Improving the interpretability of causality maps for fault identification

by

Natali van Zijl

Thesis presented in partial fulfilment
of the requirements for the Degree

of

MASTER OF ENGINEERING
(EXTRACTIVE METALLURGICAL ENGINEERING)

in the Faculty of Engineering
at Stellenbosch University

Supervisor
Dr TM Louw
Process Engineering
Stellenbosch University

Co-Supervisor/s
Prof SM Bradshaw
Process Engineering
Stellenbosch University

Prof L Auret
Process Engineering
Stellenbosch University &
Data Science and Process Manager
Stone Three Digital

December 2020

Financial assistance from Anglo American Platinum towards this research project is hereby acknowledged. Opinions expressed and conclusions arrived at, as outlined in this document, are those of the author and do not necessarily reflect those of Anglo American Platinum.

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 19 November 2020

Plagiarism Declaration

1. Plagiarism is the use of ideas, material and other intellectual property of another's work and to present it as my own.

2. I agree that plagiarism is a punishable offence because it constitutes theft.

3. I also understand that direct translations are plagiarism.

4. Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully. I understand that the reproduction of text without quotation marks (even when the source is cited) is plagiarism.

5. I declare that the work contained in this document, except where otherwise stated, is my original work and that I have not previously (in its entirety or in part) submitted it for grading in this document or another document.

Initials and surname: N van

Zijl Date: 19 November 2020

_____

Abstract

Worldwide competition forces modern mineral processing plants to operate at high productivity. This high productivity is achieved by implementing process monitoring to maintain the desired operating conditions. However, a fault originating in one section of a plant can propagate throughout the plant and so obscure its root cause. Causality analysis is a method that identifies the cause-effect relationships between process variables and presents these in a causality map which can be used to track the propagation path of a fault back to its root cause. A major obstacle to the wide acceptance of causality analysis as a tool for fault diagnosis in industry is the poor interpretability of causality maps.

This study identified, proposed and assessed ways to improve the interpretability of causality maps for fault identification. All approaches were tested on a simulated case study and the resulting maps compared to a standard causality map or its transitive reduction. The ideal causality map was defined and all comparisons were performed based on its characteristics. Causality maps were produced using conditional Granger causality (GC), with a novel heuristic approach for selecting sampling period and time window.

Conditional GC was found to be ill-suited to plant-wide causality analysis, due to large data requirements, poor model order selection using AIC, and inaccuracy in the presence of multiple different residence times and time delays. Methods to incorporate process knowledge to constrain connections and potential root causes were investigated and found to remove all spurious connections and decrease the pool of potential root cause variables respectively. Tools such as visually displaying node rankings on the causality map and incorporating sliders to manipulate connections and variables were also investigated.

Furthermore, a novel hierarchical approach for plant-wide causality analysis was proposed, where causality maps were constructed in two subsequent stages. In the first stage, a less-detailed plant-wide map was constructed using representatives for groups of variables, and used to localise the fault to one of those groups of variables. Variables were grouped according to plant sections or modules identified in the data, and the first principal component (PC1) was used to represent each group (PS-PC1 and Mod-PC1 respectively). PS-PC1 was found to be the most promising approach, as its plant-wide map clearly identified the true root cause location, and the stage-wise application of conditional GC significantly reduced the required number of samples from 13 562 to 602.

Lastly, a usability study in the form of a survey was performed to investigate the potential for industrial application of the tools and approaches presented in this study. Twenty responses were obtained, with participants consisting of Stellenbosch University final-year/postgraduate students, employees of an industrial IoT firm, and Anglo American Platinum employees. Main findings include that process knowledge is vital; grouping variables improves interpretability by decreasing the number of nodes; accuracy must be maintained during causality map simplification; and sliders add confusion by causing significant changes in the causality map. In addition, survey results found PS-PC1 to be the most user-friendly approach, further emphasizing its potential for application in industry.

Opsomming

Wêreldwye kompetisie forseer moderne mineraalprosesseringaanlegte om by hoë produktiwiteit bedryf te word. Hierdie hoë produktiwiteit word bereik deur prosesmonitering te implementeer om die gewenste bedryfskondisies te handhaaf. ’n Fout wat in een deel van ’n aanleg ontstaan kan egter regdeur die aanleg voortplant en so die grondoorsaak verberg. Oorsaaklikheidanalise is ’n metode wat die oorsaak-en-gevolg-verhouding tussen prosesveranderlikes identifiseer en hierdie in ’n oorsaaklikheidskaart toon wat gebruik kan word om die voortplantings roete van ’n fout terug na sy grondoorsaak te volg. ’n Groot hindernis vir die wye aanvaarding van oorsaaklikheidanalise as instrument vir foutdiagnose in industrie, is die swak interpreteerbaarheid van oorsaaklikheidskaarte.

Hierdie studie het maniere om die interpreteerbaarheid van oorsaaklikheidskaarte vir foutidentifikasie te verbeter, geïdentifiseer, voorgestel en geassesseer. Alle benaderings is getoets op ’n gesimuleerde gevallestudie en die resulterende kaarte is vergelyk met ’n standaard oorsaaklikheidskaart of sy transitiewe inkrimping. Die ideale oorsaaklikheidskaart is gedefinieer en alle vergelykings is uitgevoer gebaseer op sy karakteristieke. Oorsaaklikheidskaarte is geproduseer deur kondisionele Granger-oorsaaklikheid (GC) te gebruik, met ’n nuwe heuristiese benadering om steekproefperiode en tydgleuf te selekteer.

Kondisionele GC is gevind om nie gepas te wees vir aanlegwye oorsaaklikheidanalise nie, as gevolg van groot datavereistes, swak seleksie van modelorde as AIC gebruik word, en onakkuraatheid in die teenwoordigheid van veelvoudige, verskillende verblyftye en tydvertraging. Metodes om proseskennis te inkorporeer om konneksies en potensiële grondoorsake te bedwing, is ondersoek en gevind om alle konneksies wat vals is te verwyder en die groep van potensiële grondoorsaakveranderlikes te verminder, onderskeidelik. Instrumente soos om node-ordes op die oorsaaklikheidskaart visueel te vertoon en skuiwers te inkorporeer om konneksies en veranderlikes te manipuleer is ook ondersoek.

Verder is ’n nuwe hiërargiese benadering vir aanlegwye oorsaaklikheidanalise voorgestel, waar oorsaaklikheidskaarte in twee opeenvolgende fases gebou is. In die eerste fase is ’n minder gedetaileerde aanlegwye kaart gebou deur verteenwoordigers vir groepe veranderlikes te gebruik, en is gebruik om die fout na een van daardie groepe van veranderlikes te lokaliseer. Veranderlikes is gegroepeer volgens aanlegdele of modules geïdentifiseer in die data, en die eerste hoof komponent (PC1) is gebruik om elke groep te verteenwoordig (PS-PC1 en Mod-PC1 onderskeidelik). PS-PC1 is gevind om die mees belowende benadering te wees, want sy aanlegwye kaart het duidelik die ware grondoorsaakligging geïdentifiseer, en die stap-gewyse toepassing van kondisionele GC het die vereisde aantal steekproewe beduidend verminder van 13 562 tot 602.

Laastens, ’n bruikbaarheidstudie in die vorm van ’n opname is uitgevoer om die potensiaal vir industriële toepassing van die instrumente en benaderinge voorgestel in hierdie studie, te ondersoek. Twintig antwoorde is verkry, met deelnemers wat bestaan het uit Universiteit van Stellenbosch se finale jaar/nagraadse studente, werknemers van ’n industriële IoT-firma, en Anglo American Platinum werknemers. Hoofbevindinge het ingehou dat proseskennis

noodsaaklik is; om veranderlikes te groepeer verbeter interpreteerbaarheid deur die aantal nodes te verminder; akkuraatheid moet gehandhaaf word gedurende vereenvoudiging van oorsaaklikheidskaarte; en skuiwers dra by tot verwarring deur beduidende veranderinge in die oorsaaklikheidskaart te maak. Daarmee saam het die opname se resultate gevind dat PS-PC1 die meer gebruiksvriendelike benadering was, wat sy potensiaal vir toepassing verder beklemtoon.

Acknowledgements

I would like to express my appreciation to the following people and institutions for their contributions during the execution of this project:

- My supervisors, Dr Tobi Louw, Prof Steven Bradshaw, and Prof Lidia Auret, for their guidance, motivation, and feedback throughout this project. Thank you for always wanting me to be excited about my work.
- Anglo American Platinum, for their financial support of this project, and specifically Dr Dev Groenewald, for giving me a glimpse into the mind-set of engineers on-site, and what industry is looking for.
- Mieke de Jager, for her invaluable advice from the perspective of a qualitative researcher, and for always being available to discuss any aspect of my survey.
- My flatmate, Anke, for her earnest care about my well-being, and for always being keen for a snack-run to Spar.
- My mom for her patience, kindness, and countless prayers; and my dad, for his willingness to debate fine points (even which shade of blue looks better in a plot), his insightful critiques, and his compassion throughout this project.
- My fiancé, Kyle, for always being there for me, and continuously encouraging me and building me up. Thank you for being the best.
- To God, for his grace and unconditional love throughout every moment of this project, and all the character building that took place.

Table of Contents

1. Introduction 1
1.1. Background 1
1.2. Aim and objectives 1
1.3. Research scope 2
1.4. Thesis organisation 2

2. Causality analysis for fault identification 3
2.1. Connectivity and causality in processes 3
2.2. Representing connectivity and causality 3
2.3. Capturing connectivity from process knowledge 5
2.4. Capturing causality from historical process data 5
2.5. Combining process knowledge and data-based causality 8
2.6. Using causality analysis for fault identification 9
2.7. Factors affecting performance of causality analysis 10
2.7.1. Noise in the data 10
2.7.2. Type of fault 10
2.7.3. Controller interaction 10
2.7.4. Confounding variables 11
2.7.5. Time frame and parameter selection 11
2.8. Summary of causality analysis for fault identification in literature 12

3. Causality map presentation and interpretability 13
3.1. Difficulties in causality map interpretation 13
3.2. Visualization techniques and tools for causality map interpretation 14
3.2.1. Pruning 14
3.2.2. Layout styles 15
3.2.3. Separating causality maps and assigning node and edge attributes 16
3.2.4. Node importance techniques 16
3.2.5. Complexity metrics 17
3.2.6. Graph traversal 17
3.3. Dimensionality reduction in causality maps 17
3.3.1. Dimensionality reduction using variable selection 17
3.3.2. Dimensionality reduction using feature extraction 20
3.4. Modular structures in networks 22

3.5. Summary of causality map interpretability and network visualisation in literature 25

4. Causality analysis methodology and proposed approaches 30

4.1. Workflow for causality analysis 30

4.2. Procedures for incorporating process knowledge in causality maps 31

4.2.1. Constraining connections 31

4.2.2. Constraining potential root cause variables 31

4.3. Procedures for applying tools to aid in causality map interpretation 31

4.3.1. Displaying node rankings 32

4.3.2. Sliders for connections and variables 32

4.4. Proposed hierarchical approach for causality analysis 33

4.4.1. Grouping the variables 33

4.4.2. Finding a representative for each group of variables 33

4.5. Procedure for evaluating the interpretability of causality maps 33

4.6. Causality maps usability study 35

5. Case study: Tank network simulation 37

5.1. Description of overall process and control in tank network 37

5.2. Data from tank network simulation 39

5.2.1. Measured variables 39

5.2.2. Fault: Valve stiction in plant section 2 39

6. Results and Discussion 42

6.1. Definition of the ideal causality map 42

6.2. Heuristic approach for selecting sampling period and time window 43

6.2.1. Selecting sampling period and time window for a standard causality map 43

6.2.2. Selecting sampling period and time window for the plant-wide causality map in the proposed hierarchical approach 48

6.2.3. Evaluation of heuristic approach for selecting sampling period and time window 52

6.3. Causality maps incorporating process knowledge 53

6.3.1. Constrained connections 53

6.3.2. Constrained potential root causes 55

6.4. Causality maps incorporating tools to aid interpretation 56

6.4.1. Visually displaying node rankings 56

6.4.2. Connections-slider 58

6.4.3. Variables-slider 58

6.5. Hierarchical approach for causality analysis 64

6.5.1. Plant-wide causality maps with nodes representing plant sections 64

6.5.2. Plant-wide causality maps with nodes representing modules in the data 66

6.6. Survey results 68
6.6.1. Closed questions results 68
6.6.2. Thematic analysis of open-ended questions 71
6.6.3. Summary of survey findings 72

7. Conclusions and recommendations 73
7.1. Conditional Granger causality for plant-wide fault identification 73
7.2. Process knowledge, tools, and visualisation techniques for causality map interpretability 74
7.3. Hierarchical approach for plant-wide causality analysis 75

A. Partial derivation of the spectral optimisation method 81
A.1. Partial derivation of spectral optimisation method – Part 1 81
A.2. Partial derivation of spectral optimisation method – Part 2 82

B. Author's code 83
B.1. Conditional Granger causality 83
B.2. Validating data-based causality with a connectivity matrix 86

C. Tank network model development and implementation 90

D. Modelling valve stiction 92

E. Derivation of controller tuning relations 95

F. Tank network connectivity matrix 98

G. Sampling period and time window plots 107
G.1. Plots for a standard causality map 107
G.2. Plots for the proposed hierarchical approach 110

H. Survey 116
H.1. Survey questionnaire 116
H.2. Survey causality maps 135
H.3. Additional survey results 141
H.4. Notice of ethics approval 153

Glossary

Adjacency matrix Matrix representation of a graph, where rows and columns represent nodes and binary entries represent edges.

Causality map Pictorial representation of a graph with nodes representing process variables and edges representing causal relations between variables.

Causality matrix Adjacency matrix where the binary entries have been replaced with causal strengths.

Connector hub Node within a module in a graph/network that is connected to nodes in other modules and has a high degree.

Confounding variable Variable that influences both the dependent and independent variables (x and y) causing a spurious causality ($x \rightarrow y$).

Degree Number of edges adjacent to a node in a graph.

Digraph (Directed graph) Graph where edges have arrow heads to indicate the direction of relation between nodes.

Edge Element in a graph that represents a causal relation between variables (nodes) in a causality map.

Fast Fourier transform An algorithm for computing the discrete Fourier transform of a data series at all the Fourier frequencies.

Fault Abnormal event in a process that causes measured variables or KPIs to deviate from desired values, possibly causing performance or safety degradation.

Fault detection Procedure to determine whether a fault has occurred in the process or not.

Fault diagnosis Combination of fault detection and fault identification.

Fault identification Procedure to determine the type, magnitude and location of a fault.

Graph Non-empty set of elements called nodes and a set of elements called edges.

Indegree Number of edges entering a node in a directed graph.

Modular structure Graph/network property where nodes are grouped into clusters, termed *modules*, so that there are a large number of edges connecting nodes within the module, and few edges connecting nodes from different modules.

Modularity A measure of the density of edges within a module in a graph/network compared to edges between modules.

Module A cluster of nodes in a graph/network that exhibits a modular structure.

Modular structure Structure in a graph/network, when the graph/network is divided into modules with high intra-connectivity within modules and low inter-connectivity between modules.

Network See *Graph*.

Node Element in a graph that represents a process variable in a causality map.

Optimal scaling vector Vector with entries that represent the contribution of corresponding time series (variables) to the spectral envelope at the frequency under consideration.

Outdegree Number of edges exiting a node in a directed graph.

Periodogram Estimate of the spectral density of a series..

Smearing effect An effect where a fault originating in one part of a process propagates to different parts of the process, affecting numerous variables.

Spectral envelope The largest portion of power obtainable at each frequency for any linear combination of the original time series.

Standard causality map A causality map without any tools or visualisation techniques presented in this project applied to it.

Strongly connected component A subgraph of a graph, where every edge belongs to at least one cycle.

Nomenclature

Throughout this thesis, (1) bold capital letters represent matrices; (2) bold lower case letters represent vectors; and (3) regular letters represent scalars.

Symbols	
$\boldsymbol{A}$	Adjacency matrix
c	AR coefficient
F	F-statistic
$F_{Y\to X}$	G-causality index
m	Total number of edges in a graph/network
N	Sample size
$\hat{\boldsymbol{I}}$	Periodogram
n	Number of variables
$\boldsymbol{P}$	Power spectral density matrix
p	Model order
Q	Modularity
$\boldsymbol{V}$	Covariance matrix
z	Normalised time series
Greek symbols	
α	Significance level in statistical test
B	Modularity matrix
δ	Kronecker delta symbol
ϵ	Residual error
λ	Spectral envelope
$\hat{\sigma}^2$	Estimate of variance of the error
$\boldsymbol{\kappa}$	Optimal scaling vector
ω	Frequency
ω_c	Common oscillation frequency
ω_k	Fourier frequency
Subscripts and superscripts	
Full	Full model
mod	Module
Res	Restricted model
Abbreviations and acronyms	
AAP	Anglo American Platinum
AIC	Akaike Information Criterion
AR	Autoregressive
CV	Controlled variable
DOF	Degrees of freedom
GC	Granger causality
KPI	Key performance indicator
Mod-PC1	Variables grouped according to modules in the data; PC1s as representatives
MV	Manipulated variable
OC	Oscillation contribution
OLS	Ordinary least squares

Abbreviations and acronyms

PCA	Principle component analysis
PSD	Power spectral density
PS	Plant section
PS-PC1	Variables grouped according to PSs; PC1s as representatives
RCA	Root cause analysis
RSS	Residual sum of squares
SP	Sampling period
TE	Transfer entropy
TW	Time window

CHAPTER 1

Introduction

This study investigates the interpretability of causality maps for fault identification in chemical or mineral processes.

1.1. Background

Global competition forces modern chemical and mineral processing plants to operate at a high productivity. Rather than introduce costly new initiatives, process monitoring and optimisation are used to ensure optimal use of existing processes and infrastructure. Well-functioning process monitoring should maintain plants under operating conditions required to achieve the desired productivity, with minimal time spent under faulty conditions.

Advances in informative sensor technology and analytics platforms allow large amounts of data to be captured in real-time and processed in useful time throughout industrial plants (Reis and Gins, 2017) – so allowing for advanced and large-scale process monitoring. Whenever a fault occurs in a plant, the process monitoring system should therefore be able to detect, identify and rectify it (Lindner, 2019). However, fault identification in industry can still take up to hours or days (Reis and Gins, 2017), because a fault originating somewhere in a plant can propagate throughout the plant so that numerous variables show an effect of the fault (Van den Kerkhof *et al.*, 2013).

Fortunately, research from the last decade has shown data-based causality analysis techniques to be useful for fault diagnosis (Lindner, Auret and Bauer, 2017). Causality analysis uses the cause-effect relationships between variables to construct a causality map, where the propagation path of a fault can be traced back to its root cause. However, due to poor causality map interpretability, engineers struggle to understand and use causality analysis results for fault diagnosis, and it has yet to become widely accepted in industry.

1.2. Aim and objectives

The aim of this study is to improve the interpretability of causality maps for fault identification. Recent studies indicate that the incorporation of process knowledge in causality analysis aids in identification of the root cause of a fault (Landman *et al.*, 2014; Thambirajah *et al.*, 2009). Specific emphasis is therefore placed on the incorporation of process knowledge in causality maps to improve interpretability. This aim is achieved by addressing the following objectives:

1. Identify and assess ways to incorporate process knowledge in causality maps.
2. Identify, propose, and assess tools and visualisation techniques to aid in interpretation of *plant-wide* causality maps.
3. Perform a usability study to gain insights on whether causality maps with the proposed tools and visualisation techniques can be practically applied in industry.

1.3. Research scope

This study is focussed on improving the interpretability of causality maps for fault identification. The scope of this study is limited to only the fault identification step of process monitoring, therefore excluding the fault detection and process recovery steps that take place before and after fault identification. Furthermore, fault identification in this study is limited to finding the root cause of oscillations, therefore excluding other types of faults.

1.4. Thesis organisation

This thesis is organised as follows: Chapter 2 presents the relevant background information to introduce causality analysis as a method for fault identification; and Chapter 3 focusses on causality map presentation and interpretability, highlighting the vacancies in literature regarding this topic. Chapter 4 provides the methodology for causality analysis applied in this study: including a workflow for causality analysis; the procedures for incorporating process knowledge in causality maps by constraining connections and potential root cause variables; the procedures for applying tools to aid in causality map interpretation, such as displaying node rankings and using sliders to manipulate connections and variables; a novel hierarchical approach for plant-wide causality analysis; the procedure for evaluating the interpretability of causality maps; and finally the usability study. Chapter 5 presents a description of the case study used in this work; Chapter 6 presents the results and discussion; and lastly Chapter 7 presents the conclusions and recommendations for future work.

CHAPTER 2

Causality analysis for fault identification

2.1. Connectivity and causality in processes

Chemical or mineral processes consist of elements that are connected to each other via material, energy and information paths. Both connectivity and causality refer to links between these elements in a process, where change in one variable effects change in another (Yang *et al.*, 2014).

The difference between connectivity and causality comes in with each respective meaning of direction, according to the definitions by Yang *et al.*, (2014). In the context of connectivity, direction is that of the flow of material, energy or information; while in the context of causality, it is the temporal direction of causation. In its simplest terms, causality is the cause-effect relationship between variables.

Take a tank with feedback control, where tank level is the controlled variable (CV) and the valve on the outlet line is the manipulated variable (MV). There is information flow from the tank level (CV), via instrumentation, to the valve on the outlet line (MV) – so there is connectivity $CV \rightarrow MV$. Furthermore, a change in the tank level (CV) results in a change in the position of the valve on the outlet line (MV) after some time – so there is also causality $CV \rightarrow MV$. However, a change in the position of the valve on the outlet line (MV) results in a change in tank level (CV), but there is no material, energy or information flow from the valve on the outlet line (MV) to the tank level (CV) – so there is a causal connection $MV \rightarrow CV$, but no connectivity $MV \rightarrow CV$.

2.2. Representing connectivity and causality

The connectivity and causality in a process can be represented using graph theory, allowing for both visualisation and analysis. This section therefore presents a brief overview of graph theory in relation to causality.

A *graph* is defined as a non-empty set of elements called *nodes* (or vertices), and a set of elements called *edges* (or arcs) (Price, 1971). In terms of causality, the nodes represent process variables and the edges represent the causal connections between them. A graph can be either *undirected* or *directed*, with the difference being that directed graphs have arrow heads on their edges to indicate the direction of relation between nodes (Price, 1971). Nodes on either side of an edge in a directed graph are termed as either the *source* node or the *sink* node, depending on whether the edge originates from or ends at the node in question (Price, 1971). Furthermore, the number of edges adjacent to a node is defined as the *degree* of that node, and the number of edges entering

and exiting a node in a directed graph are defined as the *indegree* and *outdegree* of the node respectively (Price, 1971).

Directed graphs can also be called *digraphs* (Yang *et al.*, 2014), and they are used to represent causality, either in pictorial view or as a matrix. The pictorial view is termed a *causality map*, with a simple example of three variables (X, Y, Z) shown in Figure 1a. The most general matrix form is an *adjacency matrix*, where the rows and columns represent the nodes and binary entries represent the edges. An entry of 1 indicates that there is an edge (causal relation) between nodes (variables) and an entry of 0 indicates that there is no edge (causal relation) between the nodes (variables) (Price, 1971). Figure 1b shows the simple adjacency matrix for the graph in Figure 1a. Generally, the rows represent source elements and the columns represent sink elements.

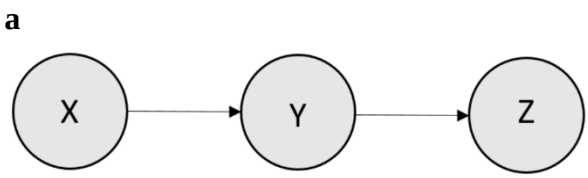


Figure 1: **a.** Example of causality map with three variables. *Nodes represent variables and edges represent casual connections.* **b.** Adjacency matrix corresponding to causality map shown in **a**.

In addition to an adjacency matrix, the causal structure of a process can also be represented as a weighted digraph, termed a *causality matrix*. This is similar to an adjacency matrix, but the matrix entries are quantitative values for the *causal strengths* (e.g. values of the G-causality index defined in Section 2.4) of each causal connection. Furthermore, it is emphasised that both adjacency matrices and causality matrices only show direct causal connections. That is why the connection from $X \rightarrow Z$ in Figure 1 has a zero-entry in the top right hand corner of the adjacency matrix, although there is an indirect causal connection from $X \rightarrow Z$ via Y.

Indirect causal connections can be presented in a *reachability matrix*, where binary entries represent paths between variables (i.e. indirect connections between variables). A reachability matrix is obtained from either an adjacency matrix or a causality matrix, according to eqn 1 (Jiang, Patwardhan and Shah, 2008):

$$\boldsymbol{R} = (\boldsymbol{A} + \boldsymbol{A}^2 + \boldsymbol{A}^3 + \cdots + \boldsymbol{A}^n)^{\#} \qquad [1]$$

Where n is the number of nodes (variables), A is the adjacency/causality matrix, and # indicates that the entries are binary (1 or 0 for true or false respectively).

Figure 2 shows the reachability matrix for the causality map and adjacency matrix presented in Figure 1. Here, the path (indirect connection) from $X \rightarrow Z$ is indicated with an entry of 1 in the top right hand corner.

$$\boldsymbol{R} = \boldsymbol{A} + \boldsymbol{A}^2 + \boldsymbol{A}^3 = \begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$$

Figure 2: Reachability matrix corresponding to the causality map and adjacency matrix shown in Figure 1.

Finally, connectivity can also be represented in matrix form, as a *connectivity matrix* (Thambirajah *et al.*, 2009). In this case, binary entries indicate whether variables are directly

connected (as opposed to whether there is a direct causal influence as is the case in an adjacency matrix), where the connectivity can be via material, energy or information flows.

2.3. Capturing connectivity from process knowledge

Process connectivity describes the paths that material, energy and information flows take through the process. Since faults spread via these connected paths, knowledge of the connections can allow a fault to be traced back to its root cause, as is further discussed in Section 2.6. This connectivity information is often available from process knowledge, but it is typically qualitative (Yang *et al.*, 2014). The existence of a connection can therefore be inferred from process knowledge, but its strength cannot.

There are various sources of process knowledge, including fundamental or empirical models, the understanding and knowledge of a plant expert, and process schematics such as piping and instrumentation diagrams (P&IDs). Plant expert knowledge along with models or process schematics can be used to manually construct adjacency matrices or graphical models such as signed digraph models (SDGs), where SDGs are simply digraphs (see Section 2.2) with a '+' or '-' assigned to each edge to represent a positive or negative influence of the source node on the sink node (Yang *et al.*, 2014). Unfortunately, these are time-consuming and complicated tasks. Furthermore, process models are often not available or accurate enough to capture the causal structure of a process (Maurya, Rengaswamy and Venkatasubramanian, 2003).

Plant schematics such as P&IDs therefore look to be the most promising source of process knowledge from which to infer connectivity, especially since the process does not need to be manual. Extensive research has been performed to automate this procedure (Arroyo Esquivel, 2017; Yim *et al.*, 2006; Thambirajah *et al.*, 2009; Yang *et al.*, 2014). Connectivity information is extracted from P&IDs by first converting them to eXtensible Markup Language (XML) files with either Computer Aided Engineering Exchange (CAEX) or XMPlanet schemas (Thambirajah *et al.*, 2009). These XML files are then used to construct connectivity matrices (see Section 2.2).

