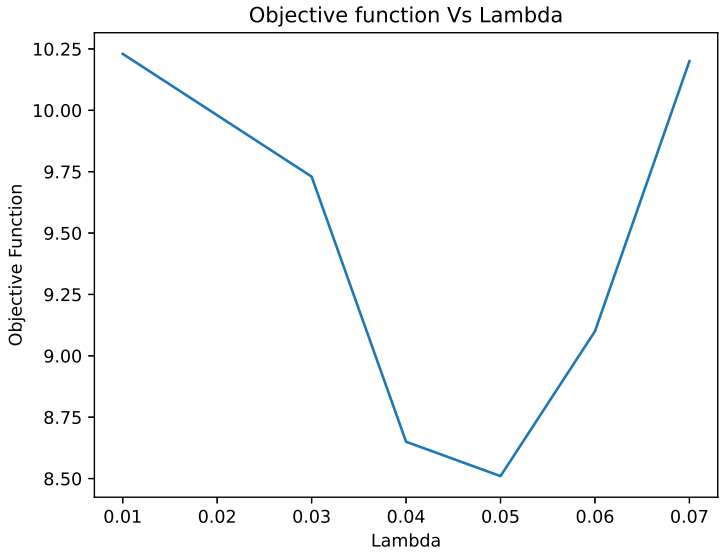


Figure 3.10: Objective function Vs Lambda for Physics-Informed Sparse Causality

be relied upon completely for the development of a cause-effect network. However, the addition of physics information to the observed data significantly improves the accuracy and reliability of the cause-effect network along with accurately determining the true source of plant-wide oscillations.

3.5 Conclusion

In this chapter, a novel physics-informed sparse causal inference approach is proposed that reconstructs the cause-effect network for source detection of plant-wide oscillations. Additionally, a novel algorithm for the surrogate generation of data sets involving periodic oscillations is also proposed. The proposed method addresses the shortcomings of existing data-based approaches like Granger causality and sparse Granger causality to detect the source when the data involves periodic oscillations. The physics information is formulated as constraints in a constrained optimization framework and the cause-effect network is reconstructed. Combining the physics information with observed data significantly improves the accuracy, reliability, and interpretability of the constructed causal network. The results from the simulated

and industrial case studies demonstrate the effectiveness of the proposed method. These promising results pave the way for further research and exploration of physics-informed causal inference beyond linear time-invariant dynamical systems.

Chapter 4

Diagnosis and Advanced Prediction of Flooding in Separation Columns

This chapter demonstrates the application of the physics-informed sparse causal inference methodology, developed in Chapter 3, to real-world industrial data, representing the third contribution of this thesis.

4.1 Introduction

In the domain of process engineering, the operating performance of chemical process equipment stands as a critical factor in ensuring safe and reliable operation. Among these various process equipment, the separation column serves as a linchpin for diverse industrial processes including distillation, absorption, and extraction. However, the occurrence of unexpected process disturbances like Flooding poses a significant challenge to its safe and reliable operation. Early flooding detection is thus crucial for a profitable and sustainable plant.

Flooding occurs when a liquid rises above a tray because of foaming or excessive downcomer fill-up [73]. This causes a significant loss in tray efficiency and hence plant profitability [74]. Flooding can be understood from an operational and design perspective. The vapor-liquid ratio (V/L) serves as a pivotal parameter for the design of a separation column. Usually, columns are designed with a higher (V/L) ratio, otherwise known as near-flooding conditions to achieve higher separation efficiency. Failing

to operate the column at optimal ratio results in significant operational challenges as a lower (V/L) ratio results in a condition called weeping and higher (V/L) results in flooding. Apart from this, differential pressure (DP) acts as another important indicator for the detection of flooding as a higher pressure drop is set to occur on the tray below the point of flooding.

Traditional approaches for identifying flooding instances rely heavily on manual observation and liquid holdup assessments, leaving room for inaccuracies and inconsistencies attributed to human error. Consequently, the reliability and precision of these methods are compromised, emphasizing the critical need for more advanced and automated flood detection techniques. Numerous studies have been dedicated to addressing the critical issue of flooding detection within separation columns. Given the detrimental impact of flooding on the overall performance and stability of these essential industrial units, the development of robust and reliable flooding detection methodologies remains a focal point in the field of chemical engineering and process optimization. Peiravan et al.[75] considered the correlation of the flooding effect with internal process variables, especially pressure drop across the column while Pihlaja [76] used the time derivative of the pressure drop. Brockkötter et al.[77] developed a Gaussian process-based data-driven model to predict the flooding state of filled liquid-liquid and high-pressure extraction towers. They tested the model's performance under different chemical systems and unstructured packing geometries, using various Gaussian process regression algorithms. Ochoa-Estopier et al.[78] presented a random forest model for the detection of flooding in distillation columns. Instead of relying on direct pressure measurements, they focussed on real-time measurements of flow rates, liquid levels, and temperatures to train a binary classification random forest model. Liu et al.[79] proposed a Convolutional Neural Network-based approach in their work. Various other machine-learning approaches were also employed to aid decision-making in industrial columns [80, 81].

Given the reliance of all the aforementioned research on data, the absence of high-

quality data would considerably compromise the results. Therefore, integrating human expertise with process data presents a more robust and precise approach to detecting flooding, ensuring greater reliability and accuracy in the results. Merely diagnosing the fault in isolation will not adequately address the requirements for mitigating process disturbances. It is imperative to employ advanced predictive methods to anticipate such faults, enabling proactive measures to prevent their occurrence and ensure uninterrupted operational efficiency. Hence in this chapter, PCA-based T^2 statistics for advanced prediction of flooding in separation columns are also proposed. Given the high dimensionality of process data, employing Principal Component Analysis (PCA) as a dimensionality reduction technique facilitates capturing maximal data variance within a lower-dimensional space [82] and the application of T^2 statistics serves as an effective anomaly detection technique [83]. The importance of automation in the process industry has increased dramatically in recent years. It has become a force in the entire chemical, oil, gas, and biotechnology industries [84]. Numerous researchers have developed various MATLAB-based toolboxes; however, the proprietary nature of MATLAB software presents certain challenges [85–87]. Hence in this chapter, we present a novel Python-based Graphic User Interface (GUI) toolbox to automate the generation of causal maps by combining process data and human knowledge.

The main contributions of this chapter are outlined as follows:

1. Applied Physics-Informed Sparse Causal Inference for Diagnosis of Flooding
2. PCA-based T^2 statistics for advanced prediction of flooding in separation columns
3. A Graphic User Interface (GUI) toolbox for automating the generation of the causal map.

The remainder of this chapter is organized as follows. Section 4.2 provides an overview of the process and problem description. Section 4.3 provides an overview

of the diagnosis of flooding using the Physics-Informed Sparse Causal Inference approach. In Section 4.4, the prediction of flooding using the PCA-based T^2 statistics is presented. In Section 4.5, the concluding remarks are presented.

4.2 Process & Problem Description

4.2.1 Process Description

Flooding was observed to occur extensively in the industrial de-propanizer column associated with a Gas Recovery Unit as shown in Figure 4.1. The column was designed to separate the mixed propane/butane stream by distillation from the De-Ethanizer bottoms. The C3/C4 stream is rich in hydrogen sulfide and hence it is treated in an amine treating unit before being fed to the de-propanizer column. The de-Propanizer bottoms are fed to an adsorber to remove Diisopropanolamine (DIPA) before being sent for storage. The de-propanizer overhead stream is sent to reflux to improve the product quality and a slipstream is routed to the drier before being sent to storage spheres.

4.2.2 Problem Description

A brief description of the problem of diagnosis and advanced prediction of flooding in the de-propanizer column is presented here. In this work, the differential pressure across the column is used as the Flooding Indication Variable (FIV) for diagnosis of flooding. Physics-Informed Sparse Causal Inference is used to identify the root cause of flooding and thereby generate the causal map. Based on the identified source variables, a PCA-based T^2 statistics methodology is developed to predict flooding events.

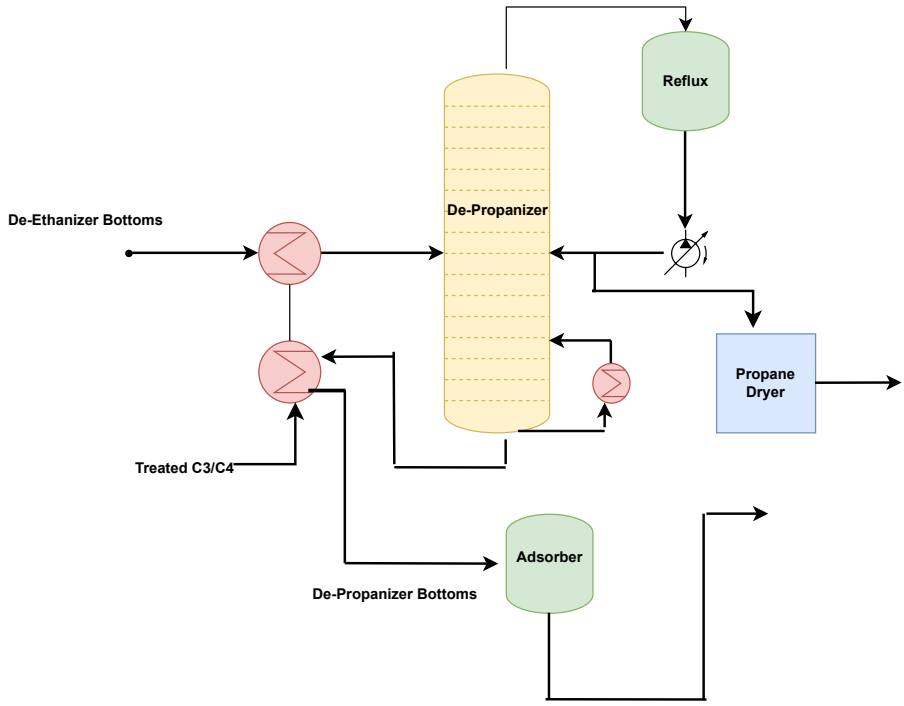


Figure 4.1: De-Propanizer System

4.3 Diagnosis of Flooding

4.3.1 Data

The industrial process data spans almost a year. The data is divided into bi-monthly data sets across ten Excel sheets. The sampling time for data collection is 1 minute. The process variables and their corresponding tags used in this work are shown in Table 4.1

4.3.2 Diagnosis

As discussed before, the proposed Physics-Informed Sparse Causal Inference approach is employed for the root source identification of Flooding. The mathematical framework for the diagnosis of flooding is given as:

$$\begin{aligned} \min_{M} \quad & ||X_t - MX_{td}||_2 + \lambda||M||_1 \\ \text{s.t.} \quad & Physics\ information \end{aligned} \tag{4.1}$$

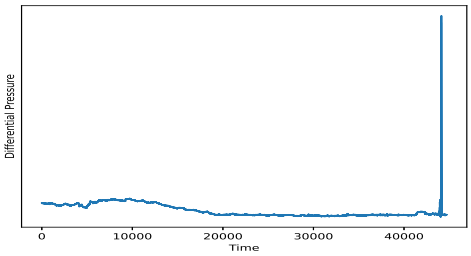
Process Variable	**Tag**
Reboiler Steam	RBS
Reflux	RFX
Feed Flow Rate	FFR
Feed Temperature	FT
Reboiler Return Temperature	RRT
Top Pressure	TP
Tray 11 Temperature	T11T
Tray 3 Temperature	T3T
Tray 37 Temperature	T37T
Tray 2 Pressure	T2P
Overhead Flow Rate	OFR
Flooding Indication Variable (Differential Pressure)	FIV
Overhead Temperature	OT
Overhead accumulated Temperature	OACT
Bottoms C4	BC4
Bottoms C5	BC5

Table 4.1: Process Variables and their Tags

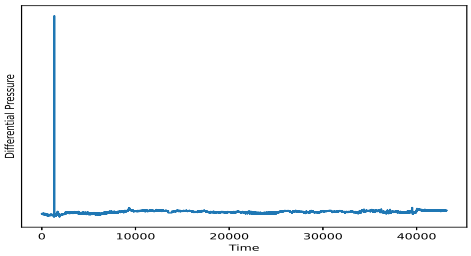
where M represents the coefficient matrix, X_t denotes the measured variables at time t, X_{td} represents the past of X with a time delay d, and λ is the hyper-parameter that dictates the sparsity level. In the context of our current investigation, focused on constructing a cause-effect network for diagnosing flooding in the De-Propanizer column, the physics information used is basic process understanding of an engineer or technician associated with the operation of a separation column. They are

- Reboiler steam causes Top pressure
- Top pressure causes FIV
- Reflux causes Top pressure
- Reflux causes Overhead temperature

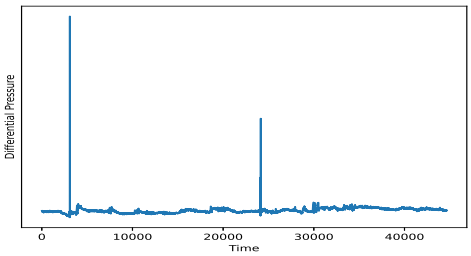
Differential Pressure (DP) being the Flooding Indication Variable was analyzed to identify the actual flooding events in the process data. The obtained plots of the DP are shown in Figure 4.2. Drawing from the insights gained through process understanding and the established alarm thresholds for flooding, it becomes apparent that a total of five flooding events transpired during the specified timeframe. To elaborate, these events unfolded in distinct periods: once during October-November, once during November-December, twice during December-January, and once during March-April.



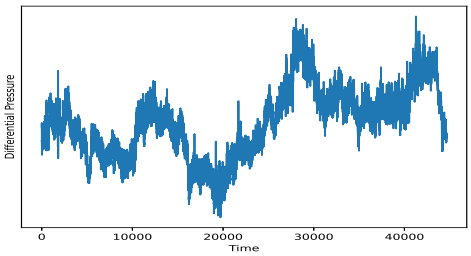
((a)) Oct-Nov



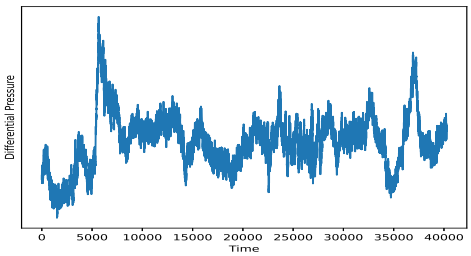
((b)) Nov-Dec



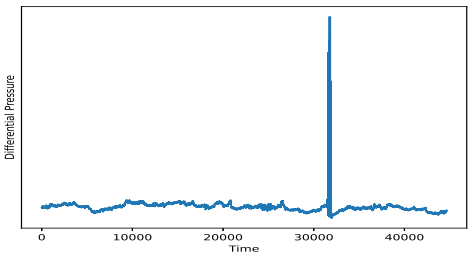
((c)) Dec-Jan



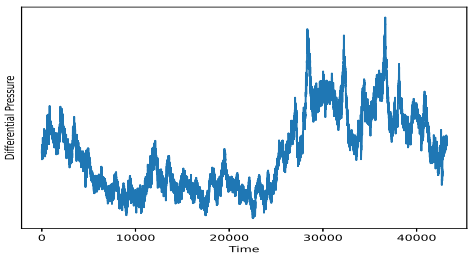
((d)) Jan-Feb



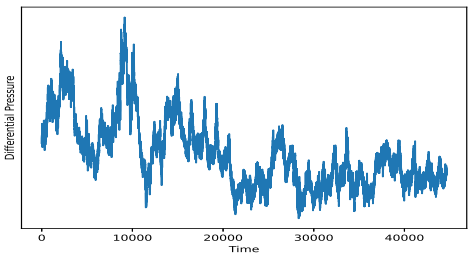
((e)) Feb-Mar



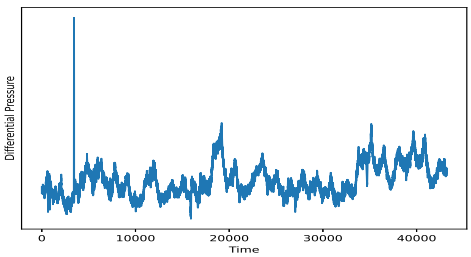
((f)) Mar-Apr



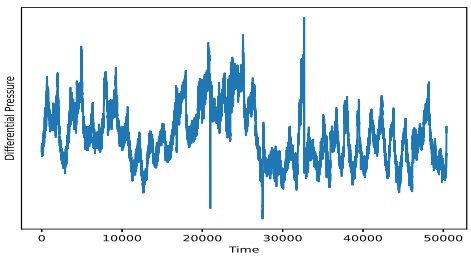
((g)) Apr-May



((h)) May-Jun



((i)) Jun-Jul



((j)) Jul-Aug

Figure 4.2: Plots of Differential pressure

After successfully identifying the instances of flooding, Physics-Informed Sparse causal Inference was employed to obtain the causal map shown in 4.3. The Differential pressure variable which is the Flooding Indication variable is presented in Red color, the source variables are presented in Green and the intermediary variables are presented in Yellow. It is evident from the causal map that a significant number of arrows are leaving the variables **Feed flow rate**, **Feed temperature**, **Reboiler steam** and **Reflux** compared to the number of entering which makes them the source variables and the others as intermediary variables. Even though the data consisted of 16 process variables, the variables Bottoms C4 and Bottoms C5 are not present in the constructed cause-effect network. This can be accounted for by the sparsity term in the objective function which removes variables associated with very weak causations.