A major drawback of this method is that plants rarely have updated copies of P&IDs available. Furthermore, often only hard copies of process schematics exist. Fortunately, a tool has been developed that can scan hard copies of process schematics to capture connectivity (Arroyo Esquivel, 2017).

2.4. Capturing causality from historical process data

In addition to the extraction of connectivity from process knowledge, causality information can be extracted from the process data captured by the numerous sensors present in a plant. The values of these continuously or even intermittently measured process variables are stored in the form of time series, and the relationships between these time series can be exploited to infer causality between variables (Yang *et al.*, 2014). A mathematical definition of causality proposed by Wiener states that "X could be termed as to 'cause' Y if the predictability of Y is improved by incorporating information about X" (Wiener, 1956).

There are various techniques that infer causality from process data, with examples including Granger causality, transfer entropy, cross-correlation, and partial directed coherence (Lindner, 2019). The two most commonly used techniques are Granger causality (GC) and transfer entropy (TE) (Lindner, 2019). However, TE is an information-theoretic technique, whereas GC is based

on simple regression statistics. For this reason, engineers can reason more easily about results obtained from GC than from TE. This section therefore presents a brief overview of GC to infer causal relationships from historical process data.

GC is a practical adaptation of Wiener's definition of causality, stating that "$x(t)$ is causing $y(t)$ if we are better able to predict $y(t)$ using all available information than if the information apart from $x(t)$ had been used" (Granger, 1969). It makes use of linear autoregressive (AR) models of stochastic variables to quantify and then compare prediction error.

Consider two time series of variables $x(t)$ and $y(t)$. The time series $y(t)$ can be predicted using a linear AR model that includes past values of only y itself, known as the *restricted model* [eqn 2] (Barnett and Seth, 2014). The time series $y(t)$ can also be predicted using a linear AR model that includes past values of both y and x, known as the *full model* [eqn 3] (Barnett and Seth, 2014). If the prediction error of the full model is smaller than that of the restricted model, x is said to Granger-cause y. This is known as *unconditional Granger causality*.

$$y(t) = \sum_{k=1}^{p} c_{y,res}(k)y(t-k) + \epsilon_{res}(t) \quad [2]$$

$$y(t) = \sum_{k=1}^{p} \{c_{y,full}(k)y(t-k) + c_x(k)x(t-k)\} + \epsilon_{full}(t) \quad [3]$$

The *c*'s are regression coefficients, which can be determined using the ordinary least squares (OLS) approach (Hill, 2011), and p is the model order, which represents the number of lagged historical data points to include. A commonly used approach to select the model order is the standard Akaike Information Criterion (AIC) (Akaike, 1974) to ensure the best model fit while preventing overfitting due to noise present in the data [eqn 4] (Duan, 2014). However, there is typically a limited sample size available for fault identification in the chemical or mineral processing industry, which can be addressed by implementing the finite sample bias-corrected form of AIC [eqn 5] as proposed by (Hurvich and Tsai, 1989) to counteract bias in small samples:

$$AIC(p) = \ln(\det(\Sigma)) + \frac{2pn^2}{N} \quad [4]$$

$$AIC_c(p) = \ln\left(\det\left(\frac{\Sigma}{N-p-1}\right)\right) + \frac{2pn^2N}{N-pn^2-1} \quad [5]$$

Where Σ is the residual covariance matrix of the full model, n is the number of variables, and N is the number of observations.

GC was quantified as a log-likelihood ratio statistic by Geweke, (1982), which is now termed the *G-causality index* [eqn 6] . The log-likelihood ratio compares the goodness of fit of two competing statistical models, where the one was found by maximisation over the entire parameter space (i.e. the full model in eqn 3, where both x and y were included) and another by imposing some constraint that represents the null hypothesis (i.e. the restricted model in eqn 2, where $c_x(1) = c_x(2) = \cdots c_x(p) = 0$). If the null hypothesis is true (i.e. if $c_x(1) = c_x(2) = \cdots c_x(p) = 0$, meaning that there is zero causality), then the two likelihoods should not differ significantly – so the log-likelihood ratio should not be significantly different from zero. Furthermore, G-causality may be meaningfully interpreted as a quantitative measure, as it is asymptotically equivalent to information-theoretic TE for a large class of joint processes (Barnett and Bossomaier, 2013), and exactly equivalent in the Gaussian case (Barnett, Barret and Seth, 2009).

$$F_{x\to y} = \ln\left(\frac{var(\epsilon_{res})}{var(\epsilon_{full})}\right) \quad [6]$$

Unconditional GC can report spurious causalities if there is a set of *confounding variables* z, which influence both x and y (Barnett and Seth, 2014). For example (Figure 3), if there is no causal influence $x \to y$, but both x and y are dependent on z, then a spurious causality $x \to y$ may be reported. These spurious causalities may be eliminated by 'conditioning out' the confounding variables, provided they are measured variables, as initially proposed by Geweke, (1982). The restricted and full models are then expanded to include the set of variables z, as shown in eqns 7 and 8 (Barnett and Seth, 2014), and the G-causality index is defined as in eqn 9 (Barnett and Seth, 2014). This is known as *conditional Granger causality*.

$$y(t) = \sum_{k=1}^{p} c_{y,res}(k)y(t-k) + c_{z,res}(k)z(t-k) + \epsilon_{res}(t) \quad [7]$$

$$y(t) = \sum_{k=1}^{p}\{c_{y,full}(k)y(t-k) + c_{z,full}(k)z(t-k) + c_{x,full}(k)x(t-k)\} + \epsilon_{full}(t) \quad [8]$$

$$F_{x\to y|z} = \ln\left(\frac{var(\epsilon_{t,res})}{var(\epsilon_{t,full})}\right) \quad [9]$$

GC has some clear advantages, such as the simplicity of its calculations which make it computationally inexpensive, and its basis in the well-understood concept of regression which make the results easy to interpret (Lindner, 2019). A disadvantage of GC is the dependency on the accuracy of linear AR models, while linearity may not always be a valid assumption (Yang *et al.*, 2014).

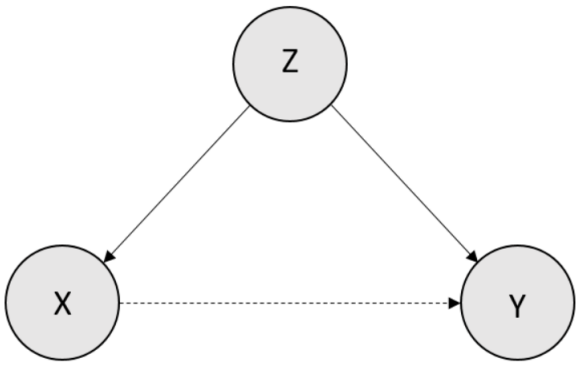


Figure 3: Example of spurious causal connection $X \to Y$ reported due to confounding variable Z. *Solid lines represent actual causal connection; dashed line represents spurious causal connection.*

Statistical significance test

Statistical significance must be established before any causal influence can be reported. This is done with a hypothesis test, where the null hypothesis states that the c_x's are zero (i.e. no causal influence $x \to y$). Statistical significance for GC is determined by applying the F-statistical test. This involves calculating the F-statistic [eqn 10] (Bressler and Seth, 2011) and comparing it to the critical F-value (F_{crit}).

$$F = \frac{\frac{RSS_{res} - RSS_{full}}{p}}{\frac{RSS_{full}}{N - 2p - 1}} \quad [10]$$

In GC analysis, the numerator degrees of freedom (DOF) are p and the denominator DOF are $(n\text{-}2p\text{-}1)$. RSS_{res} and RSS_{full} are the sum of squares of residuals of the restricted and full model respectively, and F_{crit} is the F-statistic with a significance level α. If $F > F_{crit}$ and the p-value is smaller than α, the null hypothesis can be rejected and a causal influence reported.

Significance levels α used in literature are typically 0.01 or 0.05, but this is simply convention. The significance level is essentially a threshold for which causal connections are strong enough to be displayed on the causality map and which are not. Different processes may require different thresholds for connections strength to obtain the causality map closest to the ideal one (see Section 6.1). The use of this threshold as a lever or hyper-parameter that can be manipulated to obtain a more interpretable causality map for a specific system is therefore identified as a gap in literature.

2.5. Combining process knowledge and data-based causality

Connectivity and causality can be inferred using both knowledge-based and data-based methods, as explained in Sections 2.3 and 2.4 respectively. The two approaches each have their own advantages and disadvantages (Table 1) which appear to be complimentary, allowing improvement via a hybrid approach.

Table 1: Comparison of knowledge-based methods and data-based methods. *Clear blocks indicate advantages and shaded blocks indicate disadvantages.*

Knowledge-based methods	**Data-based methods**
Reports only true connections	May report spurious connections
Reports only connectivity & known causality	May detect non-intuitive connections (e.g. connections due to energy flow)
Reports *static* causal structure: o Does not account for changing operating conditions/presence of fault o Based on potentially outdated information, e.g. P&IDs	Reports *dynamic* causal structure: o Can be applied to the latest available data to reveal the current causal structure of the process
Lower computational burden, when automated (Arroyo Esquivel, 2017)	Higher computational burden

A hybrid approach combining knowledge-based and data-based methods has therefore been attempted by some researchers. This has been done in two distinctive ways, both of which effectively constrain the paths in the reported causal structure (generally in the form of a causality map). The first way is to specify process connectivity and only consider those connections when performing data-based methods (i.e. validate knowledge-based connections using data-based methods). This reduces the amount of variable pairs to test for causality which reduces the computational burden, as well as ensure that indirect and shortcut connections (see Section 3.2.1) are ignored. Landman *et al.*, (2014) validated connectivity from XML schematics with GC, and Landman and Jämsä-Jounela, (2016) did the same with transfer entropy. Both papers report a decrease in spurious connections due to the inclusion of process knowledge.

The second way is to calculate causality using data-based methods and then use process connectivity for validation. Thambirajah *et al.*, (2009) identified possible root causes using transfer entropy and then used connectivity from XML descriptions to determine the most likely

root cause by investigating which propagation paths were physically possible. (Landman, 2019) took this a step further by only validating data-based causal connections in a controller causality map when there was at least one *direct pathway* from the source node to the sink node (see Section 2.2 for definitions of source and sink nodes) according the connectivity matrix. For this, Landman, (2019) used the definition from Jiang, Patwardhan and Shah, (2008), which states that a direct path from node i to node j exists if the output of controller i directly affects controller j without intermediate effect on any other controller.

Although hybrid approaches combining knowledge-based and data-based methods have been investigated, this was done with the aim of improving the accuracy of causality analysis results. The approaches were therefore applied to simple case studies with few variables where interpretability of the resulting causality map was not an issue. Application of these hybrid approaches on plant-wide systems with numerous variables and the effect on causality map interpretability is therefore identified as a gap in current literature. This gap is promising, as application of these hybrid approaches can decrease the number of connections to decrease graph density (see Section 2.7.1), as well as increase the trust that engineers put in causality maps, as they are not simply the result of noise-filled data but also include understandable process knowledge.

2.6. Using causality analysis for fault identification

Causality analysis can be used for various applications, one of which is process monitoring (Lindner, 2019). This section presents a brief overview of process monitoring with emphasis on fault identification, and how causality analysis can be applied in this context.

Process monitoring consists of three distinguishable steps, namely fault detection, fault identification, and process recovery (Lindner, 2019). Fault detection is the procedure to determine whether a fault has occurred in the process or not, and techniques for this have been researched and improved to the extent that a fault can be detected within an order of seconds or minutes after occurrence (Reis and Gins, 2017). Fault identification is the procedure to determine the type, magnitude and location of the fault, and can take up to hours or days (Reis and Gins, 2017). In other words, a bottleneck occurs in the fault identification step within process monitoring (Reis and Gins, 2017). Improvement in fault identification techniques such as causality analysis would therefore benefit industry, as it would decrease the time for which a fault is present in a plant.

Within the context of fault identification, causality analysis can be used in root cause analysis (RCA), which refers to determining the location of a fault. For example, if the process recovery (e.g. mineral recovery in a mineral processing plant) decreases, causality analysis can aid engineers in finding the root cause of this detrimental change in the process so that it can be corrected. This is useful, as the root cause of a fault in modern chemical and mineral processing plants is often hidden as a result of something called the “smearing effect”.

The “smearing effect” is where a fault originating in one part of the process propagates to different parts of the process (Van den Kerkhof *et al.*, 2013), causing numerous measured variables to show the effect of the fault. This is due to the high interconnectivity of process units, equipment, material flow, energy flow, and information flow, which is complicated even further by recycle streams, complex control strategies and control interaction (Lindner, 2019).

The purpose of a causality map in fault identification is therefore to identify the fault location by allowing the propagation path to be traced back to the root cause. Generally, a variable shown to

have a large causal influence is taken as the root cause. However, the smearing effect often allows numerous variables to show significant causal influences, creating a challenge in identifying a single root cause or even localising the fault to a single process unit.

Significant input is therefore required from engineers to interpret results to identify or localise a root cause, which is largely why fault identification is such a slow process. Causality map interpretability was therefore identified as an area that requires improvement by Lindner, (2019).

2.7. Factors affecting performance of causality analysis

Causality analysis can be used to identify the root cause of a fault, as discussed in 2.6. However, various factors can affect the performance of data-based causality analysis when used for this purpose, including: noise in the data, the type of fault present, controller interaction, confounding variables, and time frame and parameter selection.

2.7.1. Noise in the data

A causal connection exists when variation in one variable (input variable) causes variation in another variable (output variable) after some time. However, if there is sensor noise (i.e. uncertainty in the measurement of the property) in the input or output variable - meaning that there is variation in the output variable that was not caused by variation in the input variable - then the causal connection can be obscured and may not be detected by data-based causality techniques.

Lindner, Auret and Bauer, (2017) investigated the effect of noise on GC. The results indicated that GC is not sensitive to sensor noise. This is likely explained by the fact that GC fits *linear* models to the data, assuming that the errors follow a Gaussian distribution – so the models may have a large bias, but they should always have low variance.

2.7.2. Type of fault

Causality analysis has been used to identify the root cause of different types of faults, but most literature applies it to identify the root cause of oscillations. Oscillatory behaviour is often exhibited by chemical and mineral process due to poor controller tuning, control valve stiction, poor process and control system design, and oscillatory disturbances (Miao and Seborg, 1999; Thornhill, 2005). These oscillations tend to persist as they propagate throughout a process, resulting in long datasets that can be used for causality analysis, and making causality analysis well-suited for the diagnosis of oscillations (Lindner, 2019). All further discussions are therefore limited to causality analysis to identify the root cause of oscillations.

Lindner, Auret and Bauer, (2017) investigated the effect of different oscillations on a known causal connection between two variables. A peak in the G-causality index was observed for intermediate frequencies, while both high frequencies and low frequencies resulted in lower G-causality indices. At high frequencies, the perturbations resembled sensor noise, which obscured the causal connection. At low frequencies, propagation of the input perturbation to the output variable was so gradual that the effect was dominated by the more rapid dynamics of the sensor noise. However, this effect was somewhat compensated for with the model order selection via AIC (see Section 2.4), as a larger model order could capture slower dynamics.

2.7.3. Controller interaction

Lindner, Auret and Bauer, (2017) report that the presence of a controller in a process decreases the G-causality index of a known causal connection, specifically at lower frequencies. This is because lower frequency oscillations are more successfully attenuated by the controller, so the

impact of these oscillations on the output are less pronounced, which hides the causal connection. However, Lindner, Auret and Bauer, (2017) also reports that GC was still able to detect a known causal connection in a closed loop, even for a low frequency oscillation.

2.7.4. Confounding variables

Data-based causality techniques can report spurious connections if a confounding variable is present, as discussed in Section 2.4. Conditional GC can be used to prevent this, except if the fault is due to an unmeasured variable. This is a limitation to all data-based causality analysis techniques, as they can only use the information provided by measured variables.

2.7.5. Time frame and parameter selection

GC is a lag-based causality analysis technique (Yang *et al.*, 2014), meaning that it is based on temporal information. The performance of GC analysis is therefore dependent on parameters that affect how this temporal information is included in the analysis.

The time frame of the data included in the analysis should therefore be long enough to incorporate the process dynamics, which includes residence times and time delays of the process. The time frame is the product of the model order and the sampling period, where the model order is typically selected via AIC (see Section 2.4). Most chemical and mineral processes have unit residence times of the order of at most a few minutes, and sampling frequencies of around 10 seconds (Lindner, 2019), which gives typical model orders of around 1 – 30.

Sampling period (SP) refers to the time between samples when the data is sub-sampled from the original time series. A larger SP can be used to cover a larger time frame with fewer samples, but it also means that the included data is of a lower resolution, which can negatively affect causality analysis results. Barnett and Seth, (2017) reports detectability “black” and “sweet” spots as the sampling period interacts with the underlying process dynamics; but an exponential decay in the detection ability of causality analysis techniques as the sampling period increases beyond the causal delay.

Bauer *et al.*, (2007) investigated the minimum number of samples required to obtain a significant Transfer entropy value. Results indicated that the minimum number of samples should be set to 2000 samples if possible, but could be as low as 400 samples; and Lindner, (2019) agrees with this. To the best of the author’s knowledge, no investigation has been performed to find the minimum number of samples required to obtain a significant value for the G-causality index. However, the full autoregressive (AR) model in conditional GC predicts all variables as a function of all the other variables (see Section 2.4). The minimum number of samples included in causality analysis should therefore be sufficient to find a unique solution for the full AR model (Barnett and Seth, 2014).

SP and time window (TW) selection is therefore important to provide a long enough time frame and a sufficient number of samples for conditional GC, where TW is the total length of time captured in the dataset. However, to the best of the author’s knowledge, there is no guideline or heuristic approach for engineers to follow in SP selection for GC, and no mention of TW selection for causality analysis in literature. Typically, no decision is made regarding TW, and all available data with the fault present is used for causality analysis. However, the amount of available data may not be enough to provide a unique OLS solution for the full AR model, especially if the number of available samples is decreased by sub-sampling. A usable method or guideline for SP and TW selection for GC is therefore identified as a vacancy in current literature.

2.8. Summary of causality analysis for fault identification in literature

This section provides an overview of how causality analysis has been applied for fault identification in published literature regarding process engineering. In line with the scope of this work, Table 2 presents a summary of literature where GC was performed for fault identification of an oscillation. For a broader summary, including different causality analysis techniques and fault types, the reader is referred to Lindner, (2019).

Table 2: Summary of literature on Granger causality application for fault identification of an oscillation.

Source	Industry	Data type
Yang, Sirish and Xiao, (2010)	Chemical processing	Industrial
Duan *et al.*, (2014)	Chemical processing	Simulated
Landman *et al.*, (2014)	Chemical processing	Industrial
Yuan and Qin, (2014)	Chemical processing	Simulated & Industrial
Yang *et al.*, (2014)	Chemical processing	Industrial
Wakefield *et al.*, (2018)	Mineral processing	Simulated
Lindner, (2019)	Mineral processing	Simulated & Industrial
Suresh, Sivaram and Venkatasubramanian, (2019)	Chemical processing	Industrial

The sources are categorised according to industry in the second column in Table 2, showing that the majority of applications are in the chemical processing industry, with few in the mineral processing industry. To the best of the author's knowledge, no other examples of GC application in the mineral processing industry have been published, but it should be noted that the publications with mineral processing applications are recent, so more publications of causality analysis in this industry can be expected.

In addition, the third column in Table 2 categorises the sources according to data type, showing that GC has not only been applied to simulated case studies, but also to industrial case studies. However, it should be noted that the focus of this research was on the development and improvement of the GC technique, with little or no reference to usability or interpretability – as is further discussed in Section 3.5.

CHAPTER 3

Causality map presentation and interpretability

3.1. Difficulties in causality map interpretation

The output of causality analysis for fault identification is a causality map. This causality map is interpreted by engineers to identify the variable(s) at the start of the propagation path as the root cause variable(s) to be further investigated. However, this interpretation step can prove difficult.

Firstly, the true root cause variable(s) may not be included in the causality map, as data-based causality maps are limited to measured variables. Furthermore, the causality map may be ambiguous if there are multiple propagation paths pointing to different root cause variables, as mentioned in Section 2.6. This happened in the study by Landman, (2019), where further analysis was then required to identify the root cause variable from the group of potential root cause variables identified by the causality map. The variables that could not reach all other variables in the causality map were excluded from consideration as root cause variables, allowing one root cause variable to be identified. However, this approach may not always narrow the pool down to only one root cause variable.

Alternatively, the fault can be localised if all the potential root cause variables are associated with a single unit or control loop in the plant, and that unit or control loop can be further investigated (Lindner, 2019). However, the variables may all be associated with different units or even sections of the plant, so that there is no clear indication where to further investigate the fault. This could be the result of multiple faults occurring at the same time, but Shiozaki *et al.*, (1985) states that the probability of independent faults occurring simultaneously is small. It is therefore assumed that only one independent fault is present at a time.

Causality map interpretability also depends on the accuracy of the causality technique (GC, TE, etc.) used as well as system intricacy and interactivity, as this has an effect on the number of spurious connections present (Lindner, 2019). System intricacy and interactivity play another role, as even a large number of true connections will lead to confusion. In addition, causality map interpretability is influenced by whether the causality map is cyclic or acyclic. All root causes are identifiable in acyclic maps (Pearl, 2009), but cyclic maps show no clear start or end nodes to allow tracing back to a root cause.

Suresh, Sivaram and Venkatasubramanian, (2019) therefore propose an approach to produce an acyclic plant-wide causality map from a cyclic causality map. Variables that cause cyclic effects, such as variables related to control loops and recycle streams, are simply excluded from the causality analysis. However, this approach is time-intensive, as these variables are manually identified from a P&ID, which will likely be outdated (see Section 2.3).

3.2. Visualization techniques and tools for causality map interpretation

Currently little research has been done on causality map interpretation, in favour of focusing on improving the accuracy of the analysis methods such as GC and TE (Lindner, 2019). However, Lindner, (2019) does provide a starting point in the form of basic techniques to construct more interpretable causality maps; as well as post-construction tools for root cause identification. This section therefore provides an overview of these techniques and tools, as well as a workflow for causality map interpretation proposed by Lindner, (2019) in Figure 4.

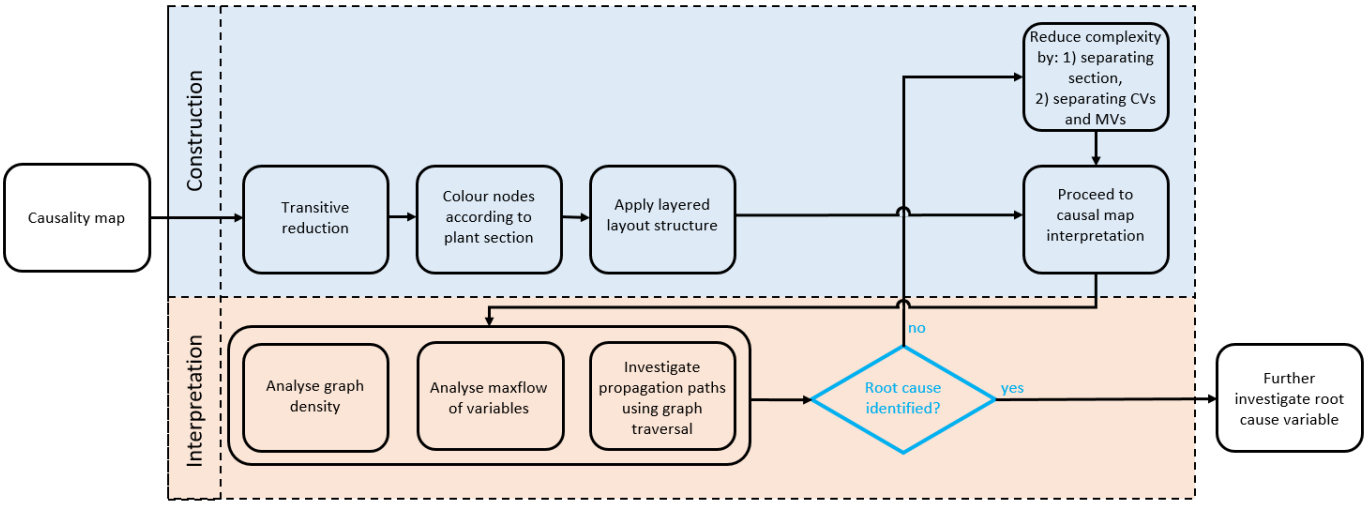


Figure 4: Workflow for causality map interpretation for fault identification, adapted from Lindner, (2019). *Techniques and tools mentioned are described in this section.*

3.2.1. Pruning

Pruning algorithms can be used to decrease the total number of connections in causality maps by removing certain connections. Commonly removed connections are *shortcut connections* (Lindner, 2019), which are causal connections due to the influence of an intermediate variable (Figure 5). This is because these shortcut connections can be true or spurious. In other words, the connection from $X \rightarrow Y$ in Figure 5 may not really exist, but only appear to exist because there are connections from $X \rightarrow Z$ and $Z \rightarrow Y$. Conditional GC (see Section 2.4) could be used to determine if these connections are true or spurious, but this increases the computational burden of the causality analysis, so pruning algorithms are often preferred.

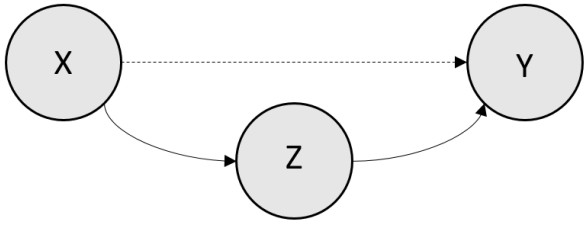


Figure 5: Example of shortcut connection from $X \rightarrow Y$. *Solid lines represent direct connections; dashed line represents shortcut connection.*

The transitive reduction algorithm is a formalized pruning algorithm for the removal of shortcut connections. The transitive reduction of graph G is another graph $G_{reduction}$ with the same number

of nodes, but with the fewest edges that still allow $G_{reduction}$ to have the same *reachability* (defined in eqn 1 in Section 2.2.) as G (Aho, Garey and Ullman, 1972). Reachability is the ability to get from one node to another within the graph.

The transitive reduction algorithm iterates through the graph using a depth-first search. The depth-first search algorithm begins at a start node that is specified by the user and discovers other nodes sequentially (Tarjan, 1972). This continues until the algorithm encounters a node with no undiscovered neighbour. The algorithm then backtracks along the discovered path to the closest previously discovered node that still has an undiscovered neighbour (Tarjan, 1972). This is implemented recursively until all the nodes that are reachable from the start node have been discovered (Tarjan, 1972). For each combination of three nodes X, Y, and Z discovered, if there are edges $X \rightarrow Z$, $Z \rightarrow Y$, and $X \rightarrow Y$, the edge $X \rightarrow Y$ is removed. The dashed line in Figure 5 would therefore be removed (Aho, Garey and Ullman, 1972).

Cyclic and acyclic graphs respond differently to the transitive reduction algorithm. The transitive reduction $G_{reduction}$ of an acyclic graph will always be a unique *subgraph* of G (Lindner, 2019), where a subgraph of G is defined as a graph that contains a subset of nodes of G and all of the edges of G that join the nodes of that subset (Price, 1971).