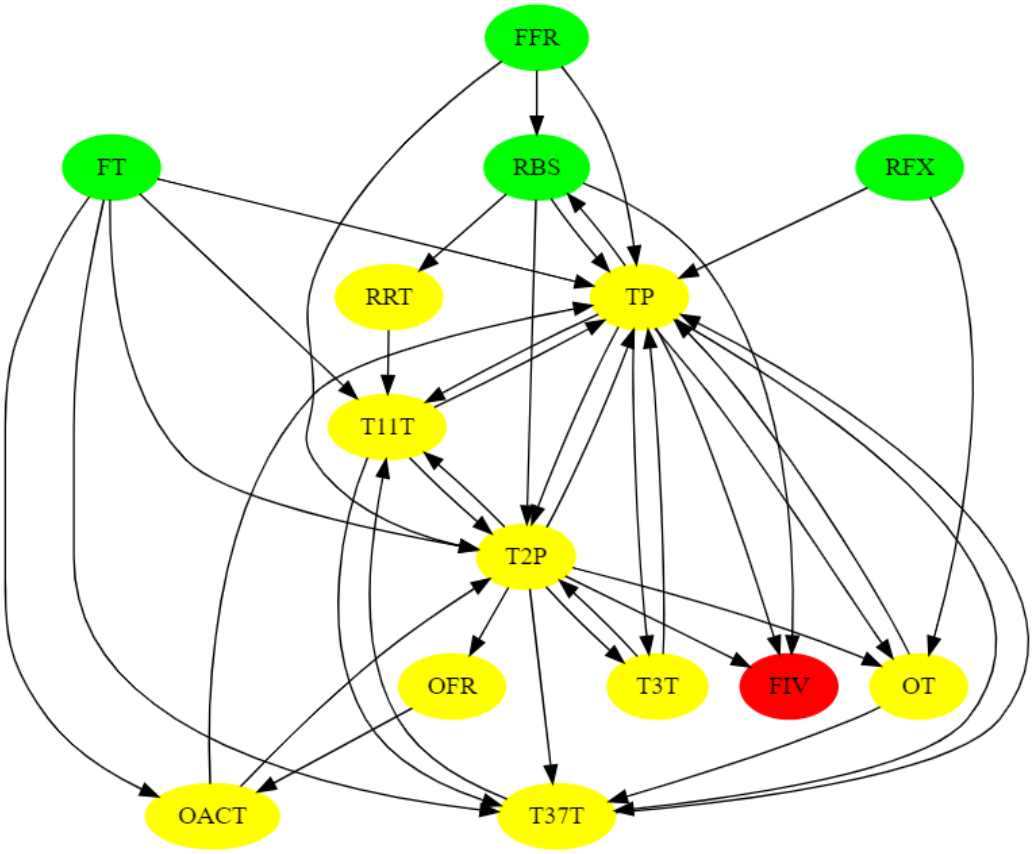
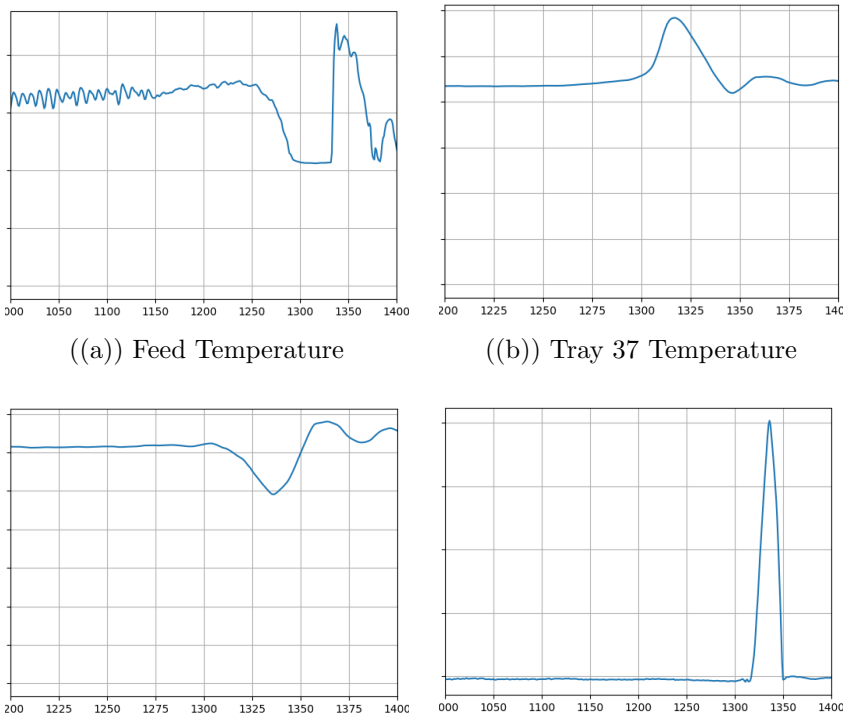


Figure 4.3: Causal Map for Diagnosis of Flooding

Once the causal map has been developed and source variables have been identified, various causal paths can be constructed starting from the source variables and ending with the Flooding indication variable. This systematic approach facilitates the verification of the causal map by analyzing whether these paths can explain the instances of flooding in the process data. For instance, one causal path constructed from the causal map in Figure 4.3 is Feed Temperature $->$ Tray 37 temperature $->$ Top Pressure $->$ FIV. This causal path explains the flooding event that transpired between October and November, as depicted in 4.4. This depiction serves as a robust verification, ensuring the precision and reliability of the generated causal map.



((a)) Feed Temperature ((b)) Tray 37 Temperature

((c)) Top Pressure ((d)) FIV

Figure 4.4: Verification of Causal Map

In the illustration above, the graphs representing the variables along the causal pathway are presented sequentially. The initial variations were detected in the Feed temperature, followed by changes in Tray 37 temperature and top Pressure, ultimately leading to the occurrence of flooding, as indicated by the FIV. Similarly, other flooding

events in the data can also be verified as discussed above.

4.4 Prediction of flooding

The accurate prediction of flooding is indispensable for maintaining operational efficiency, preventing equipment damage, and ensuring the overall effectiveness of separation columns. Predicting flooding with certainty is challenging due to the complexity of the process and the various factors involved.

4.4.1 Prediction Methodology

As observed from the Diagnosis of flooding where the causal map was developed, not all variables in the data contribute significantly to the modeling. Some variables demonstrate redundancy, while others wield substantial influence. Apart from that, dealing with the higher dimensional data in developing statistical process control tools possesses its challenges. The sheer volume of variables can lead to increased complexity, potential overfitting, and computational inefficiencies. As we navigate this terrain, it becomes imperative to employ sophisticated techniques for dimensionality reduction and feature selection. By discerning the essential variables and mitigating the impact of redundant ones, we aim to streamline the modeling process and enhance the efficacy of statistical process control tools in the context of flooding diagnosis. Hence in this work, Principal Component Analysis is used to reduce the dimensionality of the data. The T^2 statistic, often referred to as Hotelling's T-squared statistic, plays a pivotal role in detecting deviations from the expected behavior of a system or process. It is widely employed in various fields, including quality control, industrial processes, and multivariate data analysis. First proposed by Harold Hotelling [88], it is a generalization of Student's t statistic that is used in multivariate hypothesis testing. It is defined for a set of p variables $x = (x_1, x_2, \ldots, x_p)$ having mean values $\mu = (\mu_1, \mu_2, \ldots, \mu_p)$ and p×p covariance matrix, W Hotelling's T^2 statistic is given as [89] :

$$T^2 = (x - \mu)' W^{-1} (x - \mu) \tag{4.2}$$

So, in a typical online monitoring process, sensor measurements are collected sequentially and a decision of when a change occurs is made based on the value of T^2. Based on these discussed concepts the developed algorithm for the prediction of flooding is presented in Algorithm 2. The MATLAB codes for offline and online implementation are provided in the Appendix A

Algorithm 2 Prediction of flooding

1. Extract identified source variables data from the complete data set.
2. Carry out the pre-processing of data, including standardization.
3. Split the data into a training data set and a testing data set, where the training data set consists of 75% normal operation data and the testing data set consists of 25% normal operation data and flooding data.
4. Conduct Principal Component Analysis (PCA) based on training data for the 4 source variables identified and select 2 principal components (PCs).
5. Obtain the scores and loading matrices.
6. Calculate the T^2 statistics based on the equation:

$$T^2 = \sum \left(\frac{t_{i,a}}{s_a} \right)^2$$

   where $t_{i,a}$ is the score value for the i^{th} observation and a^{th} principal component, and s_a is the variance of the a^{th} principal component.

7. Calculate the T^2 statistics for the test data based on the PCA model of the training data.
8. Calculate the threshold value from the T^2 plot of the training data by keeping 1% to 2% of data points above the threshold.
9. Use the calculated threshold value for the prediction of flooding in the T^2 plot for the test data.

4.4.2 Prediction Results

The algorithm presented in the section Prediction Methodology was implemented on the available data set to predict the flooding events identified earlier to verify the efficacy of the proposed algorithm.

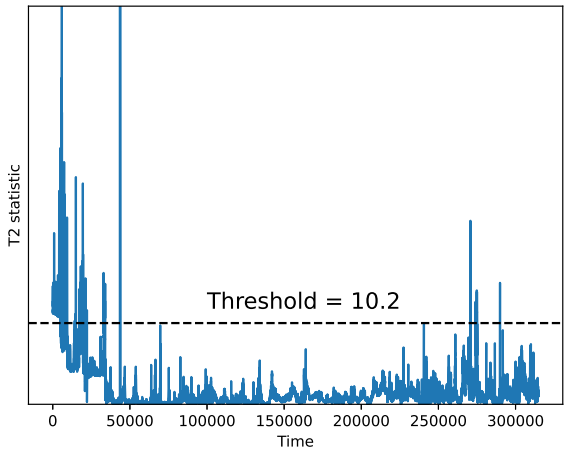
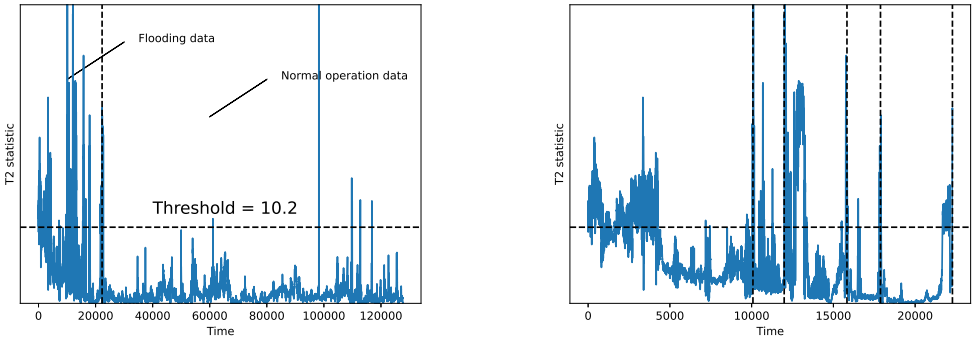


Figure 4.5: PCA based T^2 statistics plot for training data

The plot denoted as 4.5 delineates the PCA-derived T^2 statistics applied to the training dataset, exclusively populated with 75 percent of the normal operational data. The threshold value computation involved selecting the upper 1.5 percent of data points, yielding a threshold value of 10.2 as discerned from the T^2 plot of the training dataset. This determined threshold was subsequently employed for the test dataset, and the ensuing results are depicted in figure 4.6.

In Figure 4.6(a), the data points located to the left of the delineated dotted black line represent instances of flooding, while those to the right constitute 25 percent of normal operational data. For enhanced clarity, an amplified view of the flooding region is presented in Figure 4.6(b). The dotted black line within Figure 4.6(b) demarcates the temporal point at which flooding transpired. This specific occurrence of flooding is pinpointed through the analysis of the plot for Differential Pressure, identified as the Flooding-Indicating Variable (FIV). It is evident from the figure that



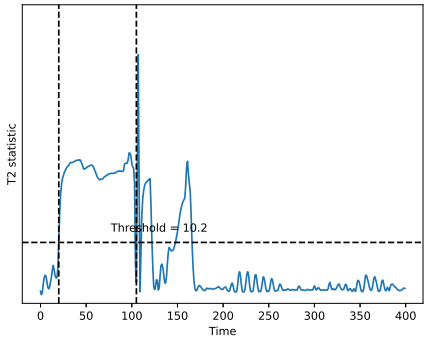
((a)) Complete test data ((b)) Enlarged Flooding region

Figure 4.6: PCA based T^2 statistics plot for test data

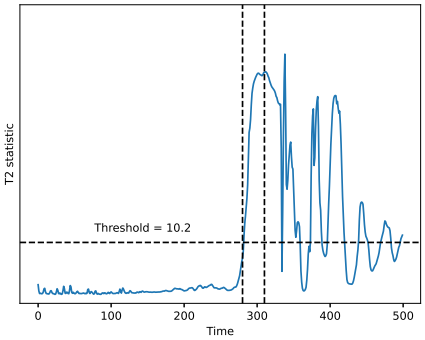
all 5 flooding events identified from FIV plots lie above the threshold value determined based on the training data. Among the five identified events, the second instance of flooding, occurring during the December-January timeframe, was attributed to an abrupt failure of the feed control valve. This failure led to instantaneous flooding of the column, rendering this particular event unpredictable. The prediction results of the other 4 events are shown in Figure 4.7.

Figure 4.7 depicts the PCA-based T^2 statistics plot designed for predicting four flooding events within the dataset. These plots specifically concentrate on the vicinity surrounding the flooding occurrences. The initial vertical dotted line denotes the moment when the T^2 value surpasses the predetermined threshold, while the subsequent dotted line signifies the exact point of flooding. This observation underscores the existence of a sufficiently extended prediction window, affording process operators ample time to undertake essential corrective measures. The detailed summary of the available prediction time for each flooding event is tabulated in Table 4.2.

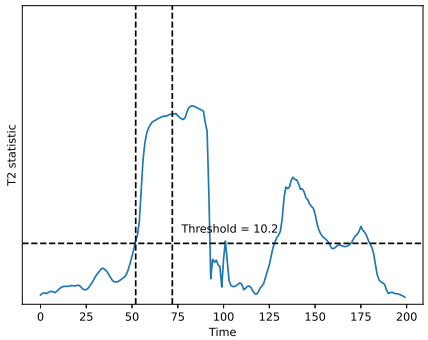
With the rapid advancement in control systems technology over the past few decades, the number of process sensors deployed for a particular plant has dramatically increased. In addition, due to the ease of configuring the alarms in control systems, the number of alarms in a plant has also gone up [90]. The false alarm rate calculated in the data set based on this developed algorithm is 2/month.



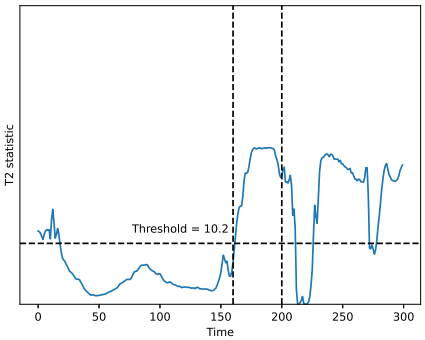
((a)) Flooding event in October - November



((b)) Flooding event in November - December



((c)) Flooding event in December - January



((d)) Flooding event in March - April

Figure 4.7: Prediction of flooding events

Flooding Event	**Time available for prediction**
October - November event	85 minutes
November - December event	30 minutes
December - January event	20 minutes
March - April event	40 minutes

Table 4.2: Flooding Prediction time

4.5 Conclusion

In this chapter, the proposed physics-informed sparse causal inference approach has been implemented on real industrial process data to identify sources of flooding in a separation column. The causal map generated and the variables identified as sources

are subsequently subjected to thorough analysis for validation. Additionally, an advanced predictive methodology is introduced to forecast flooding events enabling operators to anticipate incidents and take corrective actions within a reasonable time window. To facilitate the seamless application of the proposed methodology, a Graphic User Interface (GUI) toolbox is introduced which is discussed in detail in Appendix B. This toolbox serves a dual purpose: first, it automates the generation of the causal map, simplifying the often intricate process. Second, it enhances user accessibility by providing an intuitive interface for interacting with the physics-informed sparse causal inference model.