However, the transitive reduction for a cyclic graph G is not unique and not necessarily a subgraph of G (Lindner, 2019). Edges could be constructed that did not exist in the original graph. This is because the presence of a cycle in the original graph may have caused a connection between two specific nodes. When that cycle is removed, the connection between those two nodes is also removed, so a new edge is constructed to directly connect those two nodes in $G_{reduction}$.

3.2.2. Layout styles

Five different node layout styles are identified by Lindner, (2019). Firstly, the nodes can all be placed in a straight line, with the source nodes first and the sink nodes last, as done by Bauer *et al.*, (2007). This style can allow for quick identification of the root cause variable, but it becomes confusing for dense maps.

Secondly, a grid layout can be used, as by Landman *et al.*, (2014). This style is neat and can easily be used to mirror a plant layout or process schematic, but it can be difficult to visualize the direction of propagation (Lindner, 2019). The third style is a layered layout, as used by Sugiyama, Tagawa and Toda, (1981). This style consists of a hierarchal structure where sets of nodes make up layers, and each edge joins two nodes that belong to different layers. If the causality map has a clear start and end node, this style effectively reveals the propagation path (Lindner, 2019).

Fourth is the circle layout, as used by Yuan and Qin, (2014). It was also used by Duan *et al.*, (2015) to present results before a layered map was constructed to get a better visualization of the propagation paths. In this style, the nodes are placed in a circle around the origin, which allows for visualization of the node importance by looking at indegree and outdegree of nodes (Lindner, 2019). In turn, this allows identification of nodes as source or sink nodes respectively. This is especially useful for cyclical graphs where there is no clear start and end node.

Lastly, there is a force layout. Endpoints of edges are assigned attractive forces, and nodes are assigned repulsive forces (Lindner, 2019). This causes nodes that are not connected by an edge to be pushed away from each other due to their like forces; and nodes with many edges associated with them to attract the nodes they are connected to. The final layout therefore consists of clusters of connected nodes, which can reveal a hierarchal structure (Lindner, 2019).

Lindner, (2019) also proposes some recommendations of when to use which layout. In essence, the suggestion is to use the layered layout for sparse causal maps with minimal cyclic loops; to use the circle layout for dense maps with cyclic loops; and to use the force layout to identify important causal structures within the process.

3.2.3. Separating causality maps and assigning node and edge attributes

Complicated causality maps may be made easier to follow if nodes are grouped and coloured according to sensor or variable location. This can be done for various resolutions: namely on a plant-wide scale, where nodes are coloured according to different sections of the plant; for a single plant section, where nodes are coloured according to different units; or for the smallest resolution, where nodes are coloured according to which controller they are associated with (Lindner, 2019).

Alternatively, or in conjunction with node colouring, separate causality maps can be generated for specific groups of nodes. The example that Lindner, (2019) provides is two parallel banks of flotation cells that can be separated since they do not influence one another. This approach avoids spurious connections between the two banks and decreases the computational burden. It also leads itself to dimensionality reduction in causality maps, as further discussed in Section 3.3.

Furthermore, nodes can also be grouped, and then coloured or separated according to variable categories. These categories can be sensor type (flow, density or temperature measurements), or control categories (CVs, MVs or SPs) (Lindner, 2019).

Lastly, edge attributes can be assigned in the form of edge line thickness or edge arrow head size to represent connection strength.

3.2.4. Node importance techniques

Node importance techniques are aimed at identifying the most likely root cause variable, especially if there are multiple options presented by the causality map or if the map itself is dense. Lindner, (2019) mentions three main node importance techniques, with the first being maxflow.

Maxflow of a node is defined as the difference between the node's *outdegree* and *indegree* [eqn 11] (Yuan and Qin, 2014), which are defined as the number of edges leaving the node and the number of edges entering the node respectively (Price, 1971). The maxflow is calculated for each node and the node with the highest maxflow is considered the root cause variable. The motivation behind this is that the root cause variable would have the largest influence on the subsequent nodes (i.e. the highest maxflow) because it is the driver of the fault.

$$Maxflow = Outdegree - Indegree \quad [11]$$

The second technique is to sum the causal influence of one variable on all other variables in the system [eqn 12]. Kühnert and Beyerer, (2014) assigned a causal strength scaled between 0 and 1 to each node and then applied eqn 12, where $q_{x_j \to x_i}$ is the causal strength from x_j to x_i. This was done for each variable (or node), and a root cause priority list generated.

$$RC_j = \sum_{i=1}^{i \neq j} q_{x_j \to x_i} \quad [12]$$

Lastly, the *page rank algorithm* (Bryan and Leise, 2006) can be used, as done by Streicher, Wilken and Sandrock, (2014). The algorithm assigns scores to each node based on its influence on the rest of the causal network. The difference between this and maxflow is that the causal influence on all the other variables in the system is taken into account, instead of only on the variables (or nodes) directly surrounding the node in question.

3.2.5. Complexity metrics

Complexity metrics can prove useful to make decisions within an automated algorithm for the construction of causality maps. For example, different layout styles can be employed based on the density of the causality map (see *Layout styles* subsection earlier in this section). Lindner, (2019) proposes using *graph density* (Gibbons, 1985) for a directed graph as a complexity metric, as shown in eqn 13:

$$D = \frac{E}{V(V-1)} \quad [13]$$

Where E represents the number of edges and V represents the number of nodes in the graph. For a fully connected graph, where all nodes are connected to all other nodes via edges (Price, 1971), $E = V(V-1)$. The upper boundary of graph density is therefore one.

3.2.6. Graph traversal

Graph traversal is aimed at identifying paths between different nodes in the causality map. As with pruning, the depth-first search algorithm is typically used to move through the graph (see the *Pruning* subsection of this section) (Yang and Xiao, 2012). Different fault propagation paths can then be compared to identify the most likely candidate (Yang and Xiao, 2012). For example, causal strengths assigned to edges can be inversed and used to find the shortest path (which would represent the strongest causal link) between nodes. Shiozaki *et al.*, (1979) used the depth-first search algorithm to find the maximum *strongly connected component* in a signed directed graph (SDG) and took that component as a localized area of the root cause. A strongly connected component contains a path between all pair of nodes (Price, 1971).

3.3. Dimensionality reduction in causality maps

The number of nodes present in a causality map affects its interpretability. Section 3.2 discusses a few techniques and tools to improve causality map interpretability both during and post construction, but none of these addresses the difficulty of interpreting a causality map that has many nodes. This section therefore discusses how dimensionality reduction can be used to improve causality map interpretability, using either variable selection or feature extraction.

3.3.1. Dimensionality reduction using variable selection

Dimensionality reduction using variable selection implies that only certain types of variables are selected to include in the causality analysis and therefore the causality map. Lindner et al., (2018) did this by selecting only CVs and only MVs, and constructing separate causality maps for each group of variables, where the CV or MV in each case can be seen to represent an entire control loop.

Furthermore, variable selection using oscillation detection techniques is a known, although optional, step that can be performed prior to causality analysis. Typically, variables that oscillate at a common frequency are identified in the time series and included in the causality analysis. Landman and Jämsä-Jounela, (2016) did this by selecting variables that exhibited a peak in their power spectra at a common oscillation frequency. However, Duan et al., (2014) states that simply looking at the power spectra of variables raises the question of what is a significant oscillatory contribution at a frequency (i.e. what is a large enough peak for a variable to be selected); and opted for variable selection via the spectral envelope method developed by Jiang, Choudhury and Shah, (2007) instead. The rest of this section is therefore dedicated to discussing how the spectral

envelope can be used to (1) automatically detect a common oscillation frequency and (2) identify variables oscillating significantly at that frequency.

Concept of the spectral envelope

First, let $\boldsymbol{x}(t)$ be a multivariate time series [eqn 14], and $\boldsymbol{z}(t)$ the normalised, multivariate time series [eqn 15] such that:

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \quad [14]$$

$$\boldsymbol{z}(t) = \begin{bmatrix} \frac{x_1(t) - \mu_{x_1}}{\sigma_{x_1}} \\ \frac{x_2(t) - \mu_{x_2}}{\sigma_{x_2}} \\ \vdots \\ \frac{x_n(t) - \mu_{x_n}}{\sigma_{x_n}} \end{bmatrix} = \begin{bmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_n(t) \end{bmatrix} \quad [15]$$

Furthermore, $\boldsymbol{V_z}$ is the covariance matrix of $\boldsymbol{z}(t)$, and $\boldsymbol{P_z}(\omega)$ is the power spectral density (PSD) matrix of $\boldsymbol{z}(t)$, where the PSD of a signal describes its power distribution over frequency (Jenkins and Watts, 1968).

A scaled series $y(t, \boldsymbol{\kappa})$ can be written as a linear combination of the rows of $\boldsymbol{z}(t)$ [eqn 16], such that:

$$y(t, \boldsymbol{\kappa}) = \boldsymbol{\kappa}^* \boldsymbol{z}(t) \quad [16]$$

Where $\boldsymbol{\kappa}$ is a real or complex vector, and the symbol * refers to the conjugate transpose, meaning that the transpose of $\boldsymbol{\kappa}$ is taken and then the complex conjugate of each entry is taken. $V_y(\boldsymbol{\kappa})$ [eqn 17] and $P_y(\omega, \boldsymbol{\kappa})$ [eqn 18] are the variance and PSD matrix of $y(t, \boldsymbol{\kappa})$ respectively, which can be represented as:

$$V_y(\boldsymbol{\kappa}) = \boldsymbol{\kappa}^* \boldsymbol{V_z} \boldsymbol{\kappa} \quad [17]$$

$$P_y(\omega, \boldsymbol{\kappa}) = \boldsymbol{\kappa}^* \boldsymbol{P_z}(\omega) \boldsymbol{\kappa} \quad [18]$$

The spectral envelope $\lambda(\omega)$ is then defined as [eqn 19]:

$$\lambda(\omega) = \underset{\boldsymbol{\kappa} \neq 0}{sup} \left\{ \frac{\boldsymbol{\kappa}^* \boldsymbol{P_z}(\omega) \boldsymbol{\kappa}}{\boldsymbol{\kappa}^* \boldsymbol{V} \boldsymbol{\kappa}} \right\} \quad [19]$$

Where $\boldsymbol{V} = diag(\boldsymbol{V_z})$, and sup stands for supremum, which is the smallest number that is greater than or equal to every number in the set $\boldsymbol{\kappa} \neq 0$. The spectral envelope $\lambda(\omega)$ represents the largest portion of power (or variance) obtainable at each frequency for any linear combination of the original time series.

A constraint is imposed on $\boldsymbol{\kappa}$, so that the scaled series has unit variance [eqn 20]:

$$V_y(\boldsymbol{\kappa}) = \mathbf{1} = \boldsymbol{\kappa}^* \boldsymbol{V} \boldsymbol{\kappa} \quad [20]$$

Accordingly, the spectral envelope $\lambda(\omega)$ then represents the largest portion of power (i.e. variance) at each frequency for any linear combination of the original time series that has unit variance. Furthermore, the time series data has been normalised, meaning that all non-zero entries of $\boldsymbol{V}$ (i.e. variances) are ones, and the constraint simplifies to eqn 21:

$$1 = \boldsymbol{\kappa}^{*}\boldsymbol{\kappa} \quad [21]$$

With this constraint [eqn 21] in mind, the spectral envelope definition [eqn 19] is rewritten as eqns 22 & 23:

$$\lambda(\omega) = \underset{\boldsymbol{\kappa} \neq 0}{sup} \{\boldsymbol{\kappa}^{*}\boldsymbol{P_z}(\omega)\boldsymbol{\kappa}\} \quad [22]$$

$$\lambda(\omega)\boldsymbol{\kappa} = \boldsymbol{P_z}(\omega)\boldsymbol{\kappa} \quad [23]$$

Looking at this form of the equation [eqn 23], it follows that $\lambda(\omega)$ is the largest eigenvalue of the PSD matrix of the normalised data, and $\boldsymbol{\kappa}$ is the corresponding eigenvector. $\boldsymbol{\kappa}$ is called the optimal scaling vector, and its entries represent the contribution of each time series (i.e. variable) to the spectral envelope at that frequency.

Estimation of the spectral envelope

The crux of the spectral envelope method is to estimate the PSD of the normalised data and perform eigenvalue decomposition. The PSD matrix $\boldsymbol{P_z}(\omega_k)$ can be approximated as the periodogram $\hat{\boldsymbol{I}}(\omega_k)$ [eqn 24], which is solved using the discrete Fourier transform.

$$\hat{\boldsymbol{I}}(\omega_k) = \frac{1}{N}\sum_{j=0}^{N-1}\left\{\left(\boldsymbol{z}(t_j)e^{-2\pi i t_j \omega_k}\right) \times \left(\boldsymbol{z}(t_j)e^{-2\pi i t_j \omega_k}\right)^{*}\right\} \quad [24]$$

ω_k are the Fourier frequencies, which are normalised frequencies with units of $\left[\frac{cycles}{unit}\right]$; and are determined as $\omega_k = \frac{k}{N}$, for $k = 1, 2, \ldots, \left[\frac{N}{2}\right]$, where N is the number of observations available from the time series and $\left[\frac{N}{2}\right]$ is the greatest integer less than or equal to $\frac{N}{2}$. Again, the symbol * indicates the conjugate transpose, since multiplication by the transpose gives the magnitude of how strongly the oscillation with frequency ω_k is represented in the data.

The periodogram is not entirely consistent with the PSD matrix, so it is typically smoothed to improve the approximation. This can be done by taking a symmetric moving average of the periodogram [eqn 25]:

$$\boldsymbol{P_x}(\omega_k) = \sum_{j=-r}^{r} h_j \hat{\boldsymbol{I}}(\omega_{k+j}) \quad [25]$$

Where h_j are symmetric positive weights satisfying $h_j = h_{-j}$ and $\sum_{j=-r}^{r} h_j = 1$. h_j can be chosen as $h_j = \frac{r-|j|+1}{(r+1)^2}$ for $j = -r, \ldots, 0, \ldots, r$, where r dictates the degree of smoothness.

Take note that the smoothed periodogram (i.e. PSD matrix) is solved for at each Fourier frequency, meaning that a range of PSD matrices are produced. Eigenvalue decomposition is performed on each PSD matrix, and the spectral envelope therefore identified for each frequency.

Interpretation of the spectral envelope

The spectral envelope is plotted against frequency to produce the spectral envelope plot. Remember that the spectral envelope represents the largest portion of power (i.e. variance) obtainable at each frequency for any linear combination of the original time series. If the original time series has a common frequency component (i.e. if there is a common frequency that multiple variables are oscillating at), the spectral envelope (i.e. largest portion of power/variance) at that frequency will be larger than at other frequencies, and show up as a peak in the plot. A common oscillation frequency ω_c can therefore be detected as the frequency corresponding to a peak in the spectral envelope plot.

The variables that are oscillating at the common oscillation frequency are then identified using the Chi-square (X^2) statistical test, where the hypothesis $\kappa_j(\omega_c) = 0$ is tested for each entry of the optimal scaling vector. Only variables corresponding to entries where $\kappa_j(\omega_c) \neq 0$ are deemed to show significant power at the common oscillation frequency, and are therefore included in the causality analysis.

Details of the statistical test are outside the scope if this work, but the computational aspects are described here. Let $\lambda_1 = \lambda, \lambda_2, \dots, \lambda_n$ be the eigenvalues of the PSD matrix arranged in descending order, and $\kappa_1 = \kappa, \kappa_2, \dots, \kappa_n$ the corresponding eigenvectors. The asymptotic covariance matrix of the optimal scaling vector $\boldsymbol{\kappa}(\omega_c)$ can then be calculated as [eqn 26]:

$$V_\kappa(\omega_c) = v^{-2}\lambda_1(\omega_c)\sum_{L=2}^{n}\lambda_L(\omega_c)[\lambda_1(\omega_c) - \lambda_L(\omega_c)]^{-2}\boldsymbol{\kappa}_L(\omega_c)\boldsymbol{\kappa}_L^*(\omega_c) \qquad [26]$$

Where v is defined as [eqn 27]:

$$v = \left(\sum_{j=-r}^{r} h_j^2\right)^{-\frac{1}{2}} \qquad [27]$$

The distribution of $\frac{2|\hat{\kappa}_j(\omega_c) - \kappa_j(\omega_c)|^2}{\sigma_j(\omega_c)}$ is approximately a Chi-square distribution with two degrees of freedom; where $\hat{\kappa}_j(\omega_c), j = 1, 2, \dots, n$ is the jth entry of the estimated optimal scaling vector, $\kappa_j(\omega_c)$ is the jth entry of the optimal scaling vector, and $\sigma_j(\omega_c)$ is the jth diagonal element of the asymptotic covariance matrix (i.e. variance) of $\boldsymbol{\kappa}(\omega_c)$ calculated with eqn 26. If the test statistic, $\frac{2|\hat{\kappa}_j(\omega_c)|^2}{\sigma_j(\omega_c)} > X_2^2(\alpha)$, then the hypothesis $\kappa_j(\omega_c) = 0$ is rejected with $(1 - \alpha)$ confidence.

The spectral envelope method can therefore be used to detect a common oscillation frequency and variables oscillating at that frequency with a relatively simple implementation. Furthermore, it has a small computational burden; robustness to parameters and data selection changes; can be applied to both linear and nonlinear systems; and has straight-forward interpretation via a single plot where all the common oscillation frequencies can be easily identified (Duan et al., 2014).

3.3.2. Dimensionality reduction using feature extraction

Dimensionality reduction via the spectral envelope method (discussed in Section 3.3.1) can be useful, but this approach falls short in the case where many variables oscillate significantly at a common frequency. In such cases, dimensionality reduction using feature extraction may prove more useful.

Lindner and Auret, (2015) grouped variables displayed in a causality map into blocks of strongly connected components identified in the causality map; where a strongly connected component is defined as a group of nodes that are reachable from each other (i.e. have direct or indirect connections between each other) without violating edge direction (see Figure 6 for an example). They then performed feature extraction in the form of principal component analysis (PCA) on the variables in each block separately, and monitored statistics from these results for fault detection. Once a fault was detected, the block showing the largest effect of the fault was identified and a causality map produced for the variables in that block, which was then used to trace the propagation path back to the root cause. The feature extraction therefore allowed a group of

variables to be identified as showing the largest effect of a fault, so allowing a useful causality map with fewer nodes to be produced.

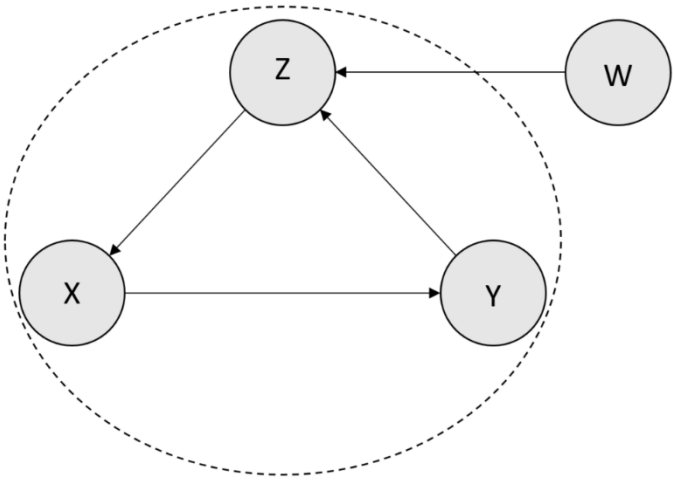


Figure 6: Example of a strongly connected component defined according to Tarjan's algorithm, adapted from Lindner, (2014). X, Y, and Z form a strongly connected component, and W is excluded because it cannot be reached without violating edge direction.

Dimensionality reduction followed by causality analysis has also been used to investigate the causal relationship between different human brain regions in the neuroscience field. Due to the limited literature within the process engineering field that touches on dimensionality reduction in the context of causality analysis, this section is expanded to also include literature from this field.

The data used for this causality analysis is blood-oxygen-level dependent (BOLD) functional magnetic resonance imaging (fMRI) signal time courses (i.e. time series produced by extracting information from successive images of the brain). These images are three-dimensional, so they consist of voxels, which are the 3D versions of pixels that make up 2D images, and each have their own energy and information features. Since each brain region in an fMRI image consists of a multitude of voxels, it is preferable to equate each brain region to one or a few nodes in the causality map, instead of equating each voxel to a node.

Goebel *et al.*, (2003) performed GC on the average values over the voxels in each brain region. Zhou *et al.*, (2009) opted to rather use principal component analysis (PCA) instead of average values, which involves identifying the directions of most variance (i.e. principal components) in the data. PCA was performed on the voxels from each brain region, and then GC was applied to the scores (i.e. the data projected onto the directions identified by PCA) obtained from a few principal components (PCs) from each brain region. It was found that this approach preserves more energy and information features in the signal when compared to the averaging approach, even when only one PC was used for each brain region. However, it was found that GC analysis based on only one PC is not enough for the brain connectivity study, and the number of PCs to include should be chosen such that 85 % of the variance in the data is retained. This heuristic was determined specifically for the investigation of brain connectivity, and should therefore not be directly translated if a version of this approach is applied to causality analysis for fault identification in the process engineering field.

Barrett, Barnett and Seth, (2010) mention that using either average values or scores from PCA are standard approaches for dimensionality reduction during an investigation of the causal

connectivity between brain regions. However, they propose that it may be more appropriate to use *multivariate Granger causality*, where causal interactions between groups of voxels compromising each brain region are considered. In this case, interactions between *sets* of variables x and y and (in the conditional multivariate case) z, are evaluated. The G-causality index (see Section 2.4) is then calculated using the determinant of the covariance matrix of the residuals instead of simply $var(\epsilon_t)$, as suggested by Geweke, (1984).

3.4. Modular structures in networks

The idea of a causality map can easily be extended to that of a network, as both represent a system of interconnected objects. One can then take advantage of this idea, by borrowing established techniques that were developed for application to networks – specifically, techniques for modularisation optimisation to identify modular structures in a network. This section therefore provides an explanation of the modular structure of a network and how this modular structure can be identified in a network (or causality map).

Many networks divide naturally into *modules*, with high intra-connectivity within modules and low inter-connectivity between modules. This is known as a *modular structure*, and a large volume of research has been conducted to develop algorithms to identify this structure in networks. Since causality maps are directed graphs with weighted edges (see Section 2.2), this section provides an overview of a modularisation optimisation algorithm developed by Blondel *et al.*, (2004) and adapted for weighted and directed networks by Newman, (2004) and Leicht and Newman, (2008) respectively.

The aim of any modularisation optimisation method is to maximise *modularity* Q [eqns 28 & 29], defined by Newman and Girvan, (2004):

$$Q = \{fraction\ of\ edges\ within\ communities\} - \{expected\ fraction\ of\ such\ edges\} \quad [28]$$

$$Q = \frac{1}{2m}\sum_i\sum_j\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta_{mod_i mod_j} \quad [29]$$

Where the expected fraction of edges is determined from a random graph conditioned on the degree sequence of the original graph.

In the original method for unweighted and undirected networks, A_{ij} is an entry of the adjacency matrix (i.e. 1 or 0; see Section 2.2); k_i and k_j are the degree (see Section 2.2) of nodes i and j respectively; and m is the total number of edges in the network. To use edge weights (see Section 2.2) instead of adjacency matrix entries for A_{ij}, k_i and k_j become the sums of the weights of edges adjacent to i and j respectively. To incorporate edge direction, eqn 29 is adapted to become eqn 30:

$$Q = \frac{1}{m}\sum_i\sum_j\left[A_{ij} - \frac{k_i^{in} k_j^{out}}{m}\right]\delta_{mod_i mod_j} \quad [30]$$

Where k_i^{in} and k_j^{out} are the indegree and outdegree (see Section 2.2) of nodes i and j respectively. Note that the division is now only by m (instead of $2m$), because distinguishing between indegree and outdegree means that edges are no longer counted twice (i.e. once for each node they are adjacent to).

In both eqn 29 and eqn 30, mod_i and mod_j are the respective labels of the modules that the nodes i and j are assigned to, and δ is the Kronecker delta symbol, which is defined as:

- $\delta_{mod_i mod_j} = 1$, if the two nodes are in the same module (i.e. $mod_i = mod_j$); and
- $\delta_{mod_i mod_j} = 0$, if the two nodes are in different modules (i.e. $mod_i \neq mod_j$).

The optimal modular structure is identified by maximising modularity over all possible divisions, where both the number of modules and module sizes can be varied. To avoid the computationally hard exhaustive optimisation of modularity, Leicht and Newman, (2008) derived an appropriate version of the spectral optimisation method that was established by Newman, (2006) for undirected networks.

First consider the simplified problem of dividing a directed network into two modules, by assigning each node to either "Module 1" or "Module 2". Here, the Kronecker symbol in eqn 30 is defined as [eqn 31]:

$$\delta_{mod_i mod_j} = \frac{1}{2}(s_i s_j + 1) \quad [31]$$

Where s_i is defined as:

- $s_i = +1$, if node i is assigned to "Module 1"; and
- $s_i = -1$, if node i is assigned to "Module 2".

The equation for modularity [eqn 30] is then rewritten as [eqn 31]:

$$Q = \frac{1}{2m}\sum_i \sum_j \left[A_{ij} - \frac{k_i^{in} k_j^{out}}{m}\right](s_i s_j + 1) \quad [31]$$

This can be further simplified [eqn 32] and written in matrix form [eqn 33]. For a step-by-step derivation from eqn 31 to eqn 32, refer to Appendix A.

$$Q = \frac{1}{2m}\sum_i \sum_j s_i \beta_{ij} s_j \quad [32]$$

$$Q = \frac{1}{2m}\boldsymbol{s}^T \boldsymbol{B}\boldsymbol{s} \quad [33]$$

$\boldsymbol{s}$ is a column vector with entries s_i for all the nodes in the network; and $\boldsymbol{B}$ is the modularity matrix, with entries β_{ij} defined as [eqn 34]:

$$\beta_{ij} = A_{ij} - \frac{k_i^{in} k_j^{out}}{m} \quad [34]$$

The goal is therefore to find $\boldsymbol{s}$ that maximises Q for a given $\boldsymbol{B}$. To do this, the spectral optimisation method for undirected networks relies on the symmetry of the modularity matrix $\boldsymbol{B}$, which is not the case for directed networks. This is fixed by adding eqn 32 to its own transpose to give [eqn 35]:

$$\begin{aligned} Q &= \frac{1}{2}(Q + Q^T) \\ &= \frac{1}{2}\left(\frac{1}{2m}\boldsymbol{s}^T \boldsymbol{B}\boldsymbol{s} + \frac{1}{2m}(\boldsymbol{s}^T \boldsymbol{B}\boldsymbol{s})^{\boldsymbol{T}}\right) \\ &= \frac{1}{4m}\boldsymbol{s}^T(\boldsymbol{B} + \boldsymbol{B}^T)\boldsymbol{s} \end{aligned} \quad [35]$$

Where the matrix $\boldsymbol{B} + \boldsymbol{B}^T$ is symmetric.

From here, Q^* is defined as [eqn 36]:

$$Q^* = \boldsymbol{s}^T(\boldsymbol{B} + \boldsymbol{B}^T)\boldsymbol{s} \quad [36]$$

Where $\boldsymbol{s} = \underset{s}{\text{argmax}}\, Q = \underset{s}{\text{argmax}}\, Q^*$ (i.e. $\boldsymbol{s}$ is the same for both Q_{max} and Q^*_{max}). For simplicity, Q^* is used for the optimisation procedure.