Chapter 5

Conclusions, Recommendations, & Future Work

This chapter outlines the conclusions drawn from the proposed methods and their implementations in various case studies. Additionally, potential future directions of research in this area are also outlined.

5.1 Conclusions

The overall underlying principle explored in this thesis is that of root cause analysis in Process Systems Engineering using Causal inference. Hence, all the proposed methods focus on various aspects of improving existing causal inference approaches for root cause analysis.

- In Chapter 2, a novel nonlinear sparse causal inference method based on the variational Gaussian process is proposed. The proposed method addresses the shortcomings of existing approaches like Granger causality, sparse Granger causality, and Gaussian process-based causality to identify and produce the true cause-effect relations among nonlinear variables. The main focus of the proposed method is to minimize spurious relations, providing a more accurate and reliable assessment of causation in nonlinear systems. This chapter contributes to the advancement of causal inference techniques, particularly in the context of

nonlinear relationships, and offers a promising avenue for enhancing the precision of causal assessments in complex systems.

- In Chapter 3, we addressed the shortcomings of existing data-based approaches like Granger causality and sparse Granger causality to detect the source when the data involves periodic oscillations. Reduced accuracy and reliability of causal maps due to the over-dependence on data is overcome by the proposed methodology which amalgamates human knowledge with data. The incorporation of physics information is a key aspect of the proposed approach, achieved by formulating constraints within a constrained optimization framework. This allows for the reconstruction of cause-effect networks, overcoming the drawbacks associated with purely data-driven methods. Notably, the methodology aims to strike a balance between empirical data and domain expertise to enhance the overall accuracy of source detection in systems with periodic oscillations. Additionally, a novel algorithm for the surrogate generation of data sets involving periodic oscillations is also proposed.

- In Chapter 4, we presented a robust framework combining the proposed Physics-Informed sparse causal inference with PCA-based T^2 statistics for the diagnosis and advanced prediction of flooding in the separation column. Successfully applied to the industrial data of a deethanizer column, the causal map generated and the variables identified as sources were subsequently subjected to thorough analysis for validation. To facilitate the seamless application of the proposed methodology, a Graphic User Interface (GUI) toolbox was also introduced. The developed methodology and toolbox will be provided to the industry for testing in real-time process operations.

5.2 Future Scope

This section outlines the various possible avenues of research in terms of extensions and further improvements to the methods presented in this thesis.

5.2.1 Extensions to the Proposed Physics-Informed Causal Inference

The proposed Physics-Informed Causal Inference in Chapter 3 is designed for linear systems and as discussed in Chapter 2, a significant number of systems in the real world tend to show nonlinear nature. Hence the Physics-Informed Causal inference can be extended to nonlinear systems by combining with the variational Gaussian Process methodology developed in Chapter 2. Modifying the objective function derived in Chapter 2 into a constraint optimization where physics information can be provided helps to overcome the issues related to the reliability and accuracy of the nonlinear data from process industries.

5.2.2 Causal Inference for Fault Prediction

In process industries, it is common to have measurements sampled at different rates thereby making the time scales different from one variable to another. Usually, in time domain causality analysis, data are assumed to be sampled at regular intervals for all the variables. Hence, this problem of causality analysis with multi-rate sampled data can be a good research topic to work on. Also, since flooding is a predominant issue in almost all process industries, transfer learning can be integrated with causal analysis, where the causal knowledge obtained from one particular process can be transferred to the process dealing with similar disturbance issues. It is also ideal to research the integration of more advanced machine learning techniques, such as deep learning or reinforcement learning, to handle complex and non-linear relationships within process data. As the operating conditions in process industries tend to change frequently due to the changing production requirements, it will be beneficial if the

model can adapt itself to these frequent changes and provide good results.

References

[1] S. Kathari and A. K. Tangirala, "A novel framework for causality analysis of deterministic dynamical processes," *Industrial & Engineering Chemistry Research*, vol. 61, no. 50, pp. 18 426–18 444, 2022.

[2] G. Sugihara *et al.*, "Detecting causality in complex ecosystems," *science*, vol. 338, no. 6106, pp. 496–500, 2012.

[3] Z. He, "Geopolitical risks and investor sentiment: Causality and tvp-var analysis," *The North American Journal of Economics and Finance*, vol. 67, p. 101 947, 2023.

[4] Z. Ghodsi, X. Huang, and H. Hassani, "Causality analysis detects the regulatory role of maternal effect genes in the early drosophila embryo," *Genomics Data*, vol. 11, pp. 20–38, 2017.

[5] A. Haug, F. Zachariassen, and D. Van Liempd, "The costs of poor data quality," *Journal of Industrial Engineering and Management (JIEM)*, vol. 4, no. 2, pp. 168–193, 2011.

[6] B. Saha and D. Srivastava, "Data quality: The other face of big data," in *2014 IEEE 30th international conference on data engineering*, IEEE, 2014, pp. 1294–1297.

[7] S. Haufe, V. V. Nikulin, K.-R. Müller, and G. Nolte, "A critical assessment of connectivity measures for eeg data: A simulation study," *Neuroimage*, vol. 64, pp. 120–133, 2013.

[8] T. Schreiber, "Measuring information transfer," *Physical review letters*, vol. 85, no. 2, p. 461, 2000.

[9] M. Le Van Quyen, J. Martinerie, C. Adam, and F. J. Varela, "Nonlinear analyses of interictal eeg map the brain interdependences in human focal epilepsy," *Physica D: Nonlinear Phenomena*, vol. 127, no. 3-4, pp. 250–266, 1999.

[10] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, "Causality detection based on information-theoretic approaches in time series analysis," *Physics Reports*, vol. 441, no. 1, pp. 1–46, 2007.

[11] M. Paluš, V. Komárek, Z. Hrnčíř, and K. Štěrbová, "Synchronization as adjustment of information rates: Detection from bivariate time series," *Physical Review E*, vol. 63, no. 4, p. 046 211, 2001.

[12] Y. Hirata and K. Aihara, "Identifying hidden common causes from bivariate time series: A method using recurrence plots," *Physical Review E*, vol. 81, no. 1, p. 016 203, 2010.

[13] L. Faes, G. Nollo, and K. H. Chon, "Assessment of granger causality by nonlinear model identification: Application to short-term cardiovascular variability," *Annals of biomedical engineering*, vol. 36, pp. 381–395, 2008.

[14] Y. Chen, G. Rangarajan, J. Feng, and M. Ding, "Analyzing multiple nonlinear time series with extended granger causality," *Physics letters A*, vol. 324, no. 1, pp. 26–35, 2004.

[15] N. Ancona, D. Marinazzo, and S. Stramaglia, "Radial basis function approach to nonlinear granger causality of time series," *Physical Review E*, vol. 70, no. 5, p. 056 221, 2004.

[16] D. Marinazzo, M. Pellicoro, and S. Stramaglia, "Kernel method for nonlinear granger causality," *Physical review letters*, vol. 100, no. 14, p. 144 103, 2008.

[17] D. Marinazzo, M. Pellicoro, and S. Stramaglia, "Kernel-granger causality and the analysis of dynamical networks," *Physical review E*, vol. 77, no. 5, p. 056 215, 2008.

[18] D. Marinazzo, W. Liao, H. Chen, and S. Stramaglia, "Nonlinear connectivity by granger causality," *Neuroimage*, vol. 58, no. 2, pp. 330–338, 2011.

[19] S. J. Schiff, P. So, T. Chang, R. E. Burke, and T. Sauer, "Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble," *Physical Review E*, vol. 54, no. 6, p. 6708, 1996.

[20] N. Nicolaou and T. G. Constandinou, "A nonlinear causality estimator based on non-parametric multiplicative regression," *Frontiers in neuroinformatics*, vol. 10, p. 19, 2016.

[21] M. Paluš and M. Vejmelka, "Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections," *Physical Review E*, vol. 75, no. 5, p. 056 211, 2007.

[22] Z. Chen, K. Zhang, L. Chan, and B. Schölkopf, "Causal discovery via reproducing kernel hilbert space embeddings," *Neural computation*, vol. 26, no. 7, pp. 1484–1517, 2014.

[23] P. Hoyer, D. Janzing, J. M. Mooij, J. Peters, and B. Schölkopf, "Nonlinear causal discovery with additive noise models," *Advances in neural information processing systems*, vol. 21, 2008.