The next step is to write $\boldsymbol{s}$ as a linear combination of the eigenvectors $\boldsymbol{v}_i$ of $\boldsymbol{B} + \boldsymbol{B}^T$ [eqn 37] (which are orthogonal due to the symmetry of $\boldsymbol{B} + \boldsymbol{B}^T$):

$$\boldsymbol{s} = \sum_i a_i \boldsymbol{v}_i \quad [37]$$

Which allows eqn 36 to be rewritten in terms of these eigenvectors [eqn 37], and further simplified [eqn 38]:

$$Q^* = \sum_i a_i \boldsymbol{v}_i^T(\boldsymbol{B} + \boldsymbol{B}^T)\sum_j a_j \boldsymbol{v}_j^T \quad [37]$$

$$Q^* = \sum_i \gamma_i (\boldsymbol{v}_i^T \boldsymbol{s})^2 \quad [38]$$

Where γ_i is the eigenvalue of $\boldsymbol{B} + \boldsymbol{B}^T$ that corresponds to the eigenvector $\boldsymbol{v}_i$. For a step-by-step derivation from eqn 37 to eqn 38, refer to Appendix A.

Assuming that the eigenvalues are labelled in decreasing order (i.e. $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n$), Q^*_{max} is achieved when $\boldsymbol{s}$ is parallel to the leading eigenvector $\boldsymbol{v}_1$. This is because the angle between two parallel lines is zero, allowing the dot product to achieve its maximum value: $\boldsymbol{v_1^T s} = \boldsymbol{v}_1 \cdot \boldsymbol{s} = |v_1||s|cos\theta = |v_1||s|\cos(0) = |v_1||s|$. However, the constraint $s_i = \pm 1$ means that the entries s_i cannot be chosen as any values to ensure that $\boldsymbol{s}$ lies parallel to $\boldsymbol{v}_1$. The values of s_i are therefore chosen to make $\boldsymbol{s}$ as close as possible to parallel to $\boldsymbol{v}_1$. Since there are only two available options for s_i, their values are chosen (and modules assigned accordingly) to achieve the same sign as the entries of $\boldsymbol{v}_1$, meaning that:

- $s_i = +1$, if $(v_1)_i > 0$;
- $s_i = -1$, if $(v_1)_i < 0$; and
- $s_i = \pm 1$, if $(v_1)_i = 0$;

Where $(v_1)_i$ refers to the i^{th} entry of v_1.

This bisection (i.e. splitting a module into two smaller modules) is then applied repeatedly until further division no longer results in an increase in total modularity of the network ΔQ. This change in total modularity is evaluated for each bisection as [eqn [39]:

$$\begin{aligned}\Delta Q &= \frac{1}{2m}\left\{\sum_{i\in g}\sum_{j\in g}(\beta_{ij} + \beta_{ji})\frac{s_i s_j + 1}{2} - \sum_i \sum_{j\in g}(\beta_{ij} + \beta_{ji})\right\} \\ &= \frac{1}{4m}\left\{\sum_{i\in g}\sum_{j\in g}(\beta_{ij} + \beta_{ji}) - \delta_{ij}\sum_{k\in g}(\beta_{ik} + \beta_{ki})\right\} s_i s_j \\ &= \frac{1}{4m}\boldsymbol{s}^T\left(\boldsymbol{B}^{(g)} + \boldsymbol{B}^{(g)T}\right)\boldsymbol{s}\end{aligned} \quad [39]$$

Where g is the module within the original network that is currently being divided, and $\boldsymbol{B}^{(g)}$ is a matrix with entries defined as [eqn 40]:

$$\beta_{ij}^{(g)} = \beta_{ij} - \delta_{ij} \sum_{k \in g} \beta_{ik} \qquad [40]$$

This spectral optimisation method provides a good guide to the optimal modular structure, but it is not always accurate when it comes to the placement of individual nodes. A fine-tuning step is therefore used between bisections, where nodes are moved back and forth between modules until modularity cannot be increased further. The complete algorithm to determine the modular structure of a weighted direct network is therefore:

1. Construct the modularity matrix $\boldsymbol{B}$, with entries $\beta_{ij} = A_{ij} - \frac{k_i^{in} k_j^{out}}{m}$.
2. Identify the largest positive eigenvalue and corresponding eigenvector of $\boldsymbol{B} + \boldsymbol{B}^T$.
3. Assign each node to a module (i.e. $s_i = +1$ or $s_i = -1$) based on the sign of the corresponding entry of the eigenvector identified in Step 2.
4. Further bisect each module, using $\boldsymbol{B}^{(g)}$, with entries $\beta_{ij}^{(g)} = \beta_{ij} - \delta_{ij} \sum_{k \in g} \beta_{ik}$.
5. Fine-tune by moving each node back and forth until modularity cannot be increased further.
6. Repeat Steps 4 and 5 until further division does not give a positive value for ΔQ.

3.5. Summary of causality map interpretability and network visualisation in literature

This section provides a broad overview of how causality map interpretability and network visualisation have been addressed in published literature. To the best of the author's knowledge, the sources listed are representative of the range of techniques and tools that have been used to improve causality map interpretability and network visualisation up to date. Table 3 presents a summary of literature where causality analysis was performed for fault identification in process engineering. The summary includes the process or case study that the causality analysis was applied to, the resolution of the causality analysis (i.e. whether a single process unit was considered or larger section of a plant), as well as the layout used for the resulting causality map, and any techniques and tools used to improve the causality map interpretability.

The studies are listed in consecutive order to identify any improvements or additions throughout the years to aid in causality map interpretability. However, it is clear that within the context of fault identification in process engineering, only Lindner, (2019) and Suresh, Sivaram and Venkatasubramanian, (2019) have conducted any research where the explicit subject was causality map interpretability. Furthermore, within other studies where causality analysis is applied, little or no reference is made to the causality map interpretability. For example, different layout styles are used, but no motivation given for the choices. This is because causality analysis research has been largely focussed on the development and improvement of the causality analysis techniques themselves, and these techniques have been tested on case studies with few variables. The largest resolution of case studies that have been used are single plant sections. Application of causality analysis to intricate and interconnected systems (e.g. multiple plant sections) involving numerous variables, with specific reference to interpretability, is therefore identified as a vacancy

in current research. This application would be especially useful to allow a fault detected in a specific plant section to be traced back to its root cause even if originated in an upstream section, which is often the case due to the smearing effect (see Section 2.6). In addition, to the best of the author's knowledge, no attempt has been made to define characteristics of the ideal causality map – hence, this is also identified as a gap in literature.

The lack of research regarding causality map interpretability within the field of process engineering can be addressed by following on the idea that causality maps are networks (mentioned in Section 3.4). This literature survey is therefore extended to network visualisation in fields outside process engineering. The summary presented in Table 4 provides the case study that the network visualisation was applied to, the aim of the network visualisation in that case, as well as the network layout and any techniques and tools used to improve the network interpretability.

Graph theoretic characteristics and parameters such as *connector hubs* and *clustering coefficient* are mentioned in some of the descriptions in Table 4, and are therefore defined here. The nodes that are connected to nodes in other modules (defined in Section 3.4) and have a high degree (see Section 2.2) are known as *connector hubs* (Newman, 2003). The *clustering coefficient* is a quantification of the number of edges that exist between the nearest neighbours of a node as a proportion of the maximum number of possible edges (Reijneveld *et al.*, 2007).

Table 3: Summary of how causality map interpretability has been addressed in literature, where causality analysis was performed for fault identification in process engineering. *Blue italic text indicates that the technique/tool is mentioned for the first time.*

Source	Process / Case study	Resolution of causality analysis	Causality map layout	Techniques and tools used to improve causality map interpretability
Bauer *et al.*, (2007)	Chemical process	Fault identification within a single process unit.	Variables arranged in a straight line	None mentioned.
Bauer and Thornhill, (2008)	Chemical process	Fault identification within a single process unit.	Two layouts presented: (1) Variables arranged in a straight line; (2) Layered layout	None mentioned.
Thambirajah *et al.*, (2009)	Chemical process	Fault identification within a single process unit.	No specific layout; appears as if nodes were simply placed in a way that prevents edges from crossing over other edges/nodes.	*Spurious connections decreased by using process knowledge to validate causality analysis results.*
Shu and Zhao, (2013)	Two case studies: (1) Water mixing process; (2) Tennessee Eastman process	Fault identification within a plant section consisting of two/three units.	For the respective case studies: (1) Layered layout; (2) Variables in a row.	*Number of edges decreased by eliminating connections that are combinations of other connections.*
Duan, (2014)	Tennessee Eastman process	Fault identification within a plant section consisting of five units.	Layered layout	*Unidirectional and bidirectional causality edges coloured differently. Source nodes with zero indegree placed on left of causality map. Arrow head sizes proportional to causal strength of connection.*
Kühnert and Beyerer, (2014)	Two case studies: (1) Simulated CSTR; (2) Experimental laboratory-scale level control system	Fault identification within a single process unit.	(1) Layered layout; (2) No clear layout	
Lindner and Auret, (2014)	Two-tank simulation with temperature and level control	Fault identification within a single process unit.	Layered layout	Nodes colour-coded to indicate symptom nodes and root nodes. Edges coloured to indicate paths from root to symptom nodes. *Note: This was done after causality map interpretation had yielded the results.*
Landman *et al.*, (2014)	Drying section of an industrial paperboard machine	Plant section consisting of 74 drying cylinders divided into six steam groups.	Layered layout	Number of connections decreased by using process knowledge in the form of a connectivity matrix to remove all indirect connections.
Yuan and Qin, (2014)	Tennessee Eastman process	Plant section consisting of five process units.	Circle layout	Unidirectional and bidirectional causality edges coloured differently.
Duan *et al.*, (2015)	Tennessee Eastman process	Plant section consisting of five process units.	Circle layout	Unidirectional and bidirectional causality edges coloured differently.

Source	Process / Case study	Resolution of causality analysis	Causality map layout	Techniques and tools used to improve causality map interpretability
Lindner and Auret, (2015)	Concentrator process	Plant section consisting of nine process units.	Layered layout	*Number of nodes and edges decreased by blocking according to strongly connected components (i.e. a single nodes represents multiple strongly connected variables). See Section 3.3.2.*
Landman and Jämsä-Jounela, (2016)	Drying section of an industrial paperboard machine	Plant section consisting of 74 drying cylinders divided into six steam groups.	Layered layout	*Bolded and dashed edges represent connections above and below causal strength thresholds respectively. Edges coloured in red if suspected to be indirect/spurious.*
Lindner *et al.*, (2018)	Two parallel flotation banks	Plant section consisting of 14 process units.	Layered layout	*Separate causality maps constructed according to variable location and variable category.*
Wakefield *et al.*, (2018)	Milling circuit simulation	Plant section consisting of four process units.	Layered layout	None mentioned.
Lindner, (2019)	Various mineral process case studies	Plant sections with the number of units in the order of ten.	See *Layout sub*section in Section 2.7.1.	See Section 3.2.
Suresh, Sivaram and Venkatasubramanian, (2019)	Tennessee Eastman process	Plant section consisting of five process units.	Layered layout	*Produce an acyclic causality map from a cyclic causality map. See Section 3.1.*

Table 4: Summary of how network visualisation has been addressed in literature in fields outside process engineering. *Blue italic text indicates that the technique/tool is mentioned for the first time.*

Source	Data / Case study	Aim of network visualisation	Network layout	Techniques and tools used to improve network interpretability
Newman, (2003)	Friendship network of children in US school. *Directed, because person A may say person B is their friend but not vice verca.*	Investigation of whether a community structure (i.e. modular structure) exists in the data.	Force layout, but resulted in what appears to be a circular layout	*Nodes colour-coded according to additional information in dense network.*
Newman, (2004)	Network of co-occurrence of words in Reuters newswire stories	Testing of module optimisation algorithm where aim is to identify communities (i.e. modules) in the network.	Modular according to communities (i.e. modules) identified in data (i.e. force layout)	*Clusters of nodes separated into modules with nodes colour-coded according to communities (i.e. modules). Edge widths represent weightings*

Source	Data / Case study	Aim of network visualisation	Network layout	Techniques and tools used to improve network interpretability
Palla *et al.*, (2005)	Two case studies: (1) Network of word associations related to cognitive sciences; (2) Molecular-biological network of protein-protein interactions	Testing of module optimisation algorithm where aim is to identify communities (i.e. modules) in the network AND incorporate the fact that some modules overlap.	Modular according to communities (i.e. modules) identified in data (i.e. force layout)	Colour-coded nodes and edges according to communities (i.e. modules). *Overlapping nodes and edges between communities emphasised in red. Edge and node weightings proportional to total number of communities they belong to. Number of edges decreased by setting a threshold connection strength.*
Blondel *et al.*, (2008)	Communication extracted from Belgian mobile phone network.	Testing of module optimisation algorithm where aim is to identify communities (i.e. modules) in the network.	Modular layout (i.e. force layout)	*Nodes represent entire communities and node size is proportional to the number of individuals in the corresponding community.* Nodes colour-coded according to additional information in dense network.
E. A. Leicht and Newman, (2008)	Randomly generated network	Testing of module optimisation algorithm where aim is to identify communities (i.e. modules) in the network.	Modular according to communities (i.e. modules) identified in data (i.e. force layout)	*Shaded backgrounds for different modules* enhance their separation.
Bullmore and Sporns, (2009)	(1) Cellular function network; (2) brain structure network	Comparison of cellular function and brain structure networks in healthy people and people with Schizophrenia.	Two layouts presented: (1) Circle layout; (2) appears to be grid layout mimicking brain structure	In the respective layouts: (1) *Dotted lines drawn around each modules in modular structure.* Nodes and edges colour-coding according to module. (2) *Nodes arranged vertically by degree and separated horizontally for clarity. Colour coding according to brain region (equivalent to plant section in process engineering). Nodes sized according to clustering coefficient.*
Pollonini *et al.*, (2010)	Brain structure connectivity	Comparison of brain structure connectivity in people with autism and control groups.	Grid layout mimicking brain structure	None mentioned.
Zhou *et al.*, (2014)	Neuron network	Search for mapping between causal connectivity and synaptic connectivity in neural networks.	Circle layout	None mentioned.

CHAPTER 4

Causality analysis methodology and proposed approaches

4.1. Workflow for causality analysis

Causality analysis in this study followed the workflow: (1) variable selection; (2) data selection; (3) the causality analysis calculation; (4) incorporation of process knowledge and tools to aid in causality map interpretation; (5) causality map construction, and (6) causality map interpretation. Steps (1) to (5) were implemented using MATLAB, (2018).

Variable selection was performed using three different methods throughout this project. In the first method, the time series of the variables were observed to identify oscillating variables (see Chapter 5). In the second method, optional variable selection was implemented via the variables-slider based on the spectral envelope method, as explained in Section 4.3.2. In the third method, variable selection was performed once-off using the spectral envelope method described in Section 3.3.1.

The various tools, visualisation techniques and approaches in this study were applied to a simulated case study of tanks in a network (described in Chapter 5). Data availability was therefore not a limiting factor, as the simulation could simply be run for as long as necessary. Due to a lack of guidance on how to select SP and TW in literature (see Section 2.7.5), a novel heuristic approach was developed in Section 6.2. Using this approach, the combination of SP and TW that results in the largest connection strength of a known connection was selected.

Pairwise conditional GC was performed to obtain the G-causality indices, using equations 7-9 in Section 2.4.1; and the G-causality indices were stored in a causality matrix (see Section 2.2). The model order was obtained using the finite sample bias-corrected form of AIC in eqn 5, and statistical significance was determined with the F-statistical test in eqn 10 using a significance level of 0.01 unless stated otherwise. GC code was developed and validated by the author to ensure understanding, but the *GCCA mode* in the MVGC toolbox (Barnett and Seth, 2014) was used to generate data-based causality results for this study. The author's code was validated by applying it to the case study described in Chapter 5 and comparing the results of the G-causality indices to those obtained when using the MVGC toolbox (Barnett and Seth, 2014) (see Appendix B).

Process knowledge was incorporated to constrain connections and potential root cause variables in causality maps, as described in Section 4.2. For the base case, the transitive reduction of a causality map (see Section 3.2.1) was obtained using the *transreduction()* function in MATLAB, (2018). The visual display of node rankings, as well as the connections-slider and variables-slider, were implemented as explained in Section 4.3.

Causality maps were constructed as digraphs from the causality matrices containing the G-causality indices. The *circle* layout (see Section 3.2.2) was used, as the process in the case study is cyclic (see Chapter 5). Causality maps were then used to track the propagation path of a fault back to its root cause, as discussed in Section 2.6.

4.2. Procedures for incorporating process knowledge in causality maps

Process knowledge was incorporated in causality maps in two different ways: to (1) constrain connections in the causality map, and (2) constrain the group of potential root cause variables.

4.2.1. Constraining connections

Process knowledge in the form of a connectivity matrix was used to validate connections in a causality map. Data-based causality analysis was performed on all variable pairs, followed by a depth-first search (see Section 3.2.1) through the connectivity matrix to corroborate the physical possibility of each identified connection (i.e. to validate each connection) and so refine the causality map. Code was written in MATLAB (MATLAB, 2018) to achieve this for both a standard causality matrix and a causality matrix where nodes represent groups of variables, and it is available in Appendix B.

Two variations of the refined causality map were constructed. In the first variation, only connections identified by the data-based causality analysis and validated by the connectivity matrix were displayed on the causality map. In the second variation, solid lines were used to represent connections validated by the connectivity matrix, and dashed lines were used to represent connections that were not validated by the connectivity matrix.

The connectivity matrix was manually constructed from process knowledge to mimic results that XML scraping from a P&ID of the case study (see Chapter 5) would yield. In addition, known causal connections between MVs and CVs in feedback control loops as well as known causal connections between tank levels and underflows were included in the connectivity matrix, to ensure that everything that is known about the process was reflected in the connectivity matrix. The connectivity matrix is available in Appendix F.

4.2.2. Constraining potential root cause variables

Process knowledge about the smearing effect was used to constrain the group of potential root cause variables. The root cause variable should be able to reach all variables that show an effect of the fault; and a causality map should consist only of variables that show an effect of the fault. As such, a reachability matrix was constructed from the adjacency matrix (see Section 2.2), and used to identify nodes that cannot reach all other nodes in the causality map, excluding any nodes with a degree of zero. The identified nodes were displayed in a faded colour on the causality map to indicate that they should be removed from immediate consideration as the root cause.

4.3. Procedures for applying tools to aid in causality map interpretation

Two different types of tools were applied to the causality maps to aid in their interpretation, namely the visual display of node rankings and sliders. The nodes were ranked using the node importance techniques in Section 3.2.4; and the sliders are a novel concept for causality maps.

4.3.1. Displaying node rankings

The nodes in the causality map were ranked according to their influence on the rest of the map, using the node importance technique based on the page rank algorithm (see Section 3.2.4). The rank of each node was visually displayed on the causality map by colouring the nodes according to their rank and providing a colour bar as key, where the colour at the top of the colour bar represents the largest influence on the map and the colour at the bottom of the colour bar represents the smallest influence on the causality map.

4.3.2. Sliders for connections and variables

Two sliders were implemented and visually displayed along with the causality maps: a connections-slider that determines which connections are displayed on the causality map; and a variables-slider that determines which variables are included in the causality analysis and subsequently displayed on the causality map. Both sliders can be manipulated until the 'most ideal' map is obtained for the system under consideration, but also have default values so that the user does not necessarily need to interact with them.

The connections slider was incorporated as the significance level of the F-test applied to the GC results, with a minimum value of 1×10^{-16}, a maximum value of 0.05, and a default value of 0.01. It acts as a threshold for how strong a causal connection must be for it to be displayed on the causality map (i.e. a lower value on the connections-slider means only the stronger connections are displayed – so fewer connections are displayed).

Two versions of a variables-slider were proposed and investigated - both based on the spectral envelope method (see Section 3.3.1). The PSD matrices for the Fourier frequencies were calculated as the smoothed periodogram with r = 1, following Duan *et al.*, (2014). The common oscillation frequency was determined as the frequency corresponding to a peak in the spectral envelop plot. The oscillation contribution of each variable was then found via the optimal scaling vector at the common oscillation frequency (i.e. the eigenvector corresponding to the largest eigenvalue of the PSD matrix).

Both versions of the variables-slider were based specifically on the oscillation contribution of variables at the common oscillation frequency, with the slider-value acting as a threshold that must be met by a variable for it to be included in the causality analysis. The first version compares the oscillation contribution of each variable to the oscillation contribution of an identified oscillating variable (IOV): $|OC - OC_{IOV}| < threshold$, where the default threshold is one, to include all variables. In the first version, the identified oscillating variable is the variable that first alerted operators/engineers to the problem (e.g. an alarm that went off in the plant). In this project, the identified oscillating variable was taken as the known root cause variable (Chapter 5) to ensure that a variable relevant to the fault was selected for the investigation. The second version acts as a threshold for how large the oscillation contribution of a variable must be in order to be included in the causality: $OC > threshold$, where the default threshold is zero, to once again include all the variables.

In both versions of the variables-slider, the optimal scalings were each divided by the maximum optimal scaling to limit the values between 0 – 1. In the first version, a slider-value of 0 includes only the identified oscillating variable itself in the causality analysis, and a value of 1 includes all variables. In the second version, a slider-value of 0 includes all the variables in the causality analysis, and a value of 1 includes only the variable with the largest oscillation contribution. The default values for the first and second versions of the variables-slider were therefore set to

1 and 0 respectively, so that all variables are included, as variable selection is an optional step in causality analysis. The variables-slider were compared to the existing, once-off, implementation of the spectral envelope method, where variables oscillating at the common oscillation frequency were identified with a 99.9 % confidence interval (i.e. $test\ statistic > 13.82$).

4.4. Proposed hierarchical approach for causality analysis

This study proposed a hierarchical approach for causality analysis, where causality maps were constructed in two subsequent stages. In the first stage, a less-detailed, plant-wide map was constructed and used to localise the root cause to a specific group of variables; followed by the second stage, where a causality map was constructed for only that group of variables and used to identify the root cause. The plant-wide map in stage one was obtained by first grouping the variables, finding representatives for each group of variables, and then performing causality analysis on these representatives, as depicted in Figure 7.

4.4.1. Grouping the variables

The variables were grouped in two different ways: using a knowledge-based method, where the variables were grouped according to plant sections (PSs); and using a data-based method, where the variables were grouped according to the causal modular structure in the causality map consisting of all variables in the plant showing an effect of the fault (a *standard causality map*). The causal modular structure was identified using a repetitive bisection method to optimise modularity, taking into account both edge weights and direction, where the tolerance for a small enough change in modularity to stop the bisection was set to 1×10^{-6} (see Section 3.5).

4.4.2. Finding a representative for each group of variables

Representatives for each group of variables were found using feature extraction in a data-based approach. Each group was represented by the first principal component (PC1) obtained from performing principal component analysis (PCA) on all its variables. Hereafter, the following abbreviations (Table 5) are used to refer to each of these approaches:

Table 5: Abbreviations used for the different hierarchical approaches.

Abbreviation	Meaning
PS-PC1	Variables grouped according to PSs; PC1s as representatives
Mod-PC1	Variables grouped according to modules in the data; PC1s as representatives

4.5. Procedure for evaluating the interpretability of causality maps

Causality map interpretability was evaluated based on the characteristics assigned to the ideal causality map (see Table 8 in Section 6.1) by answering the questions in Table 6. Process knowledge of the simulated case study with a known fault (see Chapter 5) was used to answer questions 1, 3, 4, and 9.

In each case, the interpretability of a causality map was evaluated in comparison to that of another causality map used as a base case. In Sections 6.3 and 6.4, the base case was the standard causality map without the tool or method being demonstrated; in Section 6.5 regarding the hierarchical approach, the base case was a transitive reduction of the standard causality map, as that was the main prior method used to improve causality map interpretability.

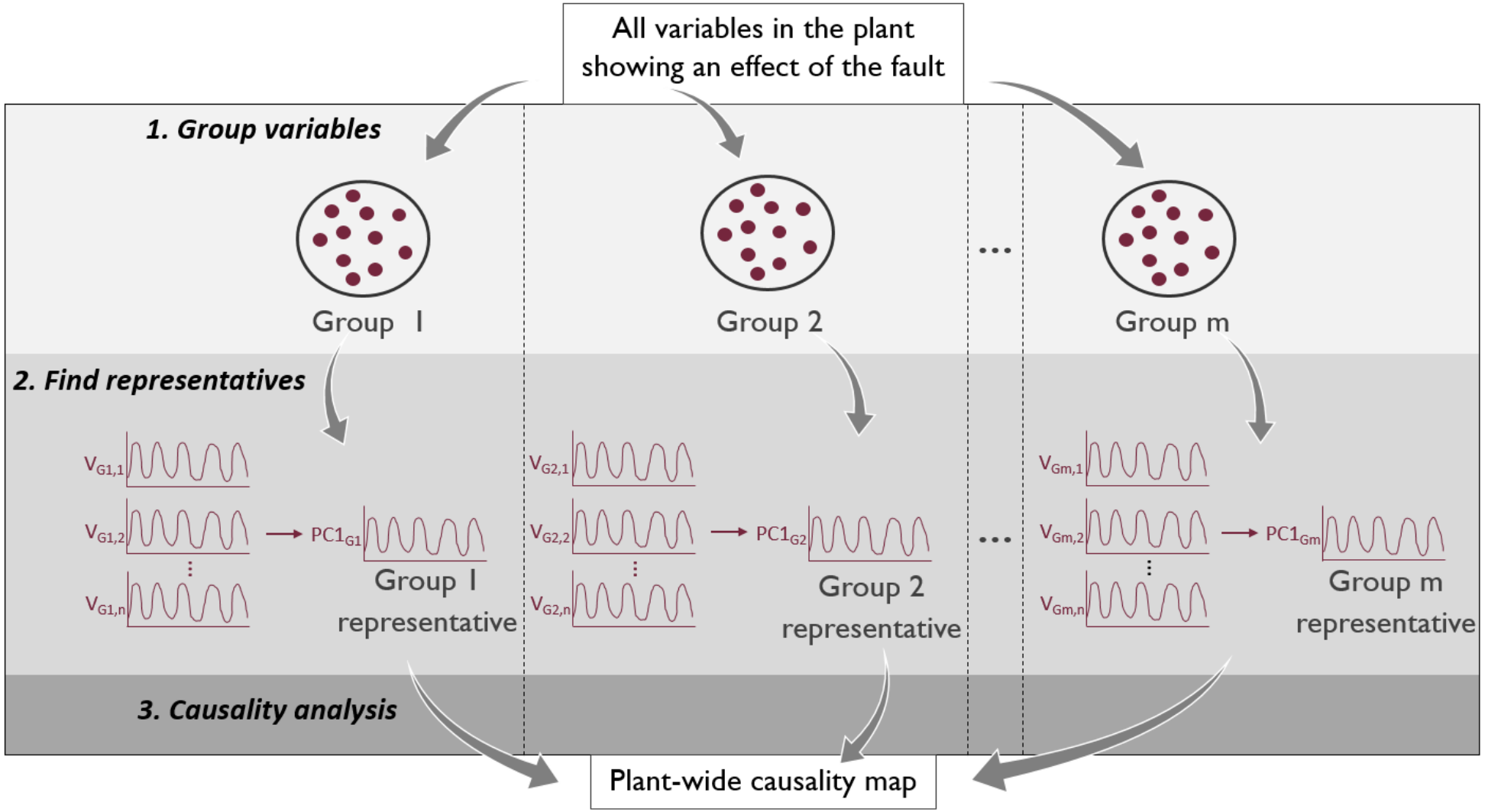


Figure 7: Flow diagram of the procedure to obtain the plant-wide causality map in the first stage of the proposed hierarchical approach. See Section 4.4.1 and Section 4.4.2 for how the variables were grouped and how representatives for each group were found respectively.

Table 6: Questions to evaluate the interpretability of a causality map

1	Is the true root cause variable present in the causality map?
2	What fraction of variables present in the causality map are oscillating at the common oscillation frequency?
3	What fraction of connections present in the causality map are spurious (i.e. what is the likelihood of a connection in the causality map being spurious)?
4	Are all the true propagation paths present in the causality map?
5	What is the graph density of the causality map (see eqn 13 in Section 3.2.5)?
6	How many nodes are there in the causality map?
7	Is the causality map acyclic or cyclic?
8	How many potential root causes are identified by the causality map?
9	Does the causality map identify the true root cause?

4.6. Causality maps usability study

The usability study was performed in the form of an online survey on the SUN Survey server, which was structured as a small test to familiarise the participants with the different approaches, followed by some closed questions and a form of expert interview (as open-ended questions in the survey) based on the participant's experience of the test. The test asked participants to identify the root cause from different causality maps: namely, the base case of transitive reduction; the hierarchical approach where grouping was done according to plant sections (PSs) and the first principal components (PC1s) were used as representatives (referred to as *PS-PC1*); and the hierarchical approach where grouping was done according to modules identified in the standard causality map and PC1s were used as representatives (referred to as *Mod-PC1).* The survey questionnaire and causality maps are available in Appendix H.

Participants were required to have some form of engineering training, and ideally to have some experience with fault identification and causality analysis. In accordance with this, purposeful sampling was applied and three specific pools of participants were invited to participate in the survey: namely final-year or postgraduate Chemical Engineering students at Stellenbosch University (SU); employees of the industrial IoT firm in the mineral processing industry; and employees from Anglo American Platinum (AAP). All relevant institutional permissions were acquired and ethical clearance was obtained; refer to Appendix H for a copy of the notice of ethics approval for this study.

The survey was not psychometrically developed, as that was not the aim of the usability study. It was developed to gain insight into which characteristics or tools are more useful and/or understandable in all causality maps, as well as gain insight into whether the base case or hierarchical approaches can be practically applied in industry. In accordance with this, participants were asked to rank the extent to which characteristic/tools played a role in their decision-making while identifying a root cause from a causality map, where the characteristics/tools were node degree (i.e. the number of arrows pointing towards and away from node(s)), edge type (solid/dashed to indicate validated/non-validated), node colour (i.e. node ranking), and edge weight. In addition, participants were asked to indicate whether they used the connections-slider and/or variables-slider (first version of the variables-slider; see Section 4.3.2), and indicate the user-friendliness of each causality map on a scale from 1 – 4, with 1 representing the least user-friendly and 4 representing the most user-friendly. The response to these questions were aggregated before analysis, although they were linked to background information where differences in trend were observed. Background information

includes the participant's university/company (SU, the industrial IoT firm, or AAP), highest level of education (National senior certificate or equivalent, Bachelor's degree, Master's degree, or PhD), and whether the participant had any prior experience with causality analysis.

The open-ended questions were analysed using thematic analysis, as recommended by Braun and Clarke, (2006), specifically for expert interviews, but also for researchers with limited experience in qualitative research (such as engineers). Coding was therefore performed by organising the data according to words or short phrases (codes) that represent recurrent features/elements in the data that are relevant to interpretability in causality maps; followed by grouping these codes into themes. The data was analysed as a rich description of all themes rising from all three open-ended questions (i.e. regarding the base case, PS-PC1, and Mod-PC1) in a deductive (theoretical) manner (i.e. top-down approach where themes were driven by the aspect of interpretability). In addition, the data was analysed for semantic (i.e. surface level) themes, where the "meaning" (epistemology) in the data was analysed as realistic (i.e. participants were taken to directly reflect their meanings).

CHAPTER 5

Case study: Tank network simulation

The case study used in this project is a dynamic simulation of a tank network developed in MATLAB and Simulink (MATLAB, 2018) specifically for this project. The aim of this simulation is not to generate time series that are accurately representative of a real-world process; rather, this simulation serves as a case study that can be fully understood and allows for in-depth understanding of its causality analysis results. This chapter provides a description of the process and control structure, as well as how data was generated from the simulation in the presence of a fault.

5.1. Description of overall process and control in tank network

The tank network consists of multiple tanks with liquid inventory. Figure 8 presents the P&ID of the system, and Appendix C contains the model development and Simulink implementation, and Appendix F contains the connectivity matrix. The tank network was developed to represent a plant with three plant sections (PSs) that have identical configurations and are separated by buffer tanks (BT 001 and BT 002). A detailed description of PS1 is therefore provided, followed by a description of the interaction between the PSs.

PS1 consists of two tanks in series (TK 001 and TK 002), with an exogenous input fed to the first tank (TK 001). The exogenous input is modelled as a random walk, because disturbances in the real world are not entirely time-independent, and a random walk represents a stochastic input excitation (where deterministic input excitation would cause problems with the AR models in GC). Both tanks have a constant product flow through a throttled centrifugal pump that leads to the next tank in series. The second tank (TK 002) also has a level-dependent underflow, of which 50 % is recycled back to the first tank (TK 001) and the rest is purged.

The buffer tanks (BT 001 and BT 002) are used to attenuate disturbances between PSs, and both have a level-dependent underflow of which 65 % is fed to the next PS. The rest of the underflow from BT 001 is purged, but the rest of the underflow from BT 002 is recycled back to TK 001 in PS1 to increase the interconnectivity between PSs. Dead times of 5 min and 10 min are artificially present in the pipes within and between PSs respectively, to incorporate the phenomena of time delay that is typically present in real-world processes.

Downstream inventory control is implemented on all the tanks within PSs (i.e. excluding buffer tanks). The CVs are the levels and the MVs are the throttling valves on the product lines. Loose linear averaging level control is implemented to maintain the CVs at their set points (SPs) of 0.5 (i.e. half full tanks) with an allowable deviation of 0.1 m/m (i.e. 10 %). Refer to Appendix E for the derivation of the controller tuning relations.

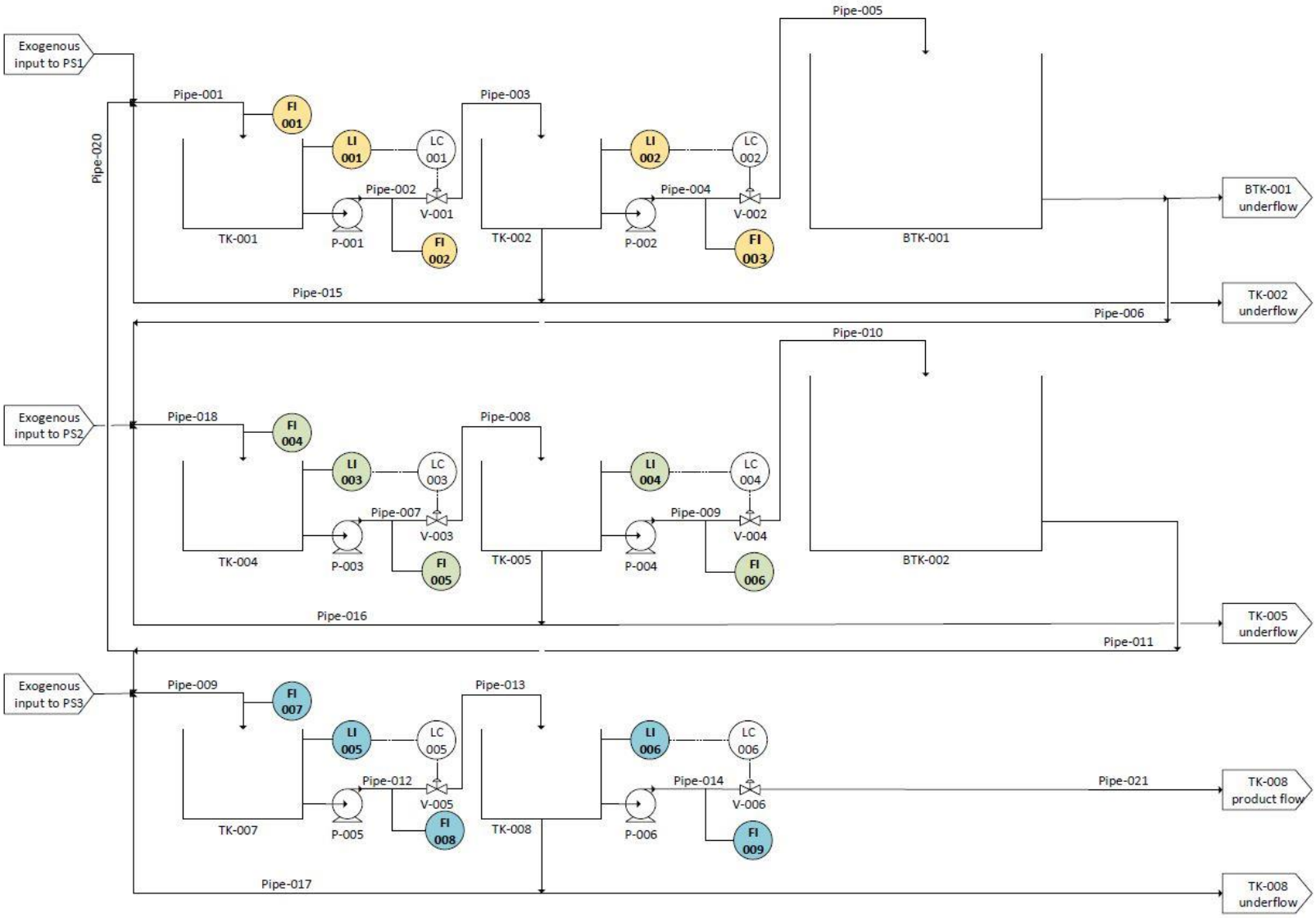


Figure 8: P&ID of tank network consisting of three sections, drawn using MS Visio (Microsoft Corporation, 2015). *Yellow, PS1; Green, PS2; Blue, PS3.*

5.2. Data from tank network simulation

The tank network simulation was used to generate time series data in the presence of a fault for fault identification using causality analysis. This section presents the measured variables that provide the time series data, and an explanation of how valve stiction was inserted as a fault in the simulation.

5.2.1. Measured variables

The process contains 15 measured variables, consisting of total inlet flowrates, product flowrates, and levels. Table 7 summarises the tag numbers, along with descriptions and units of each measured variable. In addition, sensors associated with measured variables are indicated in colour on the P&ID (Figure 8), where the different colours represent different PSs.

Table 7: List of measured variables in simulated tank network case study

Variable no.	Tag number	Description	Units
1	FI 001	Total volumetric inlet flowrate to TK 001	L/min
2	LI 001	Level of TK 001	m/m
3	FI 002	Volumetric product flowrate of TK 001 / Total volumetric inlet flowrate of TK 002	L/min
4	LI 002	Level of TK 002	m/m
5	FI 003	Volumetric product flowrate of TK 002	L/min
6	FI-004	Total volumetric inlet flowrate to TK 004	L/min
7	LI 003	Level of TK 004	m/m
8	FI 005	Volumetric product flowrate of TK 004 / Total volumetric inlet flowrate of TK 005	L/min
9	LI 004	Level of TK 005	m/m
10	FI 006	Volumetric product flowrate of TK 005	L/min
11	FI 007	Total volumetric inlet flowrate to TK 007	L/min
12	LI 005	Level of TK 007	m/m
13	FI 008	Volumetric product flowrate of TK 007 / Total volumetric inlet flowrate of TK 008	L/min
14	LI 006	Level of TK 007	m/m
15	FI 009	Volumetric product flowrate of TK 008	L/min

5.2.2. Fault: Valve stiction in plant section 2

Valve stiction was incorporated as the fault in the tank network simulation. It is a commonly occurring fault known to cause plant-wide oscillations, so it successfully demonstrates the smearing effect (see Section 2.6) in the simulation. For this project, the dynamics of valve stiction were simulated using the empirical model described in Choudhury, Thornhill and Shah, (2005). Refer to Appendix D for a full explanation of how this was modelled in Simulink (MATLAB, 2018).

The valve stiction model parameters were selected to achieve a low frequency oscillation with aggressive overshoot (see Figure 37 in Appendix D), as this maximises the chance of successfully constructing an accurate plant-wide causality map. The low frequency fault allows the slow dynamic response of the entire plant to be captured in the AR models in GC (see Section 2.4) via sub-sampling (see Section 2.7.5), without obscuring the oscillation itself between sampling points. The overshoot causes a large fault amplitude, which results in higher G-causality index values (see Section 2.7.2).

The fault was simulated on the second control valve in PS2 (V-004 in Figure 8), so that the oscillation needs to propagate both downstream and upstream to reach all variables in the plant. This incorporates the recycle stream from PS2 to PS1 in the fault propagation path, which represents the interconnectivity and often cyclic interaction present in real chemical and mineral processing plants.

The time series for all 15 measured variables (see Section 5.1.1) are presented in Figure 9, divided into their separate plant sections by colours corresponding to those in the P&ID (Figure 8). The valve stiction begins at 200 min, as indicated by the dashed line in Figure 9. The first effect of the fault is seen as a small dip in the flowrate being adjusted by the sticky valve (FI 006), which almost immediately affects its CV (LI 004), after which the oscillation propagates to all the other measured variables.

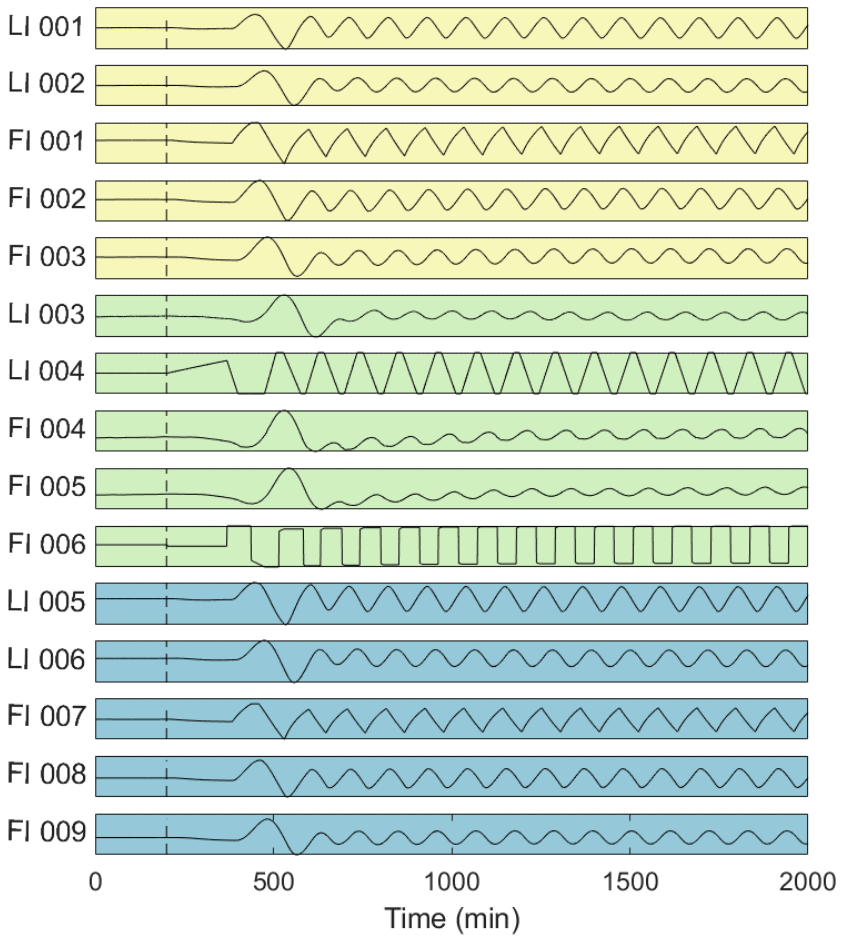


Figure 9: Time series data generated using tank network simulation with and without valve stiction. *The vertical dashed line indicates the beginning of the valve stiction. Yellow, PS1; Green, PS2; Blue, PS3.*

Propagation of the oscillation due to valve stiction is further confirmed by the spectral envelope plot (see Section 3.3.1) provided in Figure 10. The peak indicates that $0.0091 \frac{cycles}{min}$ is a common oscillation frequency, which corresponds to a period of around $110\ min$. This matches the period of oscillation observed in the time series where the valve stiction occurs (FI 006 in Figure 9). An oscillation begins and ends at around $630\ min$ and $730\ min$ respectively - again, giving a period of around $110\ min$.

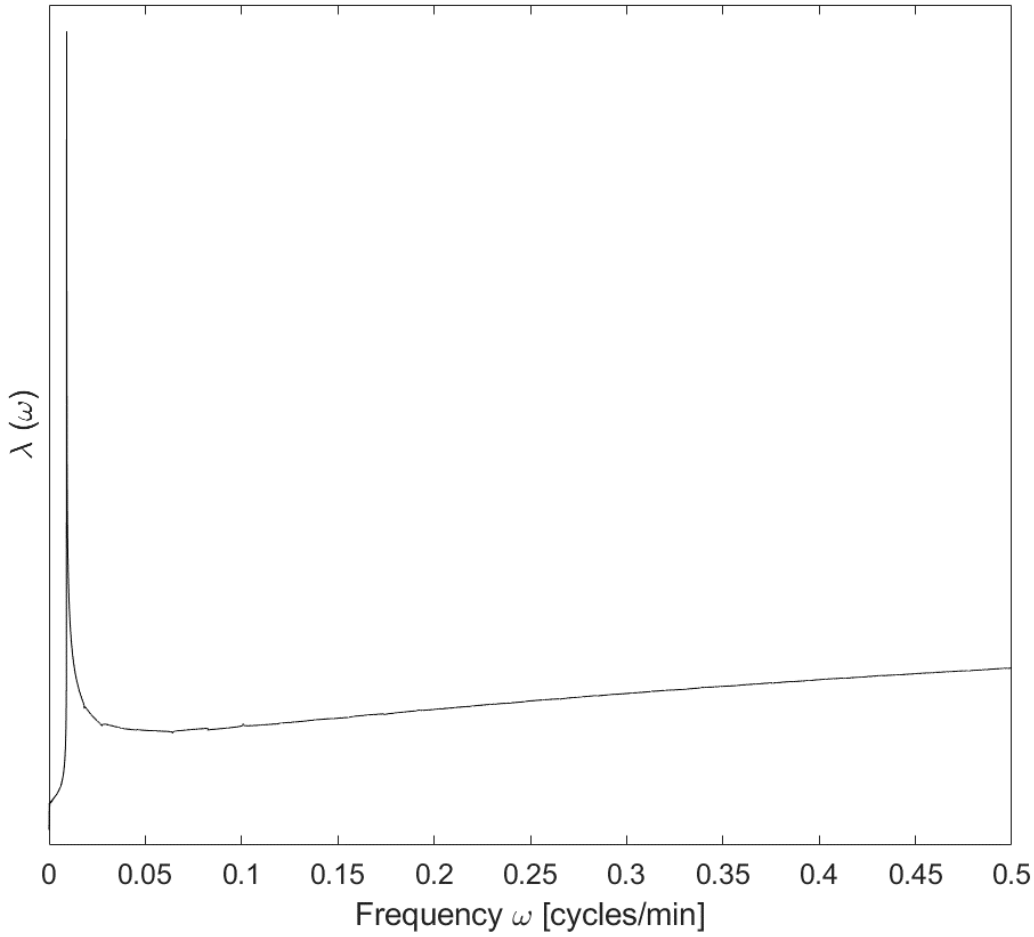


Figure 10: Spectral envelope plot.

CHAPTER 6

Results and Discussion

This chapter presents the causality maps produced using all the approaches described in Chapter 4. In all the results, edge weights represent causal strengths, as explained in Section 3.2.3.

6.1. Definition of the ideal causality map

The definition for an ideal causality map was developed using a top-down approach that consists of the three levels presented in Table 8. The desired characteristics detailed in Level 3 in Table 8 should be strived for, although they cannot always be fully achieved.

Table 8: Definition of the ideal causality map developed according to a three-level top-down approach

Level 1:
The ideal causality map is easily interpretable to successfully trace the propagation path of a fault back to its root cause.
Level 2:
The ideal causality map: o Is relevant to the current fault. o Provides confidence in its connections. o Is visually interpretable.
Level 3:
Relevance to current fault: o The true root cause variable is included in the causality map. o All variables included in the causality map show an effect of the fault. Confidence in connections: o No spurious connections are displayed in the causality map. o All true fault propagation paths are displayed in the causality map. Visual interpretability: o The causality map is not too dense (i.e. does not have too many edges). o The causality map does not have too many nodes. o The causality map clearly points to the root cause, preferably by being acyclic.

The true root cause may not always be individually represented by a node in the causality map, but then it should be included in a group of variables that is represented by a node in the map. An unmeasured true root cause variable cannot be displayed in a data-based causality map, but variables in the same control loop or associated with the same process unit may be measured and should then be included in the causality map.

As mentioned in Section 2.5, data-based causality has the possibility of identifying spurious causal connections, while knowledge-based causality may be outdated or incomplete. Take note that the ideal causality map does not require all true causal connections to be displayed, but merely for all true fault propagation paths to be displayed, as this will still allow the paths to be back-tracked to the root cause. However, it is still not always possible to prevent any spurious

connections on a causality map, whilst also ensuring that all true fault propagation paths are displayed. In response to this dilemma, uncertain connections should be identified as such on the map.

An acyclic causality map is preferable, because the root cause is easily identifiable if the propagation path ends at that specific variable. As mentioned in Section 3.1, cyclic causality maps introduce ambiguity in deciding which variable is the root cause. If an acyclic causality map cannot be achieved, tools such as edge weights and node colours discussed in Section 3.2 can be used to point to the root cause variable.

6.2. Heuristic approach for selecting sampling period and time window

Sampling period (SP) and time window (TW) selection is an important step to ensure a long enough time frame and a sufficient number of samples for conditional GC, but a guideline to perform this selection was identified as a vacancy in current literature (see Section 2.7.5). This section therefore presents a usable heuristic approach for the selection of SP and TW for causality analysis for plant-wide fault identification, for both a standard causality map and the proposed hierarchical approach (see Section 4.5).

6.2.1. Selecting sampling period and time window for a standard causality map

The proposed heuristic approach follows the following steps:

1. Using process knowledge, identify and select a *known* causal connection between two variables showing an effect of the fault (in this work, variables found to oscillate at the common oscillation frequency; see Sections 3.3.1 and 4.1).
2. Select a range of SPs to investigate (discussed in subsection below).
3. Select a range of TWs to investigate (discussed in subsection below).
4. Calculate the causal strength (in this work, the G-causality index; see Section 2.4) for the known causal connection using each combination of the SPs and TWs selected to investigate.
5. Identify and select the combination of SP and TW that results in the largest connection strength of the known connection.

For the case study in this project (see Chapter 5), the known connection $FI\ 006 \rightarrow LI\ 004$ was selected, as $FI\ 006$ is the flowrate where the fault is introduced, and it is an obvious connection, since they form a control loop, where $FI\ 006$ is the MV and $LI\ 004$ the CV.

Range of sampling periods to investigate

The smallest SP to investigate was selected as the SP provided by the sensor (1 min in this project), and the largest SP was selected as the largest SP where an acceptable resolution of the time series data is maintained (see Section 2.7.5). In this project, a strict condition for acceptable resolution of the data was applied, where a SP was deemed too large if it results in a significant change in the waveform observed in the time series. This condition could potentially be relaxed, but not past the point where a false longer period wave appears in the sampled time series data due to an occurrence called the aliasing effect. The resolution of the time series data for different SPs was determined by visually inspecting the time series plots of $FI\ 006$ (i.e. any one of the

variables found to oscillate at the common oscillation frequency) provided in Figure 11. The full set of plots is available in Appendix G.

The largest SP to be investigated was therefore selected as $SP_{27\ min}$, based on the fact that the time series with $SP_{27\ min}$ does not show any sharp peaks like the one observed just after 200 min with $SP_{28\ min}$. The high SP of 27 min is possible without a significant change in the waveform, because the valve stiction (fault) oscillation has long period (around 110 min; see Chapter 5) in this case. If the fault oscillation had a shorter period, an increase in SP would result in a significant change in waveform and the aliasing effect more quickly.

A short fault oscillation period would be problematic if the system being analysed has slow dynamics, as the AR time frame (i.e. model order x SP) in the causality analysis needs to be long enough to capture the process dynamics (see Section 2.7.5). This is specifically applicable if *plant-wide* causality analysis is performed, because variables spread throughout an entire plant take longer to affect each other than variables in only one area of a plant. Plant-wide causality analysis incorporates a multitude of residence times, likely added up in series, as well as recycle streams, which further increase the overall time characteristic (i.e. slow down overall plant dynamics).

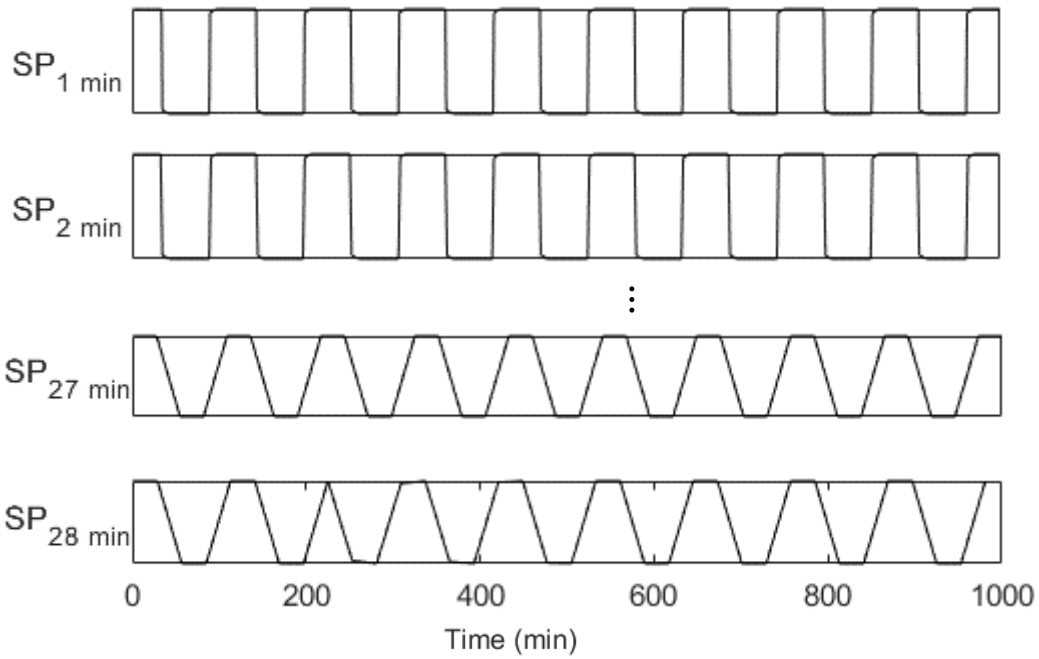


Figure 11: FI 006 time series with various sampling periods.