[24] J. Pearl *et al.*, "Models, reasoning and inference," *Cambridge, UK: CambridgeUniversityPress*, vol. 19, no. 2, p. 3, 2000.

[25] C. E. Rasmussen, C. K. Williams, *et al.*, *Gaussian processes for machine learning*. Springer, 2006, vol. 1.

[26] E. Schulz, M. Speekenbrink, and A. Krause, "A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions," *Journal of Mathematical Psychology*, vol. 85, pp. 1–16, 2018.

[27] P.-O. Amblard, O. J. Michel, C. Richard, and P. Honeine, "A gaussian process regression approach for testing granger causality between time series data," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2012, pp. 3357–3360.

[28] G. Feng, J. G. Quirk, and P. M. Djurić, "Detecting causality using deep gaussian processes," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 472–476. DOI: 10.1109/IEEECONF44664.2019.9048963.

[29] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," *Journal of the American statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.

[30] M. I. Jordan, Z. Ghahramani, T. S. Jaakkola, and L. K. Saul, "An introduction to variational methods for graphical models," *Machine learning*, vol. 37, pp. 183–233, 1999.

[31] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: journal of the Econometric Society*, pp. 424–438, 1969.

[32] H. Lütkepohl, *New introduction to multiple time series analysis*. Springer Science & Business Media, 2005.

[33] P. A. Valdés-Sosa *et al.*, "Estimating brain functional connectivity with sparse multivariate autoregression," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1457, pp. 969–981, 2005.

[34] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[35] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009, vol. 2.

[36] C. E. Rasmussen, "Gaussian processes in machine learning," in *Summer school on machine learning*, Springer, 2003, pp. 63–71.

[37] C. Williams and C. Rasmussen, "Gaussian processes for regression," *Advances in neural information processing systems*, vol. 8, 1995.

[38] A. Wismüller, A. M. Dsouza, M. A. Vosoughi, and A. Abidin, "Large-scale nonlinear granger causality for inferring directed dependence from short multivariate time-series data," *Scientific reports*, vol. 11, no. 1, p. 7817, 2021.

[39] N. Thornhill, "Locating the source of a disturbance," in *Process Control Performance Assessment: From Theory to Implementation*. Springer London, 2007, pp. 199–225.

[40] N. Thornhill, "Finding the source of nonlinearity in a process with plant-wide oscillation," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 3, pp. 434–443, 2005.

[41] W. Bialkowski, "Dreams versus reality: A view from both sides of the gap: Manufacturing excellence with come only through engineering excellence," *Pulp & Paper Canada*, vol. 94, no. 11, pp. 19–27, 1993.

[42] D. B. Ender, "Process control performance: Not as good as you think," *Control Engineering*, vol. 40, no. 10, pp. 180–190, 1993.

[43] N. F. Thornhill and A. Horch, "Advances and new directions in plant-wide disturbance detection and diagnosis," *Control Engineering Practice*, vol. 15, no. 10, pp. 1196–1206, 2007.

[44] M. Bauer and N. F. Thornhill, "A practical method for identifying the propagation path of plant-wide disturbances," *Journal of Process Control*, vol. 18, no. 7, pp. 707–719, 2008.

[45] P. Duan, T. Chen, S. L. Shah, and F. Yang, "Methods for root cause diagnosis of plant-wide oscillations," *AIChE Journal*, vol. 60, no. 6, pp. 2019–2034, 2014.

[46] T. Yuan and S. J. Qin, "Root cause diagnosis of plant-wide oscillations using granger causality," *Journal of Process Control*, vol. 24, no. 2, pp. 450–459, 2014.

[47] Q. Chen, X. Xu, Y. Shi, X. Lang, L. Xie, and H. Su, "Mncmd-based causality analysis of plant-wide oscillations for industrial process control system," in *2020 Chinese Automation Congress (CAC)*, 2020, pp. 5617–5622.

[48] A. S. Sundaramoorthy, S. K. Varanasi, B. Huang, Y. Ma, H. Zhang, and D. Wang, "Sparse inverse covariance estimation for causal inference in process data analytics," *IEEE Transactions on Control Systems Technology*, vol. 30, no. 3, pp. 1268–1280, 2021.

[49] R. Raveendran, B. Huang, and W. Mitchell, "A variational bayesian causal analysis approach for time-varying systems," *IEEE Transactions on Control Systems Technology*, vol. 29, no. 3, pp. 1191–1202, 2020.

[50] R. Raveendran and B. Huang, "Variational bayesian approach for causality and contemporaneous correlation features inference in industrial process data," *IEEE transactions on cybernetics*, vol. 49, no. 7, pp. 2580–2590, 2018.

[51] L.-F. Deng, J.-G. Wang, J.-R. Su, Y. Yao, and J.-L. Liu, "Root cause diagnosis of plant wide oscillations using kernel granger causality," in *2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*, 2020, pp. 812–816.

[52] X. Li, J. Wang, H. Shao, Y. Yao, and L. Liu, "Root cause diagnosis of plant wide oscillations based fuzzy kernel granger causality," in *2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS)*, 2021, pp. 1405–1411.

[53] H.-S. Chen, Z. Yan, Y. Yao, T.-B. Huang, and Y.-S. Wong, "Systematic procedure for granger-causality-based root cause diagnosis of chemical process faults," *Industrial & Engineering Chemistry Research*, vol. 57, no. 29, pp. 9500–9512, 2018.

[54] M. Bauer, J. W. Cox, M. H. Caveness, J. J. Downs, and N. F. Thornhill, "Finding the direction of disturbance propagation in a chemical process using transfer entropy," *IEEE transactions on control systems technology*, vol. 15, no. 1, pp. 12–21, 2006.

[55] B. Lindner, L. Auret, M. Bauer, and J. W. Groenewald, "Comparative analysis of granger causality and transfer entropy to present a decision flow for the application of oscillation diagnosis," *Journal of Process Control*, vol. 79, pp. 72–84, 2019.

[56] C. Zou and J. Feng, "Granger causality vs. dynamic bayesian network inference: A comparative study," *BMC bioinformatics*, vol. 10, no. 1, pp. 1–17, 2009.

[57] J. Yu and M. M. Rashid, "A novel dynamic bayesian network-based networked process monitoring approach for fault detection, propagation identification, and root cause diagnosis," *AIChE Journal*, vol. 59, no. 7, pp. 2348–2365, 2013.

[58] L. Cai and Y. Zhu, "The challenges of data quality and data quality assessment in the big data era," *Data science journal*, vol. 14, 2015.

[59] L. H. Chiang, B. Jiang, X. Zhu, D. Huang, and R. D. Braatz, "Diagnosis of multiple and unknown faults using the causal map and multivariate statistics," *Journal of Process Control*, vol. 28, pp. 27–39, 2015.

[60] L. H. Chiang and R. D. Braatz, "Process monitoring using causal map and multivariate statistics: Fault detection and identification," *Chemometrics and intelligent laboratory systems*, vol. 65, no. 2, pp. 159–178, 2003.

[61] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, 2021.

[62] K. Kashinath *et al.*, "Physics-informed machine learning: Case studies for weather and climate modelling," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, p. 20 200 093, 2021.

[63] T. Zhang, D. Wang, and Y. Lu, "Rheologynet: A physics-informed neural network solution to evaluate the thixotropic properties of cementitious materials," *Cement and Concrete Research*, vol. 168, p. 107 157, 2023.

[64] S. A. Niaki, E. Haghighat, T. Campbell, A. Poursartip, and R. Vaziri, "Physics-informed neural network for modelling the thermochemical curing process of composite-tool systems during manufacture," *Computer Methods in Applied Mechanics and Engineering*, vol. 384, p. 113 959, 2021.

[65] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. Doyne Farmer, "Testing for nonlinearity in time series: The method of surrogate data," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1, pp. 77–94, 1992.

[66] M. R. Allen and L. A. Smith, "Monte carlo ssa: Detecting irregular oscillations in the presence of colored noise," *Journal of climate*, vol. 9, no. 12, pp. 3373–3404, 1996.

[67] M. Small, D. Yu, and R. G. Harrison, "Surrogate test for pseudoperiodic time series data," *Physical Review Letters*, vol. 87, no. 18, p. 188 101, 2001.

[68] M. Palus and D. Hoyer, "Detecting nonlinearity and phase synchronization with surrogate data," *IEEE Engineering in Medicine and Biology Magazine*, vol. 17, no. 6, pp. 40–45, 1998.

[69] G. Lancaster, D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska, "Surrogate data for hypothesis testing of physical systems," *Physics Reports*, vol. 748, pp. 1–60, 2018.

[70] J. P. Romano and E. Lehmann, *Testing statistical hypotheses*, 2005.

[71] Janez, M. Palu, and A. Stefanovska, "Detecting couplings between interacting oscillators with time-varying basic frequencies: Instantaneous wavelet bispectrum and information theoretic approach," *Phys. Rev. E*, vol. 81, p. 036 207, 3 2010.

[72] R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: A case study on electroencephalographic signals," *Phys. Rev. E*, vol. 65, p. 041 903, 4 2002.

[73] M. King, *Process control: a practical approach*. John Wiley & Sons, 2016.

[74] H. Z. Kister, *Distillation operation*. McGraw-Hill Education, 1990.

[75] H. Peiravan, A. R. Ilkhani, and M. J. Sarraf, "Preventing of flooding phenomena on vacuum distillation trays column via controlling coking value factor," *SN Applied Sciences*, vol. 2, pp. 1–11, 2020.

[76] R. K. Pihlaja and J. P. Miller, *Detection of distillation column flooding*, US Patent 8,216,429, 2012.

[77] J. Brockkoetter, J. Ahndorf, and A. Jupke, "Prediction of flooding in packed liquid-liquid and high-pressure extraction columns using a gaussian process," *Chemie Ingenieur Technik*, vol. 93, no. 12, pp. 1907–1916, 2021.

[78] L. M. Ochoa-Estopier, S. Gourvénec, R. Cahors, N. Behara, and J.-B. Scellier, "Prediction of flooding in distillation columns using machine learning," *Digital Chemical Engineering*, vol. 7, p. 100 098, 2023.

[79] Y. Liu, Y. Jiang, Z. Gao, K. Liu, and Y. Yao, "Convolutional neural network-based machine vision for non-destructive detection of flooding in packed columns," *Sensors*, vol. 23, no. 5, p. 2658, 2023.

[80] M. Mojto, K. L'ubuškỳ, M. Fikar, and R. Paulen, "Data-based design of inferential sensors for petrochemical industry," *Computers & Chemical Engineering*, vol. 153, p. 107 437, 2021.

[81] J. Oeing, L. M. Neuendorf, L. Bittorf, W. Krieger, and N. Kockmann, "Flooding prevention in distillation and extraction columns with aid of machine learning approaches," *Chemie Ingenieur Technik*, vol. 93, no. 12, pp. 1917–1929, 2021.

[82] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[83] R. L. Mason and J. C. Young, *Multivariate statistical process control with industrial applications*. SIAM, 2002.

[84] S.-L. Jämsä-Jounela, "Future trends in process automation," *Annual Reviews in Control*, vol. 31, no. 2, pp. 211–220, 2007.

[85] A. K. Seth, "A matlab toolbox for granger causal connectivity analysis," *Journal of neuroscience methods*, vol. 186, no. 2, pp. 262–273, 2010.

[86] M. Lindner, R. Vicente, V. Priesemann, and M. Wibral, "Trentool: A matlab open source toolbox to analyse information flow in time series data with transfer entropy," *BMC neuroscience*, vol. 12, pp. 1–22, 2011.

[87] Z.-X. Zang, C.-G. Yan, Z.-Y. Dong, J. Huang, and Y.-F. Zang, "Granger causality analysis implementation on matlab: A graphic user interface toolkit for fmri data processing," *Journal of neuroscience methods*, vol. 203, no. 2, pp. 418–426, 2012.

[88] H. Hotelling, "The generalization of student's ratio," in *Breakthroughs in statistics: Foundations and basic theory*, Springer, 1992, pp. 54–65.

[89] Y.-M. Chou, R. L. Mason, and J. C. Young, "Power comparisons for a hotelling's t2 statistic," *Communications in Statistics-Simulation and Computation*, vol. 28, no. 4, pp. 1031–1050, 1999.

[90] P. Goel, A. Datta, and M. S. Mannan, "Industrial alarm systems: Challenges and opportunities," *Journal of Loss Prevention in the Process Industries*, vol. 50, pp. 23–36, 2017.

Appendix A: MATLAB Codes for Chapter 4

A.1 Offline Implementation

A sample MATLAB code for calculating loadings and score matrix is shown below:

```
% Sample dataset (replace this with your own data)
data = randn(100, 5);  % 100 samples with 5 features each

% Step 1: Mean center the data
data_mean_centered = data - mean(data);

% Step 2: Standardize
data_standardized = data_mean_centered ./ std(data_mean_centered);

% Step 3: Calculate the covariance matrix of the standardized data
covariance_matrix = cov(data_standardized);

% Step 4: Perform PCA using the built-in pca function
[coeff, ~, ~] = pca(data_standardized);

% Step 5: Calculate PCA scores
scores = data_standardized * coeff;

% Display the results
disp('Principal Component Loading Matrix:');
disp(coeff);

disp('PCA Scores:');
disp(scores);
```

A.2 Online Implementation

A sample MATLAB code for calculating T^2 statistic is shown below:

```
% Sample dataset (replace this with your own data)
data = randn(100, 4);  % Replace this with source variable data

% Loading matrix obtained from PCA model of offline simulation
coeff = transpose([Obtained loading matrix from PCA]);

% Calculation of score matrix
score = data * coeff;

% Step 5: Calculate the T2 statistics for each data point
T2 = sum(score.^2 ./ std(score).^2, 2);

% Display the T2 statistics
disp('T2-Statistics:');
disp(T2);

% Plot the T2 statistics plot
plot(T2)
ylim([0, 40]);
yline(10.2,'--','Threshold');    % Threshold value;
```

Appendix B: Graphic User Interface (GUI) Toolbox

A Graphical User Interface (GUI) toolbox was developed as an integral facet of Chapter 4 of this work. The GUI toolbox serves as a sophisticated front-end for interacting with the underlying algorithms and functionalities, offering researchers and practitioners a versatile and user-friendly interface implemented using Python and tailored to the specific requirements of this work. The impetus for developing the GUI toolbox arises from the inherent limitations of conventional command-line interfaces in managing the nuanced intricacies.

B.0.1 Implementation Details

The GUI toolbox is realized through the utilization of Python, a versatile programming language, coupled with robust GUI development frameworks. The primary frameworks employed include Tkinter and Pygubu, chosen for their widespread adoption, community support, and seamless integration capabilities. Tkinter provides a lightweight solution that seamlessly integrates with Python's standard library, while Pygubu offers a more feature-rich and aesthetically pleasing interface, enhancing the overall user experience. The GUI toolbox operates on an event-driven programming paradigm, where user actions and system events trigger specific responses. This approach ensures responsiveness as users interact with the interface. Event handlers are created to manage user inputs, such as button clicks, or menu selections, ensuring smooth communication between the graphical interface and the underlying computational backend.

B.0.2 Interface and Features

The GUI toolbox consists of three windows, where the first window allows the user to upload the process data and also provides the user with the platform to visualize the time series plots of the process parameters in the data as shown in Figure B.1. More specifically, the first window comprises the title label, a button to upload the process data, a drop-down menu from which a specific variable can be selected to visualize the time series plot on the frame to the right of the menu, and a button to add the physics-information which ensures a seamless transition to the subsequent window.