It seems intuitive to achieve the long AR time frame required for plant-wide causality analysis by simply allowing a high model order in GC, but that leads to overfitting due to the large number of parameters in the AR models, and subsequently to poor GC results. Alternatively, a long AR time frame can be achieved with a smaller model order by using a larger SP; but if the fault oscillation has a short period, this can lead to the aliasing effect. Plant-wide GC therefore only works if the ratio between the fault oscillation period and the process dynamics is large enough to allow the process dynamics to be captured in the time frame by using large enough SP without occurrence of the aliasing effect.

Range of time windows to investigate

In industry, the TW is limited by the size of the available dataset where the fault is present. However, it should be kept in mind that conditional GC relies on fitting the time series data to

AR models, which requires numerous parameters. In this project, the range of TWs to investigate was therefore selected according to the number of samples required to produce a unique OLS solution for the full model in conditional GC (see Section 2.4) for each SP, which was determined to be 13 562 samples. This is the number of samples required to ensure that there are more samples available than the number of parameters required for the largest allowable model, which was limited to 60 to prevent overfitting. This calculation was performed and a TW selected accordingly for each SP, with the results summarized in Table 9. Since the larger TWs become increasingly unrealistic, the largest TW selected to investigate was limited to 75 days.

From Table 9, it is clear that large amounts of data with the fault present are required to implement conditional GC, falling in the order of weeks and even months. This amount of data with a fault present is not likely to be available in industry, and it is unacceptable for the fault to be present for weeks or months before conditional GC can identify its root cause so it can be addressed. Furthermore, this discourages the idea of automating the standard version of conditional GC in real-time, which is the ultimate desire for a fault identification technique in process monitoring. However, the counter-argument that these constraints are imposed by the arbitrarily set maximum model order (of 60 in this case) should be noted, because it allows for a potential solution where the maximum allowable model order is determined by the available data, and not vice versa.

Table 9: Time windows required to produce 13 562 samples for the selected range of sampling periods

SP (min)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
TW (min)	13562	27124	40686	54248	67810	81372	94934	108496	122058
TW (days)	9	19	28	38	47	57	66	75	85
SP (min)	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
TW (min)	135620	149182	162744	176306	189868	203430	216992	230554	244116
TW (days)	94	104	113	122	132	141	151	160	170
SP (min)	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
TW (min)	257678	271240	284802	298364	311926	325488	339050	352612	366174
TW (days)	179	188	198	207	217	226	235	245	254

Effect of time window and sampling period on the known connection strength

The effects of TW and SP on the strength of the known causal connection $FI\ 006 \rightarrow LI\ 004$, the model order selected by AIC, and the resulting AR time frame (model order x SP) are shown in Figure 12. The largest connection strength is at $TW_{38\ days}$ and $SP_{27\ min}$, which corresponds to an AR time frame of 135 min. This combination of TW and SP was selected to perform all further causality analysis to produce a standard causality map in this project. It can be observed that the connection strength does not exhibit a singular trend as a function of either TW or SP, which could correspond to the "black" and "sweet" spots as the sampling period interacts with the underlying process dynamics, reported by Barnett and Seth, (2017) (see Section 2.7.5).

The selected model order is observed to increase with increasing TW, and decrease with increasing SP. This is because the model order is selected using AIC, which rewards goodness of fit versus the ratio of number of parameters to number of samples. As the TW increases, the number of samples increases, so AIC allows a larger model order to be selected; and as the SP increases, the number of samples decreases, so AIC allows a smaller model order to be selected. Since the largest TW started with the most samples, it can work with the larger model order for longer.

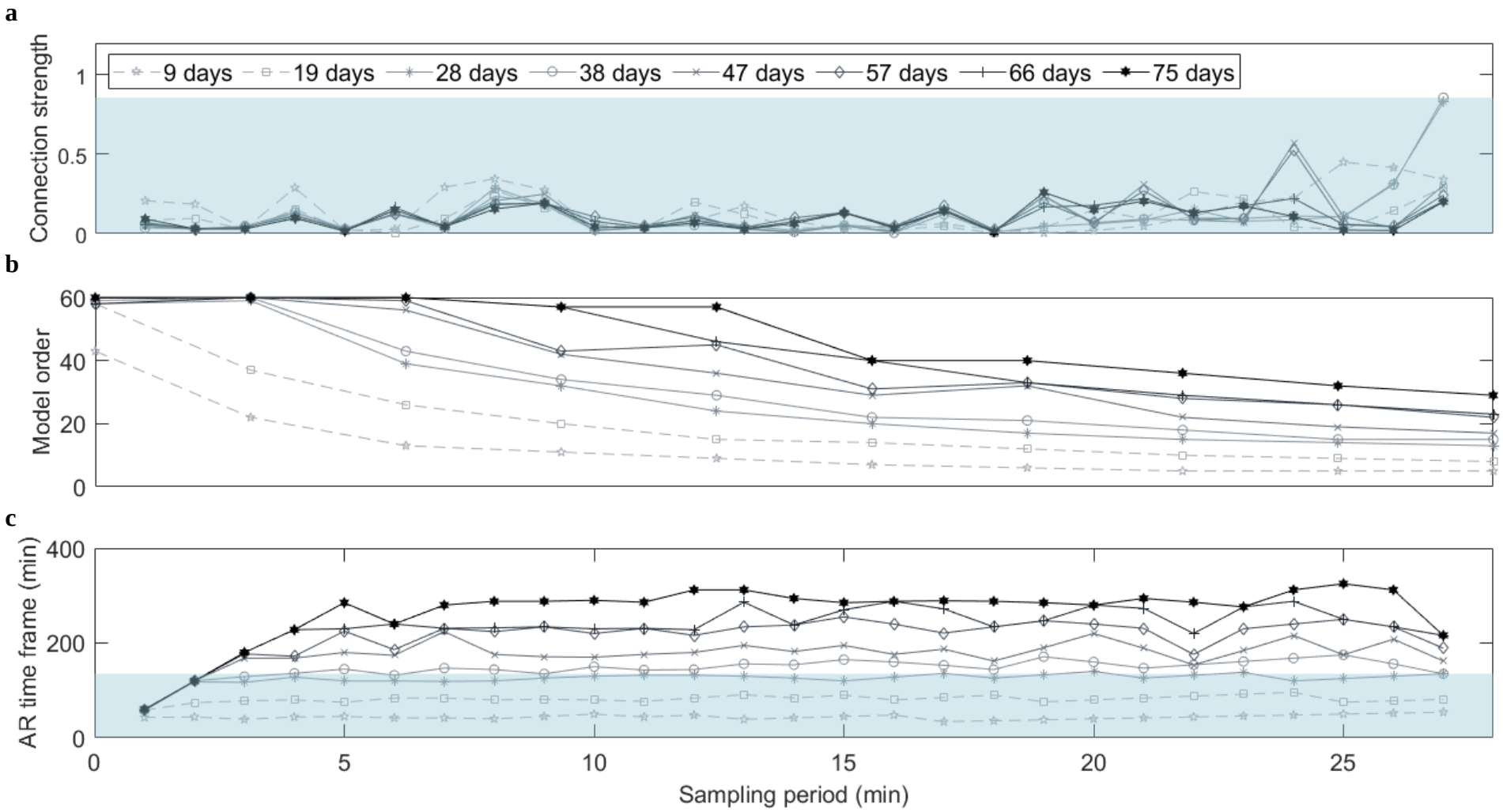


Figure 12: The effects of time window and sampling period on the strength of **a**: the known causal connection FI 006→LI 004, **b**: the model order selected by AIC, and **c**: the resulting autoregressive (AR) time frame. *Darker lines indicate larger TWs. The shaded background lines up with the largest connections strength(s) and the corresponding AR time frame.*

However, the model order does not always decrease with increasing SP. At some points, model order is observed to increase with increasing SP, such as at $SP_{5\ min}$ and $TW_{57\ days}$, and $SP_{7\ min}$ and $TW_{47\ days}$. This is because the selected model order falls within the range where variation in AIC with an increase in model order becomes negligibly small and prone to be affected by noise, as illustrated for the mentioned cases in Figure 13. The difference in AIC for the model orders within this range is negligibly small, so the choice of model order within that range is almost arbitrary, which is why the selected model order is sometimes observed to increase where it is expected to decrease. In addition, AR time frame is the product of SP and model order, so these unexpected increases that are observed in model order are also seen in the corresponding AR time frames.

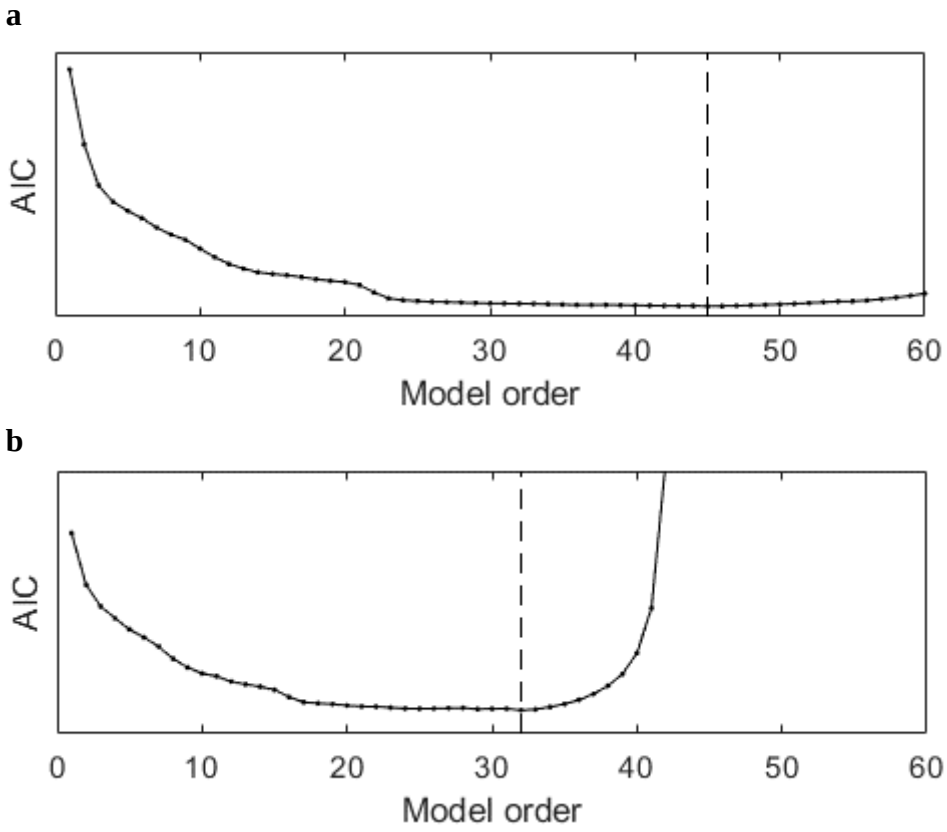


Figure 13: AIC plots for the range of allowable mode orders, with vertical dashed lines indicating the selected model order. **a**. AIC plot for $SP_{5\ min}$ and $TW_{57\ days}$. **b**. AIC plot for $SP_{7\ min}$ and $TW_{47\ days}$. *For this plot, AIC is only calculated for a model order up to 45, because too few samples are available for a unique OLS solution for the full model at higher model orders for this TW.*

The AR time frame is also observed to increase along with increasing TW, without flattening off as a function of TW at any point, which is contrary to expectation. The AR time frame represents the length of time captured by the AR model in GC, so it is expected to correspond to the time it takes to capture the dynamic response of *LI* 004 to a change in *FI* 006, as that is the connection being investigated. This deviation from the expected behaviour can likely be attributed to the heuristic nature of the AIC.

The dynamic response of *LI* 004 to a step change in *FI* 006 was therefore investigated, with the results presented in Figure 14. The full response of *LI* 004 takes around 400 min, but the majority of the response is captured in just under 200 min – which lines up with the AR time frame corresponding to the peak connection strength observed in Figure 12. The AR time frame is therefore expected to increase with increasing TW until the AR time frame reaches around

200 min, and then remain constant for any further increase in TW. However, Figure 12 shows that the AR time frame keeps increasing with TW at a fixed SP, and does not flatten off as a function of TW. This is likely because AIC does not penalize the number of parameters heavily, and therefore leads to incorrectly large model orders (and hence incorrectly large AR time frames) when large TWs (i.e. large amounts of samples) are available.

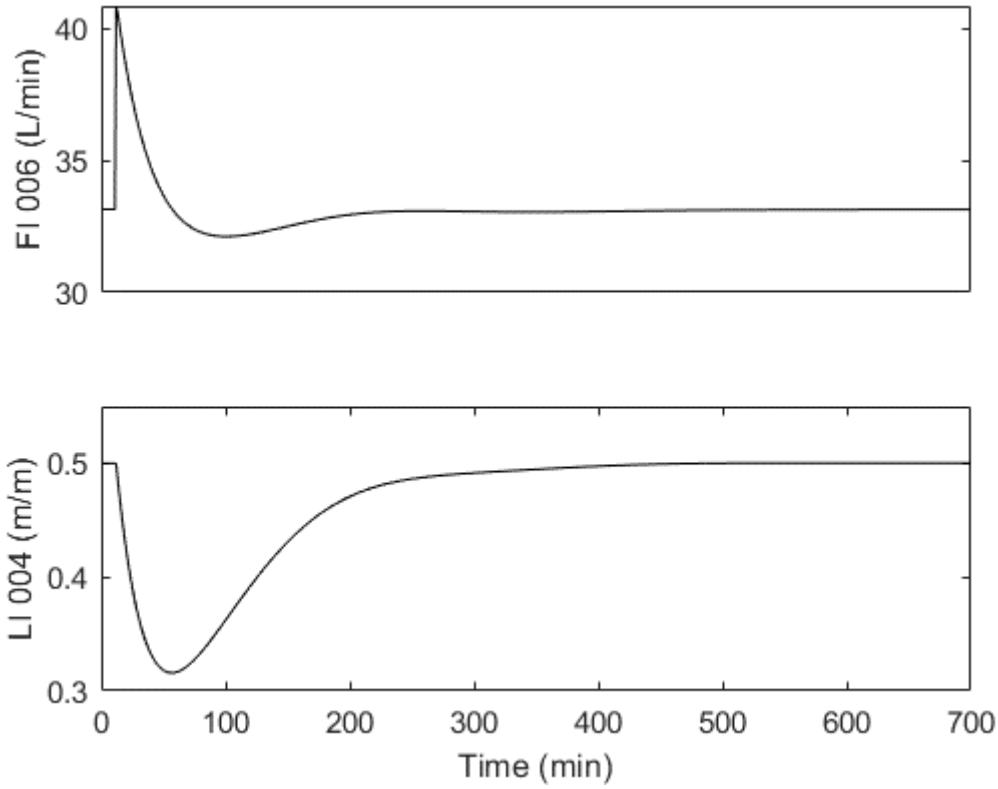


Figure 14: Dynamic response of LI 004 to a step change in FI 006 at 10 min.

6.2.2. Selecting sampling period and time window for the plant-wide causality map in the proposed hierarchical approach

The heuristic approach for selecting SP and TW (explained in Section 6.2.1) was adapted for the plant-wide causality map in the proposed hierarchical approach where PCA is applied for dimensionality reduction (see Section 4.5). The strength of a *known* causal connection is still investigated as a function of both SP and TW, but the connection is now between the representatives (i.e. the first principal components) of two groups of variables. In this project, the connection $PS2 \rightarrow PS1$ was used for the PS-PC1 approach, where variables are grouped according to PSs and the first principal components (PC1s) are used as the representatives for each PS (see Section 4.5).

Range of sampling periods to investigate

The range of SPs to investigate again range from the SP provided by the sensor (1 min in this project) to the largest SP that still captures the fault oscillation, where the scores of PC1 of $PS2$ with different SPs are provided in Figure 15. $SP_{7\ min}$ was selected as the largest SP to be investigated, as the oscillation loses its round shape at the second peak at the next SP (i.e. $SP_{8\ min}$). Since $SP_{27\ min}$ was selected for the standard causality map, $PS2$'s scores with $SP_{27\ min}$ is also provided in Figure 15, but the shape of the fault oscillation is clearly lost at this high SP. The full set of plots at all the SPs is available in Appendix G.

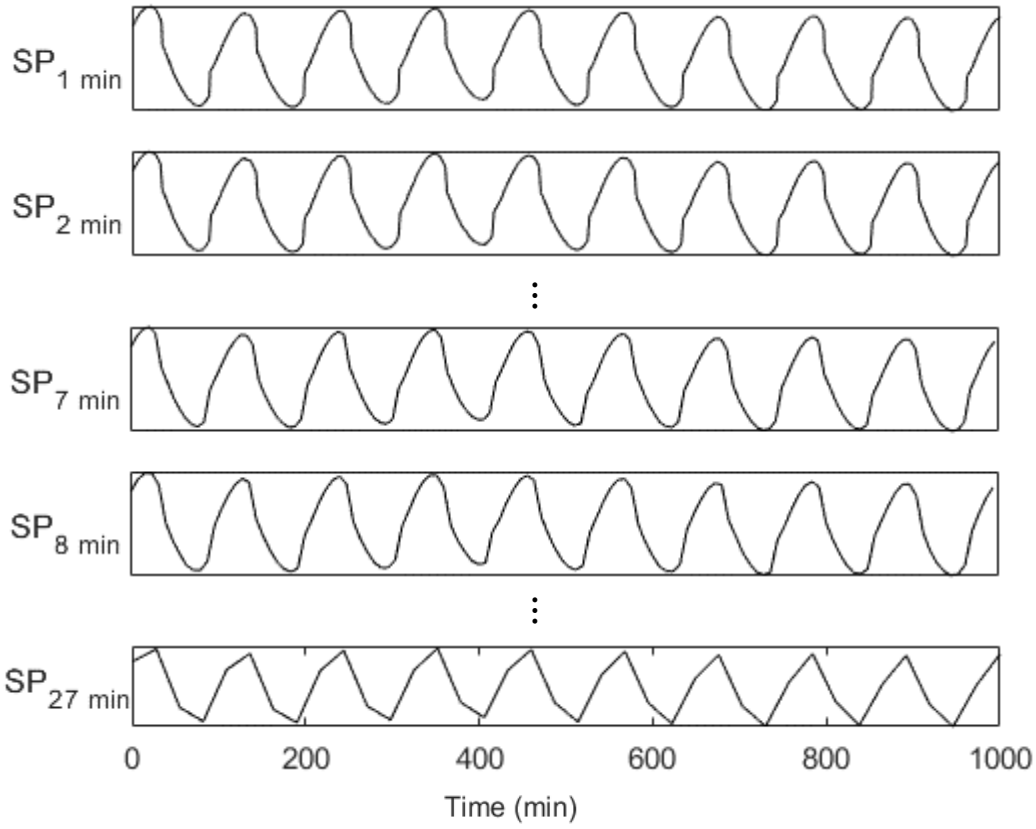


Figure 15: Plant section 2 time series scores with various sampling periods.

Range of time windows to investigate

The plant-wide map for the proposed hierarchical approach has undergone dimensionality reduction, which means that there are fewer variables involved in the conditional GC calculation and therefore fewer parameters are required. A unique OLS solution can be achieved for the full model with only 602 samples, which is significantly fewer than the 13 562 samples required for the standard causality map (see Section 6.2.1) and is more likely to be available in industry. However, here a distinction should be made between PS-PC1 and Mod-PC1 (see Section 4.4), as Mod-PC1 relies on grouping variables according to modules identified in the standard causality map – so this approach still requires the full 13 562 samples necessary to produce the standard causality map (see Section 6.2.1), while PS-PC1 does not. Table 10 summarises the TWs required to provide 602 samples for each SP, and Table 11 summarises the TWs required to provide 2000 samples for each SP, as that is the minimum number of samples suggested by Bauer et al., (2007) (albeit for transfer entropy and not specifically GC).

Table 10: Time windows required to produce 602 samples for the selected range of sampling periods

SF	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
TW (min)	602	1204	1806	2408	3010	3612	4214	4816	5418
TW (days)	0.4	0.8	1.3	1.7	2.1	2.5	2.9	3.3	3.8
SF	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
TW (min)	6020	6622	7224	7826	8428	9030	9632	10234	10836
TW (days)	4.2	4.6	5.0	5.4	5.9	6.3	6.7	7.1	7.5
SF	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
TW (min)	11438	12040	12642	13244	13846	14448	15050	15652	16254
TW (days)	7.9	8.4	8.8	9.2	9.6	10.0	10.5	10.9	11.3

Table 11: Time windows required to produce 2000 samples for the selected range of sampling periods

SF	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
TW (min)	2000	4000	6000	8000	10000	12000	14000	16000	18000
TW (days)	1.4	2.8	4.2	5.6	6.9	8.3	9.7	11.1	12.5
SF	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
TW (min)	20000	22000	24000	26000	28000	30000	32000	34000	36000
TW (days)	13.9	15.3	16.7	18.1	19.4	20.8	22.2	23.6	25.0
SF	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
TW (min)	38000	40000	42000	44000	46000	48000	50000	52000	54000
TW (days)	26.4	27.8	29.2	30.6	31.9	33.3	34.7	36.1	37.5

Effect of time window and sampling period on the known connection strength

The effects of TW and SP on the strength of the known causal connection $PS2 \rightarrow PS1$, the model order selected by AIC, and the resulting AR time frame (model order x SP) are shown in Figure 16. The AR time frame is once again observed to increase with increasing TW for all TWs investigated, indicating that AIC does not perform well for model order selection when provided with an unnecessarily large number of samples - and a minimum of 2000 samples should not be enforced for conditional GC using AIC for model order selection. To improve clarity on the plot, only the data for the shorter TWs are provided in Figure 16 but the rest of the data is plotted in Figure 45 and Figure 46 in Appendix G.

Model order is again observed to increase at some points where it is expected to decrease, as a result of the flattening out of the AIC curve as a function of model order (see Section 6.1.1). To illustrate this once again, AIC plots for the points at $SP_{3\ min}$ and $TW_{1.3\ days}$, and $SP6_{\ min}$ and $TW_{2.9\ days}$ are available in Appendix G.

Large connection strengths are observed at $SP_{2\ min}$ for $TW_{0.8\ days}$, as well as $SP_{5\ min}$ for both $TW_{2.1\ days}$ and $TW_{2.5\ days}$. The decision of which combination of SP and TW to select was therefore based on both process knowledge and convenience. Firstly, a $SP_{2\ min}$ for $TW_{0.8\ days}$ was disregarded, as it provides an AR time frame of 44 min, which only captures the downwards slope of the dynamic response of LI 004 to a step change in FI 006 and so misses distinctive part of the response (see the step test results in Figure 14). A step test may not be available in industry, but the plant operator/engineer may have process knowledge that can indicate a lower cut-off point for the AR time frame that should be captured in the known causal connection.

A $SP_{5\ min}$ for both $TW_{2.1\ days}$ and $TW_{2.5\ days}$ provide an AR time frame of 130 min which is close to the AR time frame of 135 min at the largest connection strength for the original map (see Figure 12 in Section 6.2.1). This again indicates that an AR time frame of around 130-135 min is a 'sweet spot' where the sampling period interacts with the underlying process dynamics. $TW_{2.1\ days}$ was selected over $TW_{2.5\ days}$ at $SP_{5\ min}$ for the plant-wide causality map where variables are grouped according to PSs and PC1s are used as the representatives, simply because a smaller TW equates to a smaller data requirement, which is more likely to be available in industry and decreases the computational load of the causality analysis calculation.

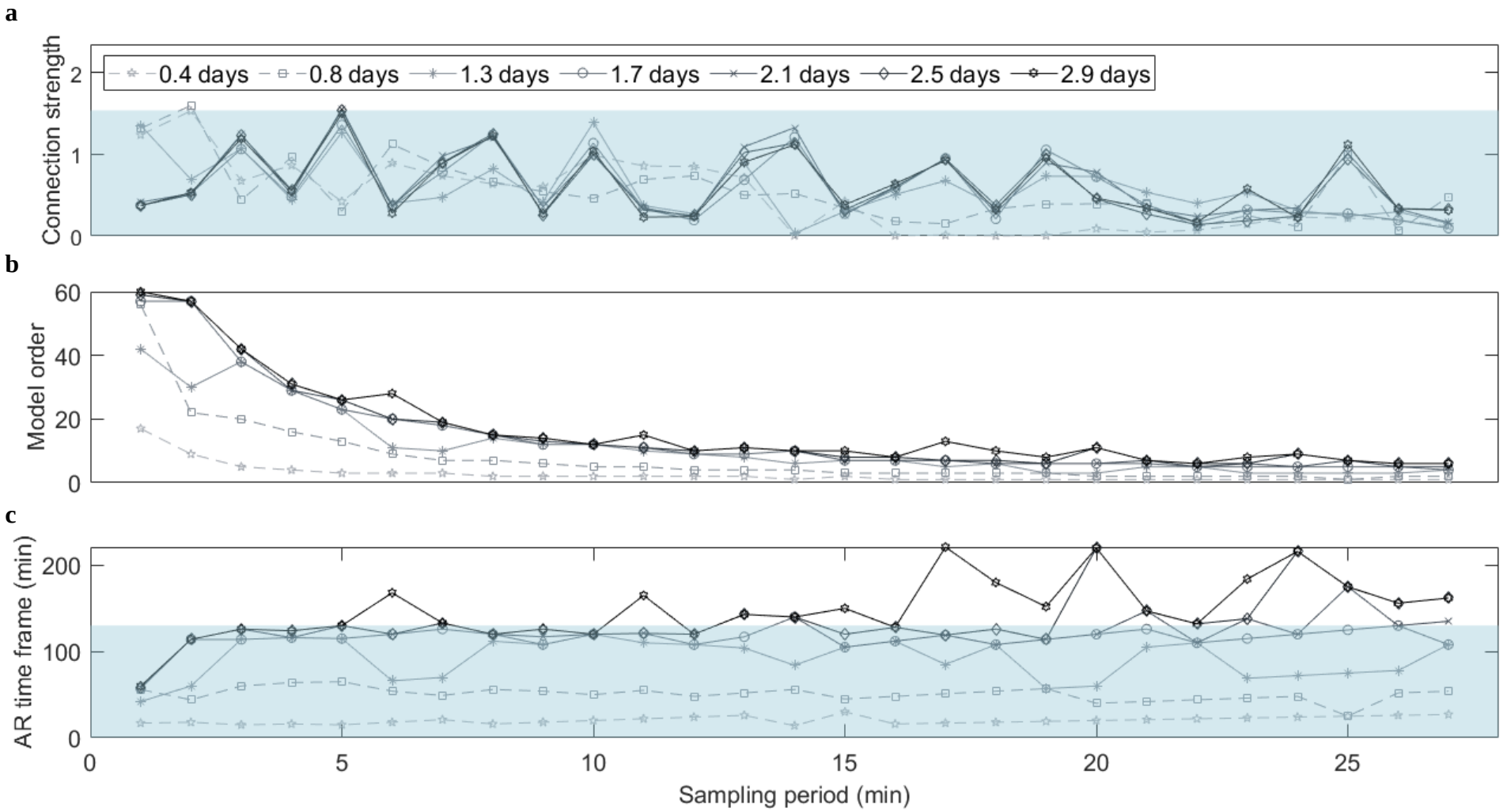


Figure 16: The effects of time window and sampling period on **a**: the strength of the known causal connection PS2→PS1, **b**: the model order selected by AIC, and **c**: the resulting autoregressive (AR) time frame. *Darker lines indicate larger TWs. The shaded background lines up with the largest connections strength(s) and the corresponding AR time frame.*

The proposed hierarchical approach where variables are grouped according to modules identified in the standard causality map (Mod-PC1; see Section 4.5) incorporates less process knowledge than the PS-PC1 approach where grouping is done according to PSs. In Mod-PC1, there is no clear direction on which known connection strength to test against SP and TW, as variables from different PSs can be grouped in the same module. A possible solution is to use the SP and TW selected for the standard causality map, but this is likely not the best option. As discussed previously, AIC selects incorrectly large model orders (and hence AR time frames) when the ratio of samples to parameters is large. Since the plant-wide map for this hierarchical approach has undergone dimensionality reduction (i.e. has fewer variables than the standard causality map), the combination of SP and TW selected for the standard map will likely result in an incorrectly large AR time frame and therefore poor GC results. The best option available is therefore to use the SP and TW selected for the proposed hierarchical approach where the variables are grouped according to PSs, as a similar number of variables are likely to be present in the causality analysis for the plant-wide map.