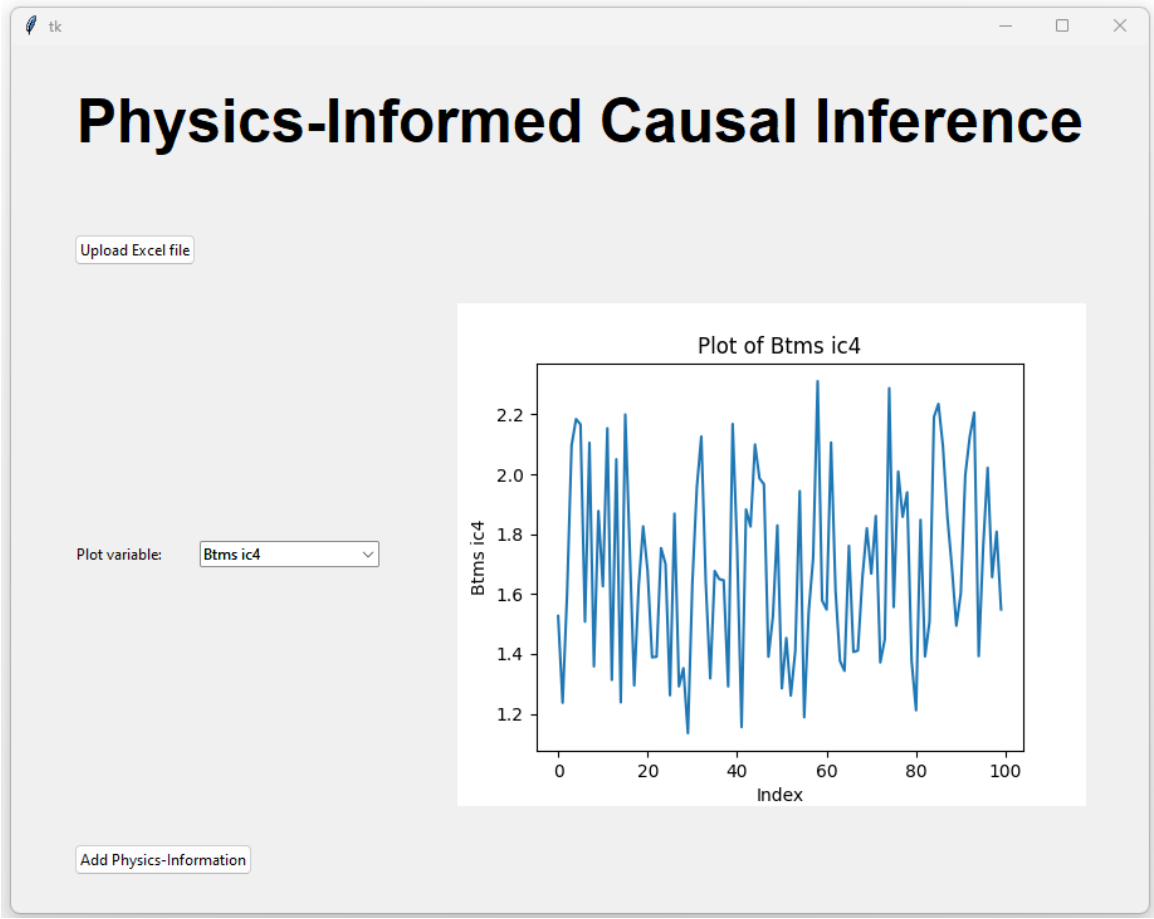


Figure B.1: First window of the GUI toolbox

After uploading and visualizing the process data, the user can go on to add the physics information by clicking the 'Add Physics-Information' button by which the subsequent window appears as shown in Figure B.2. This window consists of two drop-down menus to select the variables and one drop-down menu to select whether

one of them is causing the other or not. After selecting the variables and causation, the user can click on the 'Add' button to add that physics information into the computational framework. Once the user adds the required physics information, it then appears in the textbox at the bottom of the window for the user to verify at the end. The user can add any number of physics information they have about the process under consideration.

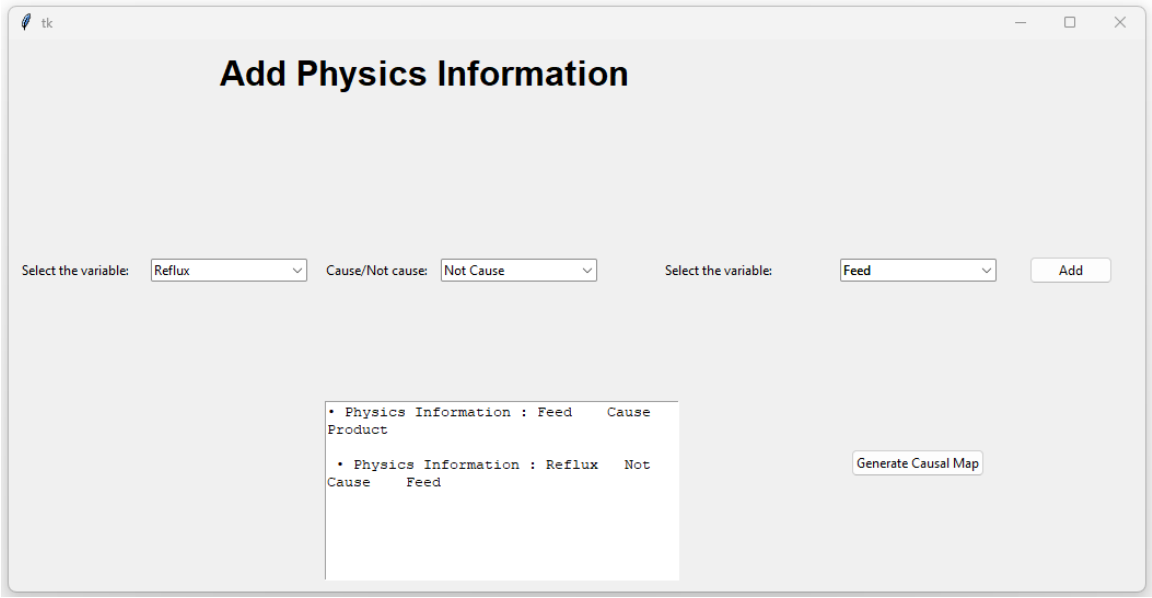


Figure B.2: Second window of the GUI toolbox

After adding the physics information, the user can go on to click the 'Generate Causal Map button' at the bottom right section of the window so that the third window appears with the generated causal map as shown in Figure B.3. This window consists of a frame where the generated causal map is shown along with options for Diagnostic and Predictive analysis of the causal map. More specifically, there are designated buttons in the window that enable the user to identify direct causal variables and indirect causal variables for both diagnosis and prediction. For example, if the user wants to identify the variables directly causing the variable 'Tray 10T', selecting that variable from the drop-down menu below the label Direct Causality for Diagnosis provides a list of such variables in the textbox at the bottom. In a similar vein, users can acquire indirect causative variables for diagnostic purposes, and this extends seamlessly to predictive analysis.

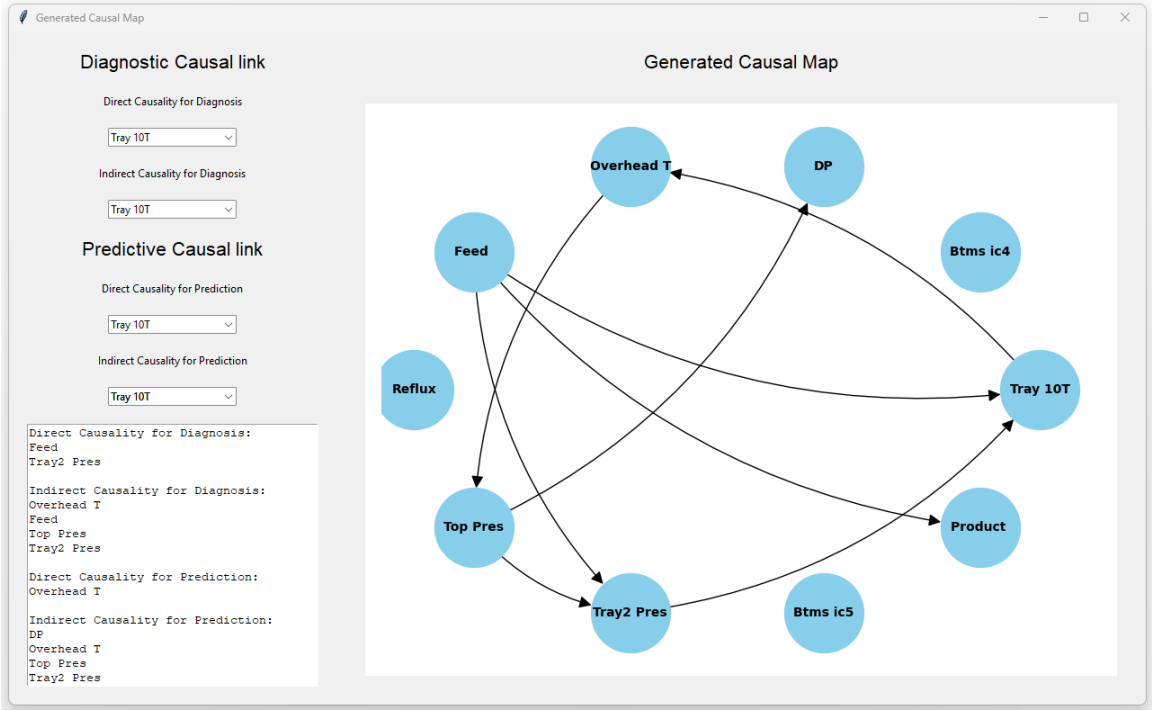


Figure B.3: Third window of the GUI toolbox