6.2.3. Evaluation of heuristic approach for selecting sampling period and time window

There does not seem to be a general SP and TW that will produce the most accurate causality map for a range of scenarios; rather, selecting the SP and TW should be done on a case-by-case basis, potentially with some form of the heuristic approach presented in this section. However, Figure 17 indicates that the heuristic approach as presented in this section, applied in conjunction with conditional GC, does not provide the ideal causality map for the case study used in this project (see Chapter 5). The standard causality map is used for evaluation of the heuristic approach presented in this section, as the causality maps for proposed hierarchical approaches are discussed separately in Section 6.5.

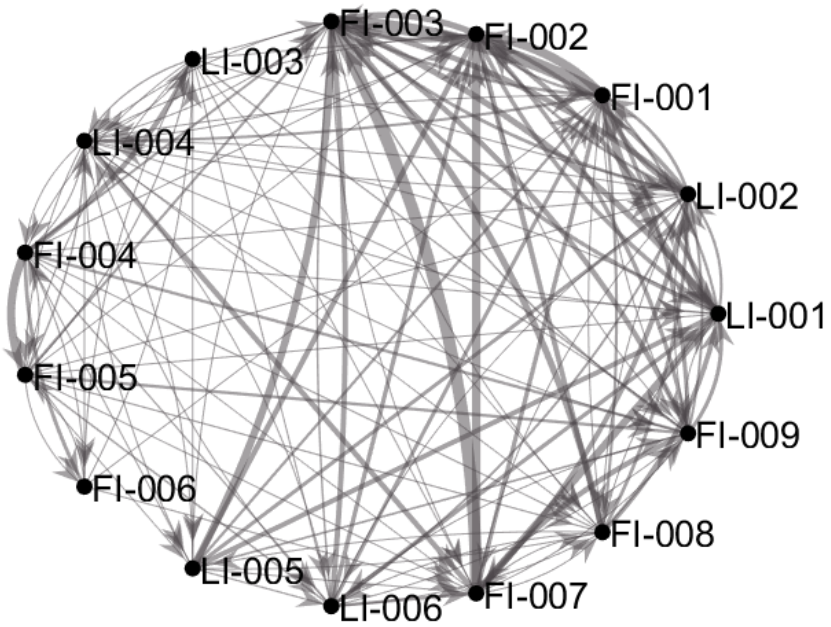


Figure 17: Standard causality map constructed using the SP and TW selected via the heuristic approach presented in this section. *The root cause is a faulty valve on line FI 006.*

Although the causality map (Figure 17) is dense and difficult to interpret, the thicker lines (indicating stronger causal connections; see Section 3.2.3) are not in close proximity to *FI* 006

and $LI\ 004$, which are the variables in the control loop where the true root cause is (see Chapter 5). In fact, based on the causality map in Figure 17, the root cause variable appears to be somewhere in PS1, while the true root cause is in PS2. This is despite the fact that a connection ($FI\ 006 \rightarrow LI\ 004$) focal to the true root cause was tested against SP and TW in the heuristic approach. In industry, there may be ambiguity in the decision of which known connection strength to investigate. Although a causal connection may be true, it may not be focal during the present fault and should therefore not be highlighted in the causality map for fault identification – but this will not be known prior to identification of the fault. This highlights the importance of incorporating process knowledge in causality analysis for fault identification, which is further discussed in Section 6.3.

The heuristic approach presented in this section attempted to improve the use of conditional GC for fault identification so that the true root cause can be accurately identified. However, this goal has not been achieved, likely because plant-wide causality analysis involves a variety of different residence times and time delays. It is therefore not recommended to use standard conditional GC for *plant-wide* fault identification.

6.3. Causality maps incorporating process knowledge

Data-based and knowledge-based causality analysis methods can be combined in a hybrid approach, to exploit the advantages of both methods (see Section 2.5). This section provides the results of incorporating process knowledge in data-based causality maps in two different ways: namely, by using connectivity information from the P&ID to constrain connections to improve accuracy and trust in the causality map; and by incorporating process knowledge to provide additional insights to constrain potential root causes.

6.3.1. Constrained connections

The causal connections identified by conditional GC were validated using the connectivity matrix (see Appendix F) obtained from the P&ID of the process (see Chapter 5). The data-based connections were validated using the connectivity matrix (see Section 4.2.1), as opposed to vice versa (i.e. validating knowledge-based connections using data-based causality analysis), to avoid restricting connections in the causality map to those present in the P&ID (see Section 2.5), as P&IDs in industry are often outdated (see Section 2.3). The causality map in Figure 18 contains only the connections that were validated by the connectivity matrix; while the causality map in Figure 19 contains all the connections identified by data-based conditional GC, but displays the validated connections as solid lines (to indicate certain connections) and the non-validated connections as dashed lines (to indicate uncertain connections).

These two causality maps (Figure 18 and Figure 19) are compared to the standard causality map where no process knowledge is incorporated (see Figure 17 in Section 6.1.3). No true propagation paths are excluded from any of the three causality maps, and the true root cause ($FI\ 006$) is present in all three causality maps, but none of them successfully identify it as the root cause. This again indicates that standard conditional GC does not work well for fault identification on a plant-wide basis, as already mentioned in Section 6.1.3.

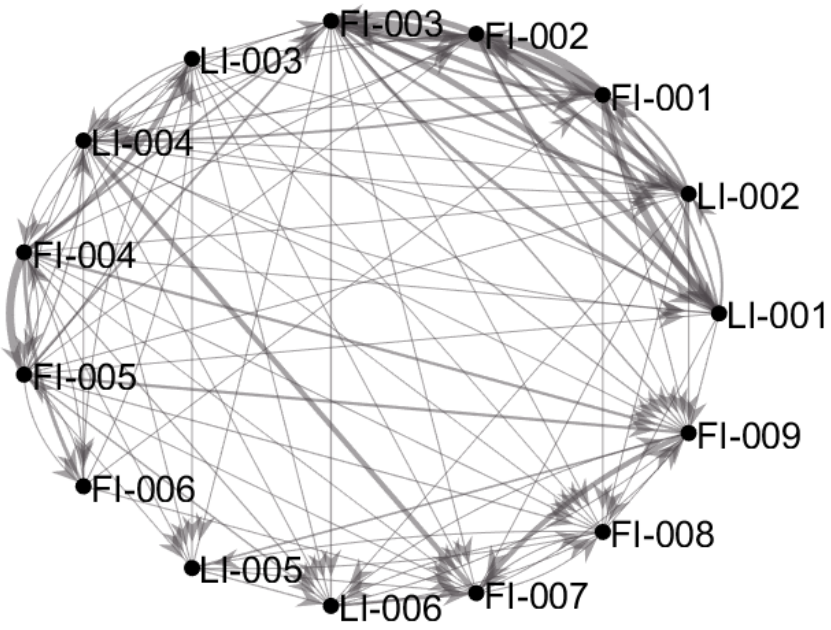


Figure 18: Plant-wide causality map showing only connections validated by the connectivity matrix. *The root cause is a faulty valve on line FI 006.*

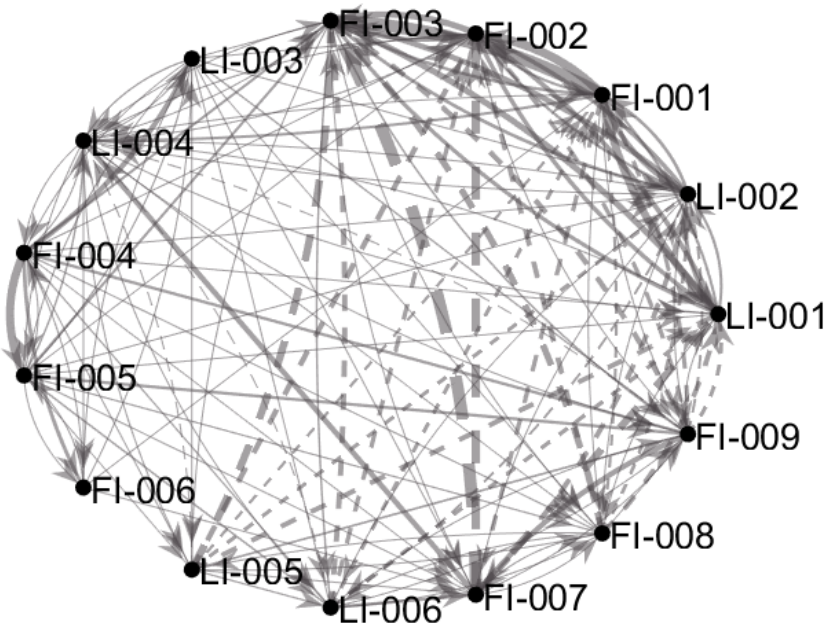


Figure 19: Plant-wide causality map. *Solid lines represent connections validated by the connectivity matrix, and dashed lines represent connections that were not validated by the connectivity matrix. The root cause is a faulty valve on line FI 006.*

The two causality maps incorporating process knowledge (Figure 18 and Figure 19) appear to identify three potential root cause variables (*FI* 001, *FI* 002, *FI* 004), based on the fact that they each have thick edges pointing from them to other nodes. This is fewer than the five potential root cause variables (*LI* 005, *FI* 007, *FI* 001, *FI* 002, *FI* 004) identified from the standard causality map (Figure 17 in Section 6.1.3), based on the fact that they each have thick edges pointing from them to other nodes, and *LI* 005 has a large outdegree and small indegree. Fewer potential root cause variables are identified from the causality maps incorporating process knowledge, because the connections leaving *LI* 005 and the strong

connection leaving $FI\ 007$ are identified as spurious or at least uncertain. However, it should be noted that there is no clear root cause variable in any of these three causality maps, so the process of identifying potential root cause variables from them is subjective and may differ from person to person.

All three maps have 15 variables, as these methods of incorporating process knowledge do not affect the number of variables that are included in the causality maps. Since all the variables in the case study oscillate at the common oscillation frequency (see Chapter 5), all the variables present in all three causality maps fulfil that condition. In addition, all three the maps are cyclic, as the process contains control loops and recycle streams (see Chapter 5) – so even removing all spurious connections still leaves a cyclic causality map.

The major differences between the causality maps are the fraction of connections that are spurious and the graph density. Spurious connections were identified as connections from variables in PS3 to variables in PS1 or PS2, as the variables in PS1 and PS2 can all affect each other and the variables in PS3 due to control loops and recycle streams, but there is no recycle stream or control loop from PS3 to any upstream PSs (see Chapter 5). The standard causality map and the map that shows non-validated connections as dashed lines have 28 spurious connections out of a total of 139 connections, meaning that 20 % of the connections present in these maps are spurious; while the map that displays only validated connections has no spurious connections out of its 111 connections. Since the P&ID used for this case study (and therefore also the connectivity matrix) accurately represents the process, it is noted that all 28 spurious connections, and no true connections (since known causal connections that are not captured via connectivity were added to the connectivity matrix; see Section 4.2.1) were removed via validation with the connectivity matrix. However, it should be noted that the map with dashed lines does identify all the spurious connections as uncertain connections (i.e. displays them all as dashed lines).

The graph density of the standard map and the map that shows non-validated connections as dashed lines is 0.66, while the graph density of the map showing only validated connections is only 0.3. The number of nodes in each map stays the same, so the difference in density is purely due to a decrease in the number of edges when non-validated edges are removed.

In summary, both causality maps with connections constrained according to process knowledge (Figure 18 and Figure 19) are an improvement over the standard causality map with no process knowledge (Figure 17 in Section 6.1.3). They both increase trust in causality maps as a tool for fault identification by users, as all spurious connections are either removed or identified as uncertain respectively – so a user will not notice a spurious connection due to their own process knowledge and lose confidence in the causality map. The causality map that shows only validated lines (Figure 18) is less dense, which aids in interpretability, but it runs the risk of excluding true connections due to an outdated P&ID or the presence of process mechanisms such as heat or mass transfer lacking in a P&ID (see Section 2.5); while the causality map that shows non-validated connections as dashed lines (Figure 19) does not lose any information, but is more dense and hence more difficult to interpret.

6.3.2. Constrained potential root causes

The potential root causes in the causality map were constrained according to which nodes can reach all other nodes in the causality map – based on the assumption that all nodes present in the causality map show an effect of the fault and must therefore be reachable by the root cause via the smearing effect (see Section 2.6). It is therefore advisable to perform variable selection

(see Section 3.3.1) prior to causality analysis if this approach will be implemented. Only PS2 was used to demonstrate this method of incorporating process knowledge, because the data-based plant-wide causality map consists entirely of a strongly connected component (i.e. all the nodes are reachable from each other without violating edge direction; see Section 3.3.2), so no variables would be excluded from the pool of potential root causes.

The resulting causality map is provided in Figure 20, where the only difference from the standard causality map (Figure 17 in Section 6.1.3) is that *LI* 003 is faded to draw attention away from it as it should be excluded from consideration as a potential root cause variable. The faded variable is not completely removed from the causality map, because only measured variables are displayed on a causality map and it is possible that the root cause variable is not measured. In such a case, a causality map is still useful, as it can be used to localise the fault, but the variable(s) that the fault is localised to may not necessarily be able to reach all the variables affected by the smearing effect. It should also be noted that this tool will be more useful in larger systems, as *LI* 003 has an outdegree of zero and would therefore already be excluded from the pool of potential root cause variables in this case.

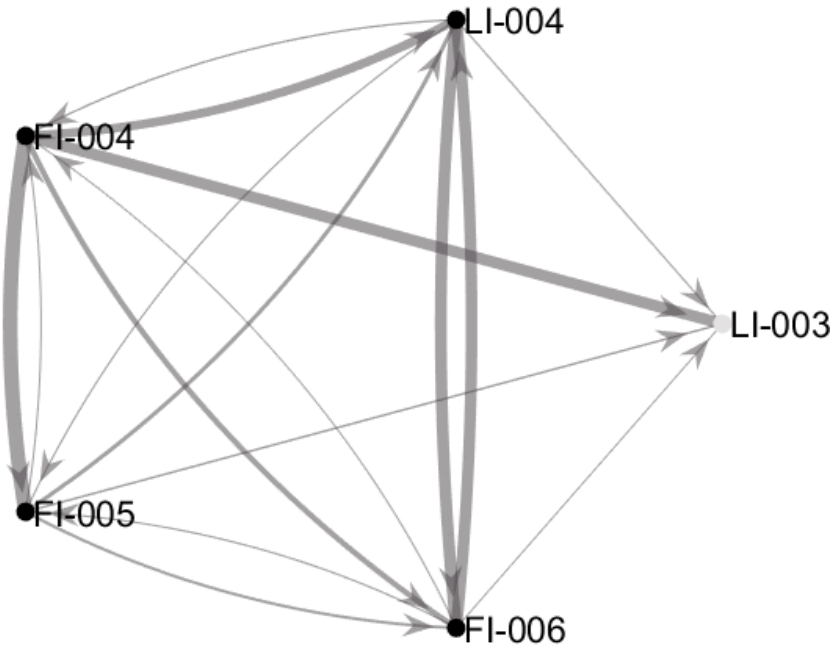


Figure 20: Causality map of PS2, where nodes that cannot reach all other nodes in the map are faded and should be excluded from consideration as root cause variables. *The root cause is a faulty valve on line FI 006.*

6.4. Causality maps incorporating tools to aid interpretation

The interpretability of causality maps can be improved by using various tools. This section therefore presents the results where novel tools are incorporated: namely visually displaying node rankings and incorporating sliders for connections and variables.

6.4.1. Visually displaying node rankings

Nodes in causality maps have been ranked using the page rank algorithm before (see Section 3.2.4), but here those rankings are displayed visually on the causality map using a colour bar (Figure 21). Only PS2 was used to demonstrate this tool, because this smaller map is more

accurate than the plant-wide map, and ranking the nodes only adds value to an accurate causality map.

The only difference from the standard causality map is the colouring of the nodes, so the other characteristics are not discussed. Visually displaying the node rankings is especially useful in this case, as the causality map is cyclic, so there is no clear root cause. In the standard causality map, three potential root causes would likely be identified (*FI* 004, *LI* 004, *FI* 006); while the node rankings help narrow it down to two potential root causes (*FI* 004, *FI* 006), based on the fact that they each have red nodes, which indicates that they have a large effect on the causal structure. It should be noted that the true root cause (FI 006) is included in the pool of potential root causes in both cases, but some users may identify only *FI* 004 as the root cause variable (which would be incorrect) in the causality map with coloured nodes, as it is a darker red, meaning that it is ranked as having the largest influence on the causal structure.

This emphasises that ranking the nodes alone is not sufficient to clearly point to the true root cause. The user should still have sufficient process knowledge to identify the correct root cause variable from the causality map. For example, the control loop consisting of *FI* 006 and *LI* 004 is shown with thick lines (i.e. strong causal connections), while the control loop consisting of *FI* 005 and *LI* 003 (which is more upstream) is shown with thin lines (i.e. weak causal connections); and the only other thick lines in the causality map originate at *FI* 004 and point downstream. *FI* 004 therefore only has a large effect on the causal structure because it is upstream from all the other variables. *FI* 006 and *LI* 004 stand out because they are downstream from the other variables, but are still shown with thick lines (i.e. strong causal connections in the presence of the current fault), which indicates that they are likely where the root cause is located (which is true).

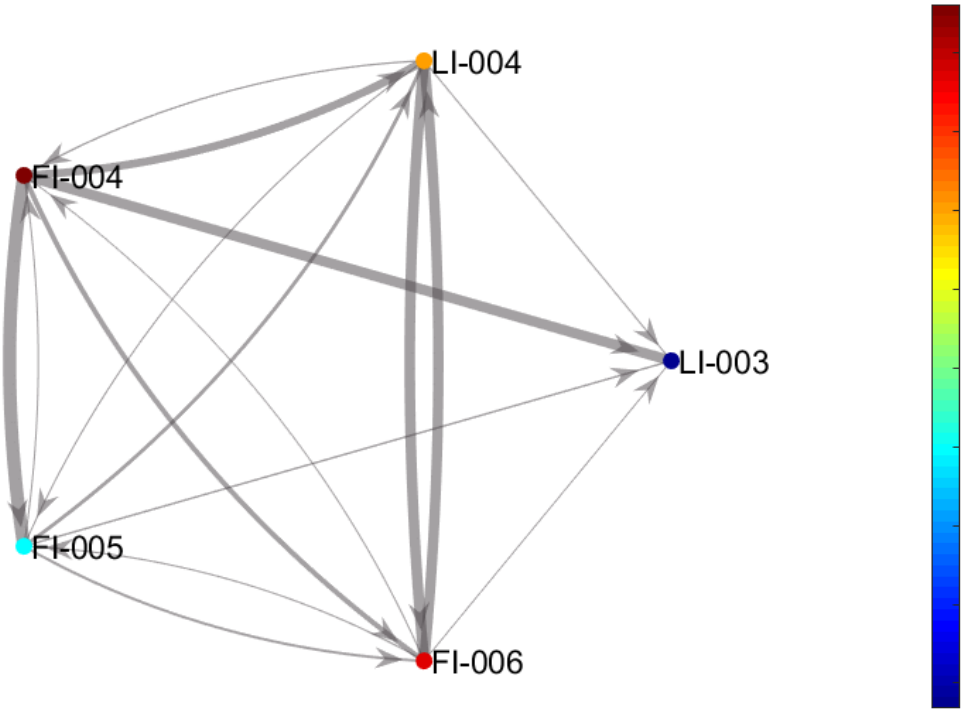


Figure 21: Causality map of PS2, where nodes are coloured to indicate their rank according to the page rank algorithm. *The root cause is a faulty valve on line FI 006.*

6.4.2. Connections-slider

The connections displayed in a causality map are dependent on the significance value (α) in the statistical test (see Section 2.4), but there is no indication of what significance value provides the causality map closest to the ideal causality map, or even whether one specific significance value can provide the most ideal causality map in all scenarios. The idea is to provide a slider for the significance level that the user can manipulate to obtain the causality map closest to the ideal map, as this may be at a different significance level for different scenarios. This idea is communicated in this document by presenting the plant-wide causality map for the case study described in Chapter 5 for three different significant values, where the default value provides the standard causality map (Figure 22).

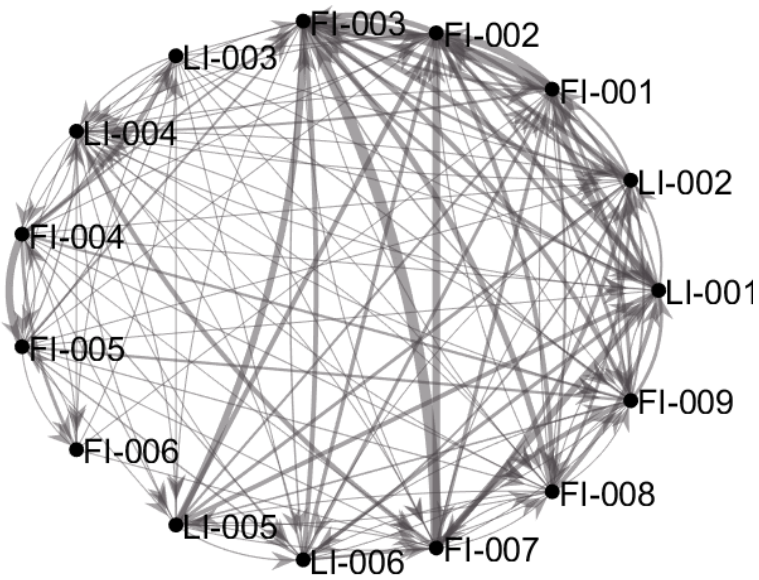
However, a lower significance value is observed to lead to more potential root cause variables, which is contrary to expectation. The number of potential root causes increases from five ($LI\ 005$, $FI\ 007$, $FI\ 001$, $FI\ 002$, $FI\ 004$) - identified based on thick edges pointing from them to other nodes and $LI\ 005$ having a large outdegree and small indegree - in the map with $\alpha = 0.05$ or $\alpha = 0.01$, to seven ($LI\ 005$, $LI\ 006$, $FI\ 007$, $FI\ 008$, $FI\ 001$, $FI\ 002$, $FI\ 004$) in the map with $\alpha = 1 \times 10^{-16}$ – again, based on line thickness and outdegree. This is likely because the decrease in the number of edges allows the user to see more clearly what the indegree and outdegree of each node is. Again, it should be noted that there is no clear root cause variable in any of these three causality maps, so the process of identifying potential root cause variables from them is subjective and may differ from person to person. In addition, none of the causality maps identify the true root cause variable – again, indicating that standard conditional GC does not work well for fault identification on a plant-wide basis, as already mentioned in Section 6.1.3.

6.4.3. Variables-slider

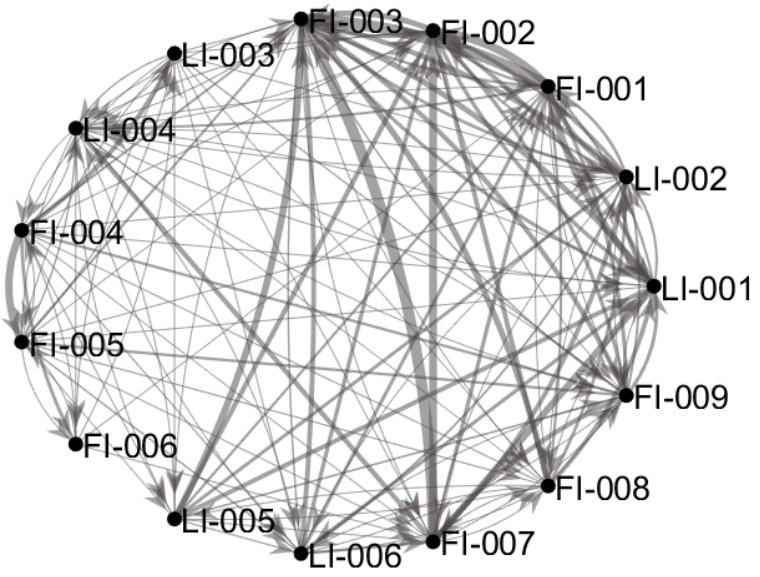
Variable selection prior to causality analysis has been done before (see Section 3.3.1), but again, the ideal causality map could potentially be more closely achieved if the user is able to manipulate the causality map. This section therefore presents the results for two versions of a slider for variable selection, both based on the idea of combining the ability of the spectral envelope method to aid in root cause identification (see Section 3.3.1) with a causality map. The sliders are demonstrated on the plant-wide causality map for the case study described in Chapter 5, but with the low significance level $\alpha = 1 \times 10^{-16}$, so that the different variables included in the causality maps can be more easily identified in the less dense map.

The ideas of both variables-sliders are communicated by presenting the plant-wide causality map for three different threshold values, where the default value provides the standard causality map. The first version of the variables-slider is based on how close a variable's oscillation contribution must be to the oscillation contribution (OC) of FI-006 (the true root cause variable) in order to be included in the causality analysis (Figure 23); while the second version is based purely on how large the OC of each variable is (Figure 24) (see Section 4.3.2).

$\alpha = 0.05$



$\alpha = 0.01$
(default)



$\alpha = 1 \times 10^{-16}$

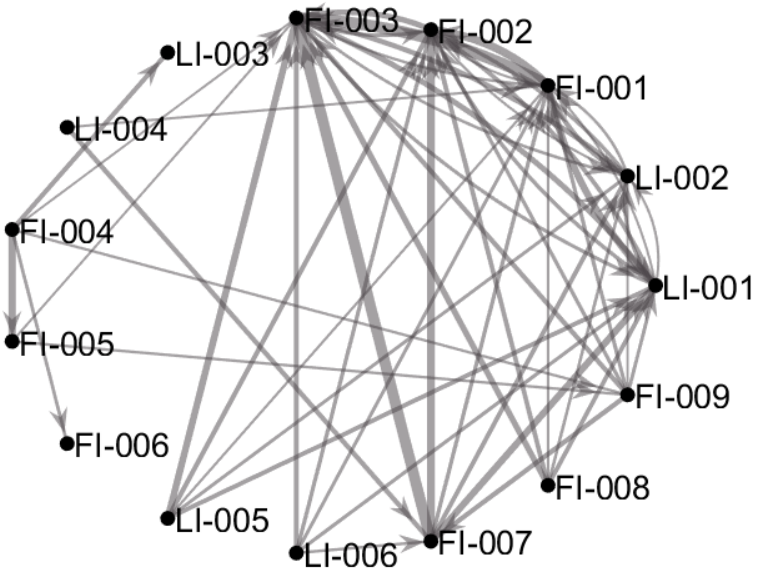
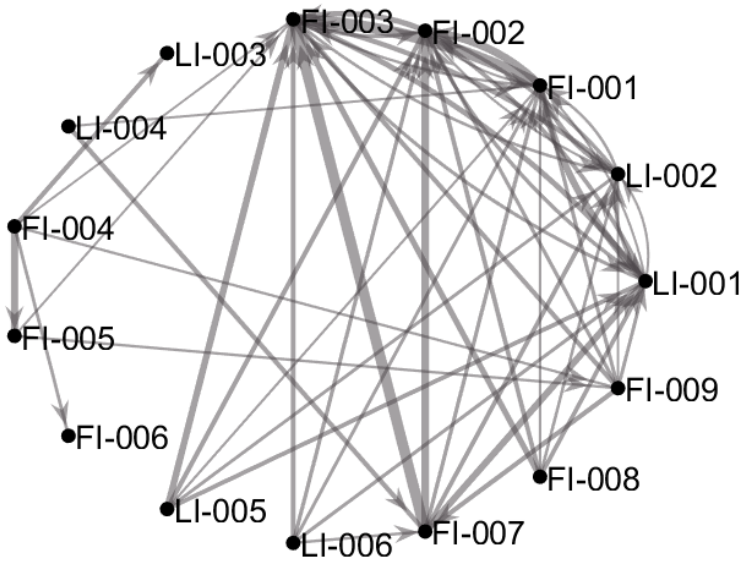
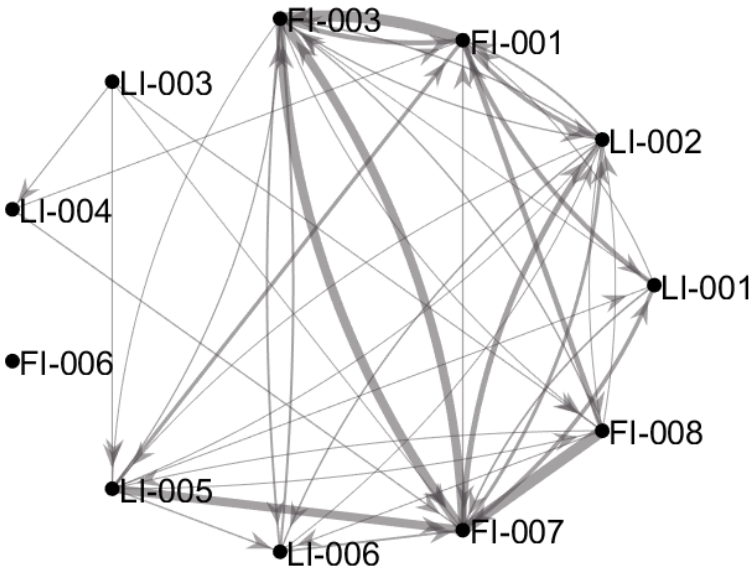


Figure 22: Plant-wide causality map for three different significant values to represent the connections-slider. *The root cause is a faulty valve on line FI 006.*

$|OC - OC_{FI\ 006}| < \mathbf{1}$ (default)



$|OC - OC_{FI\ 006}| < \mathbf{0.2}$



$|OC - OC_{FI\ 006}| < \mathbf{0.05}$

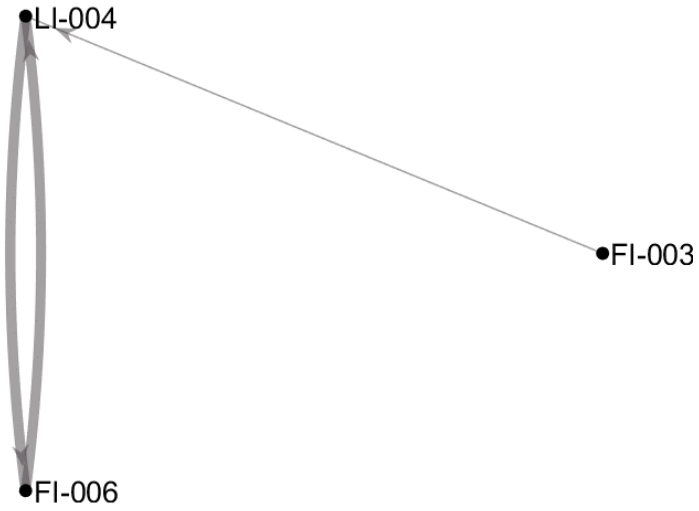
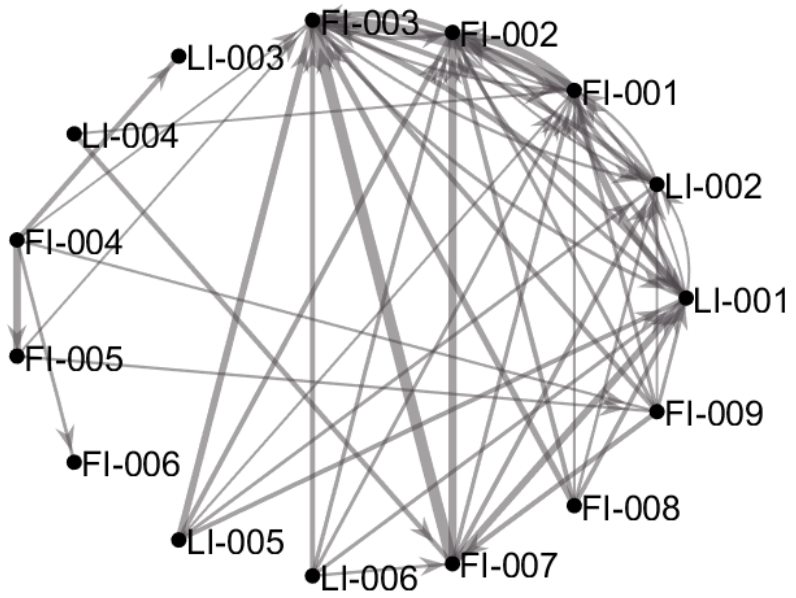
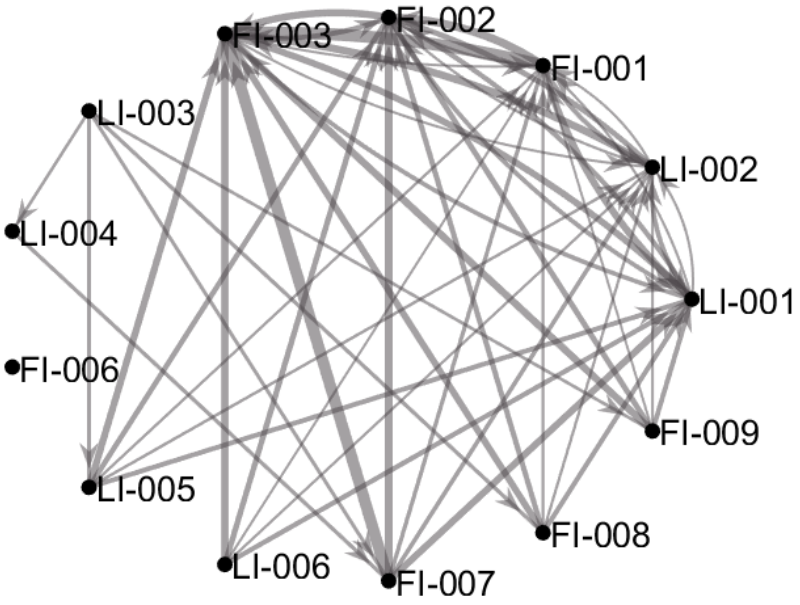


Figure 23: Plant-wide causality map for the first version of the variables-slider, with the threshold $|OC - OC_{FI\ 006}|$ set to three different values. *The root cause is a faulty valve on line FI 006.*

$OC > \mathbf{0}$
(default)



$OC > \mathbf{0.6}$



$OC > \mathbf{0.9}$

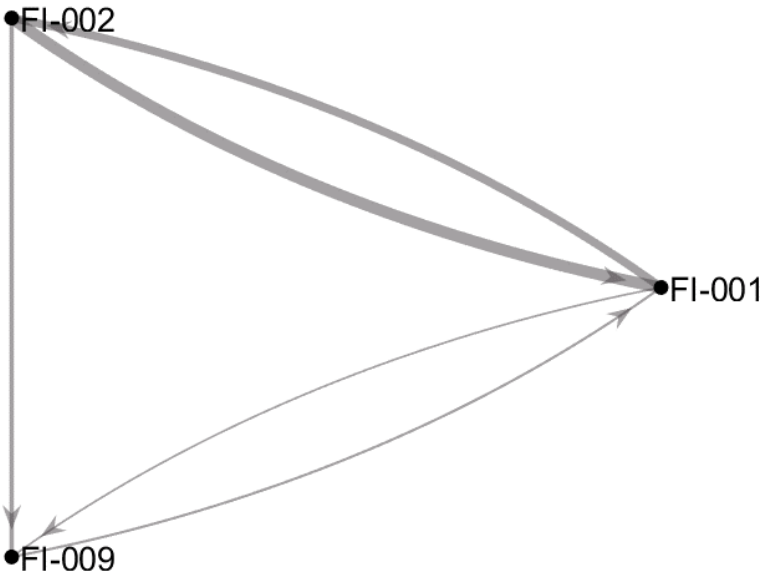


Figure 24: Plant-wide causality map for the second version of the variables-slider, with the threshold OC set to three different values. *The root cause is a faulty valve on line FI 006.*

All the causality maps with both versions of the variables-slider are cyclic, since the process itself is cyclic; and all the variables in all the maps are oscillating at the common oscillation frequency (see Chapter 5). The number of variables for each version of the variables-slider decreases as the threshold becomes stricter. However, for the first version of the variables-slider, the graph density increases (from 0.66 to 0.81 to 1) with decreasing number of variables, despite appearing less complicated to the eye; and for the second version of the variables-slider there is no clear trend in graph density (which changes from 0.66 to 0.65 to 1) with decreasing number of variables.

This is because the variables-sliders affect both the number of edges and the number of nodes in the causality map, as the conditional GC calculation is performed again for each subset of variables, and the results depend on the variables included in the analysis. In addition, fewer nodes improve interpretability, but increase graph density. A complexity metric based on graph density, that includes a term that penalises a graph for a large number of nodes should therefore be considered to evaluate causality maps.

There is no trend in the fraction of spurious connections for either version of the variables-slider. For the first version of the slider, the fraction changes from 0.20 to 0.26 to 0; and for the second version of the slider, the fraction changes from 0.20 to 0.29 to 0.33. In each case, the first map is the default setting (which equates to the standard map), which is why they have the same fraction of spurious connections. The lack of trend is because the variables-sliders can affect both the number of spurious connections identified by conditional GC and the total number of connections. If a variables-slider threshold is made stricter, fewer variables are included in the causality analysis, which means that there are more potentially confounding variables, so more spurious connections are expected to be identified. However, fewer variables being included in the causality analysis also means that the complete graph (where all possible connections are present) has fewer connections, so fewer total connections are expected.

For the first version of the variables-slider (Figure 23), the true root cause variable ($FI\ 006$) is included in each causality map, as the threshold is based on its oscillation contribution – so if only one variable was included, it would be the true root cause in this case (Figure 25a). However, this will not be the case if the threshold was based on a different variable. This approach therefore cannot be applied in industry, since the root cause is not known *a priori*.

In the first two maps (i.e. $|OC - OC_{FI\ 006}| < 1$ or $|OC - OC_{FI\ 006}| < 0.2$) for the first version of the variables-slider (Figure 23), the true root cause (FI 006) has an outdegree of zero, so the true propagation path from it to the rest of the variables is not shown. Furthermore, the first two maps identify six ($LI\ 005$, $FI\ 007$, $FI\ 001$, $FI\ 008$, $FI\ 002$, $FI\ 004$) and four ($LI\ 005$, $FI\ 007$, $FI\ 008$, $FI\ 001$) potential root cause variables respectively – again based on thick edges pointing away from these nodes and/or large outdegrees and small/no indegrees, but bearing in mind that the identification is subjective and may differ from person to person. The last map ($|OC - OC_{FI\ 006}| < 0.05$) correctly identifies the correct control loop consisting of $FI\ 006$ and $LI\ 004$ as the root cause location. However, this is likely because the true root cause variable was chosen as the comparison variable in this case. In industry it would be unclear which variable to compare the oscillation contributions to. In addition, the fact that two out of the three maps do not identify the correct root cause location, make it unlikely that the user would correctly identify the root cause location. In fact, the inconsistency in the potential root causes identified using the different maps may serve to confuse the user.

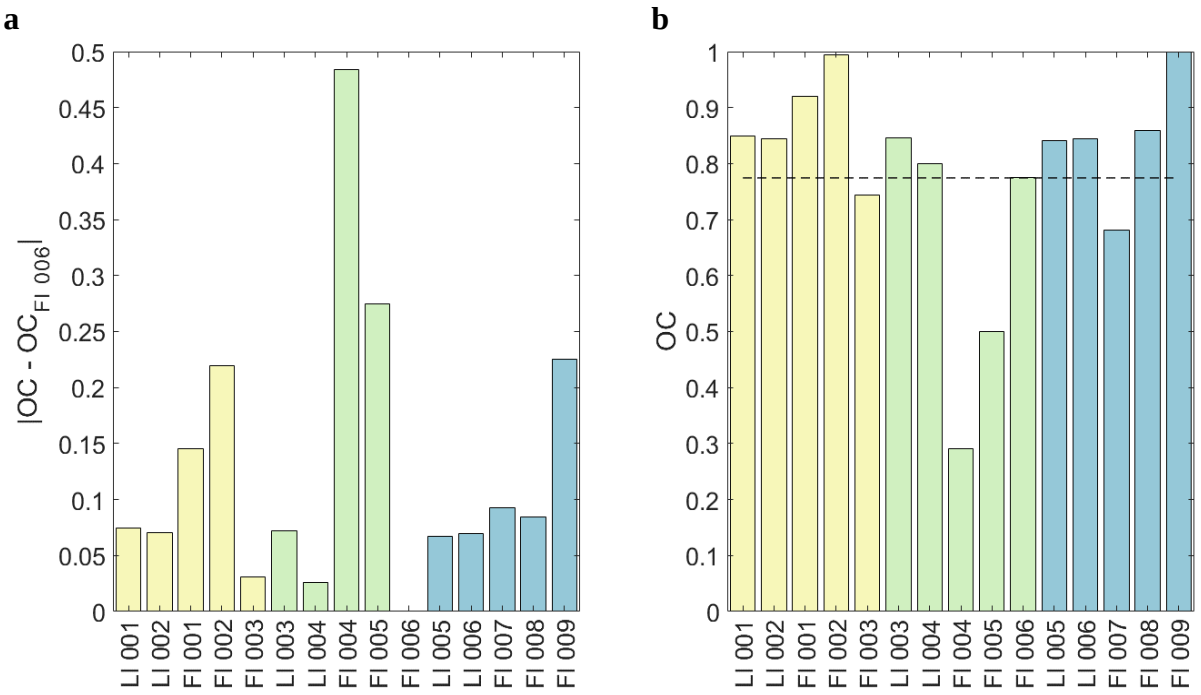


Figure 25: **a**. $|OC - OC_{FI\ 006}|$ for each variable. **b**. OC for each variable. *The dashed line indicates the OC of the true root cause variable (FI 006). The bars are coloured according to PSs (see Chapter 5).*

For the second version of the variables-slider (Figure 24), the true root cause is not included in all the maps. This is contrary to expectation, because literature indicates that the spectral envelope method can be used to narrow down the pool of potential root cause variables based on their OCs. It is therefore expected that the true root cause variable would have the largest OC, but this is clearly not the case, as shown in Figure 25b, where the dashed line indicates the OC of the true root cause (FI 006).

In summary, neither of the two versions of the variables-slider work effectively to aid in interpretation of the causality map(s). However, the existing way of using the spectral envelope method for variable selection is still useful. A statistical test is performed once-off to determine which of the variables are oscillating significantly at the common oscillation frequency, and only those variables are included in the causality analysis (see Section 3.3.1).

In this case, the p-value associated with the test statistic for each variable is larger than 0.999, indicating that all the variables are oscillating significantly at the common oscillation frequency. This is likely because the valve stiction parameters were selected in order to produce an exaggerated oscillation in the case study for this project (see Chapter 5). This may not always be the case, and this method for variable selection prior to the causality analysis calculation can therefore be valuable, because it decreases the number of variables included in the causality analysis. It is especially helpful if conditional GC is being used. As explained in Section 6.2, conditional GC involves a large number of parameters that need to be estimated, which requires large datasets, which may not always be available in industry. Furthermore, this method has only one hyperparameter (the significance value in the statistical test), so it can be applied with confidence, potentially even in an automated variable selection step in industry. This is especially helpful, as conditional GC suffers from hyperparameter selection (especially model order; see Section 6.2).

6.5. Hierarchical approach for causality analysis

The interpretability of causality maps can be improved by decreasing their number of nodes, as mentioned in Section 3.3. This section therefore presents the results for a hierarchical approach, where the variables are grouped according to either PSs or modules in the data, and dimensionality reduction is applied to each group of variables using PCA (see Section 4.5). Only the plant-wide causality maps produced in the first stage of the hierarchical approach (hereafter referred to as *representative causality maps*) are discussed, as the detailed maps in the second stage are simply standard causality maps consisting of the relevant groups of variables.

The representative causality maps are compared to the transitive reduction (see Section 3.2.1) of the standard causality map, provided in Figure 26. Node ranking has been applied to the transitive reduction and the representative maps, to aid in interpretation of the cyclic maps; but only the representative maps use dashed edges to indicate uncertain connections (see Section 6.3.1), as it does not make sense to validate the edges in a transitive reduction with a connectivity matrix (because if non-validated edges are removed, propagation paths that would have been validated may also be removed). In addition, the connections-slider and potential root cause constraint explained in Sections 4.3.2 and 4.2.2 respectively are incorporated in each causality map where their incorporation results in changes in the causality map.

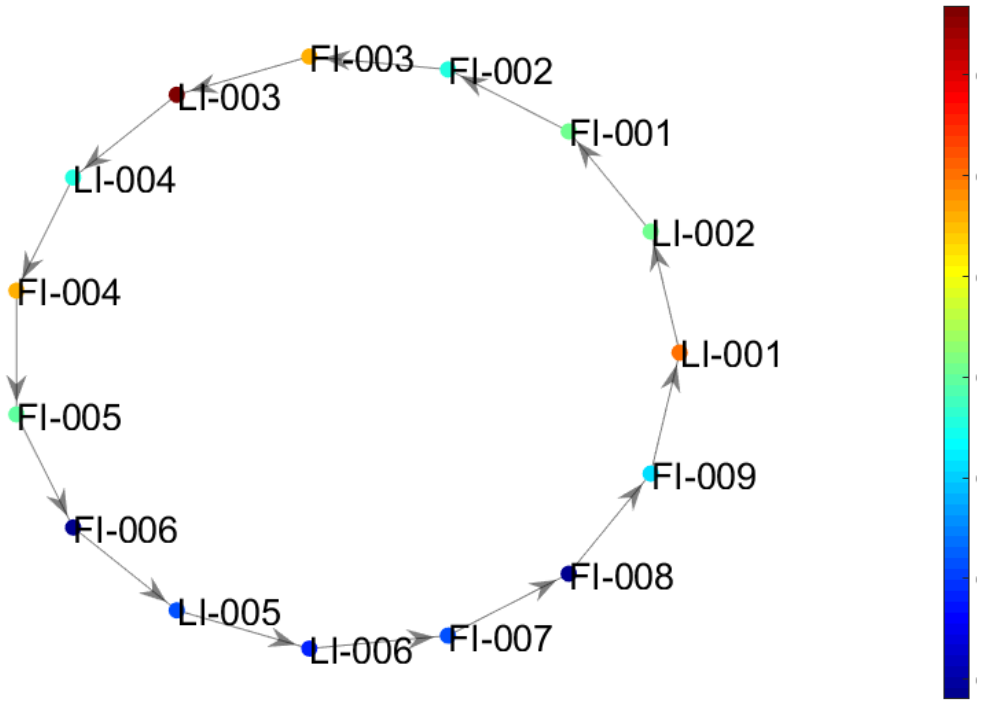


Figure 26: Transitive reduction of the plant-wide causality map. *The root cause is a faulty valve on line FI 006.*

6.5.1. Plant-wide causality maps with nodes representing plant sections

The representative maps, where grouping was done according to PSs and the first principal components (PC1s) were used as representatives of each PS (PS-PC1 causality maps), are shown in Figure 27. The true root cause variable (*FI* 006) is present in all the PS-PC1 maps, as well as the transitive reduction – since *FI* 006 is a measured variable, and neither of these approaches excludes any variables.

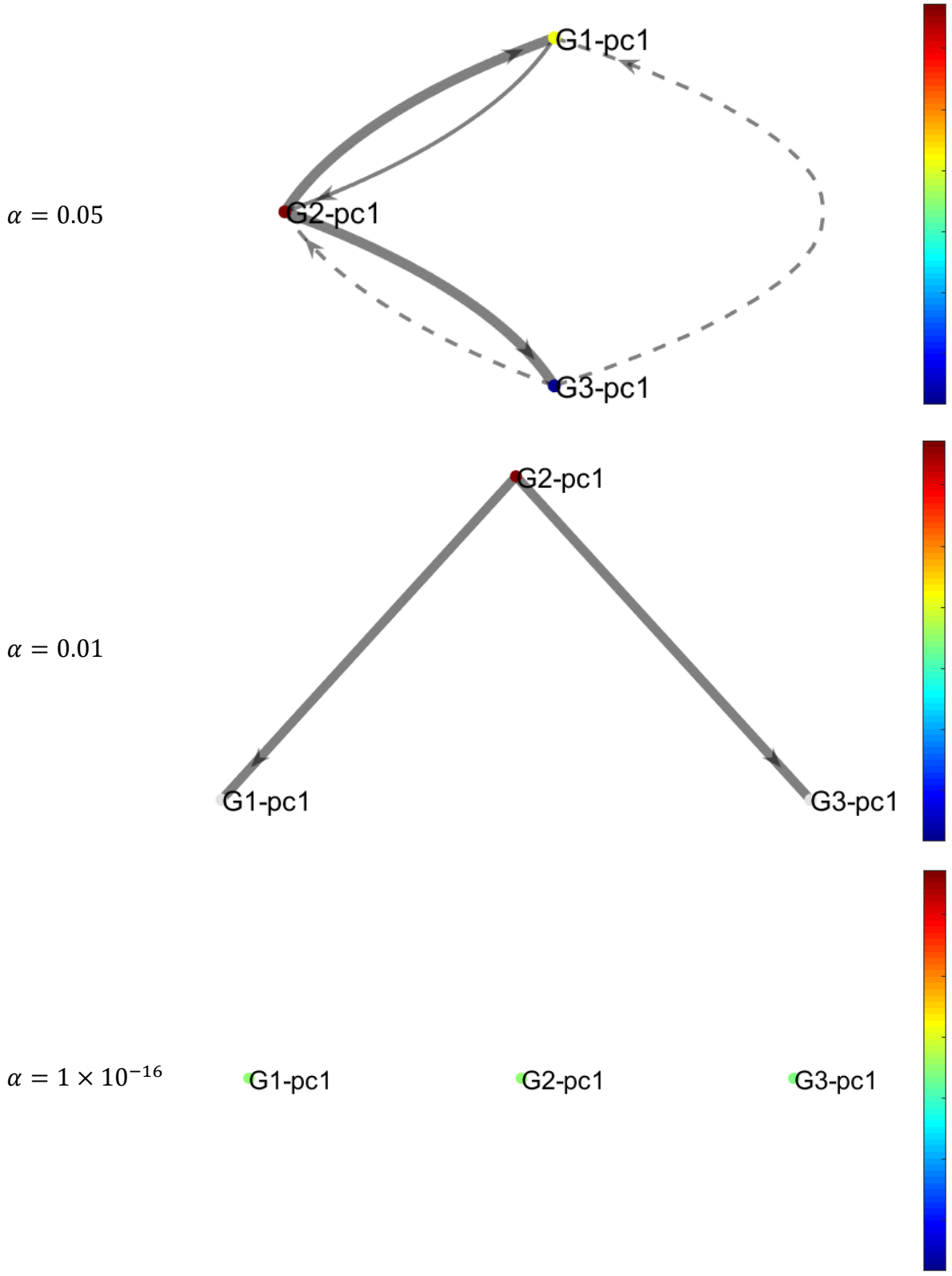


Figure 27: The representative map, where grouping is done according to PSs and PC1s used as representatives *(PS-PC1 causality maps)* for three significance levels. *The root cause is a faulty valve located in G2.*

The PS-PC1 map identifies the true root cause location (PS2, named G2-pc1 on the causality maps) at each significance level, except the lowest one, where the map has no edges. This map is even acyclic at a significance level of 0.01, which is the default level; and both other nodes (G1 and G3) are faded to exclude them as potential root causes. This is an improvement to the transitive reduction map, which is completely cyclic, making it difficult to identify any root cause. If any root cause had to be identified from the transitive reduction, it would likely be $LI\ 003$ (which is incorrect), as that is the only node coloured in red, indicating that it has the largest influence on the causal structure.

Furthermore, the PS-PC1 map only has two spurious connections at a significance level of 0.05. The transitive reduction may only have one spurious connection, but both spurious connections in the PS-PC1 are identified as uncertain via the connectivity matrix (see Section 6.3.1). In addition, the PS-PC1 map has only three nodes, while the transitive reduction has 15 nodes.

The PS-PC1 maps with significance levels of 0.05 and 0.01 have graph densities of 0.83 and 0.33 respectively, while the transitive reduction has a graph density of 0.067. Again, this is not an indication that the transitive reduction is more interpretable than the PS-PC1 map; but rather that graph density is not a good complexity metric, as it does not penalise the number of nodes (see Section 6.4.3).

6.5.2. Plant-wide causality maps with nodes representing modules in the data

The representative maps, where grouping was done according to modules identified in the data and PC1s were used as representatives of each module (Mod-PC1 causality maps), are shown in Figure 28. The true root cause variable ($FI\ 006$) is present in all the Mod-PC1 maps, because $FI\ 006$ is a measured variable, and this approach does not exclude any variables. However, the Mod-PC1 maps do not identify the true root cause location. At significance levels of 0.05 or 0.01, the Mod-PC1 map does not identify any location for the root cause, as there is little difference in line thickness between the two edges, and the nodes are both coloured green. At a significance level of 1×10^{-16}, two variables, namely G1 or G3 could be identified as potential root causes, based on node colouring and line thickness respectively – which are both incorrect, as the true root cause is in G2 (see Table 13).

None of the Mod-PC1 maps have any spurious connections (while the transitive reduction has one spurious connection), but this is not necessarily a positive attribute of this approach. Since each group of variables contains at least one variable from PS1 or PS2 (which have causal connections to all other variables), it is impossible for any connection in any of these causality maps to be spurious. For this same reason, it is difficult to incorporate process knowledge in this causality map, because the variables from PS3 (which do not have a causal effect on any variables in PS1 or PS2) are spread throughout G1, G2, and G3; so none of the connections between representative nodes can be eliminated based on known connectivity. In summary, none of the Mod-PC1 maps are necessarily wrong; they are just not useful. This further reinforces the importance of incorporating as much process knowledge as possible in causality analysis.

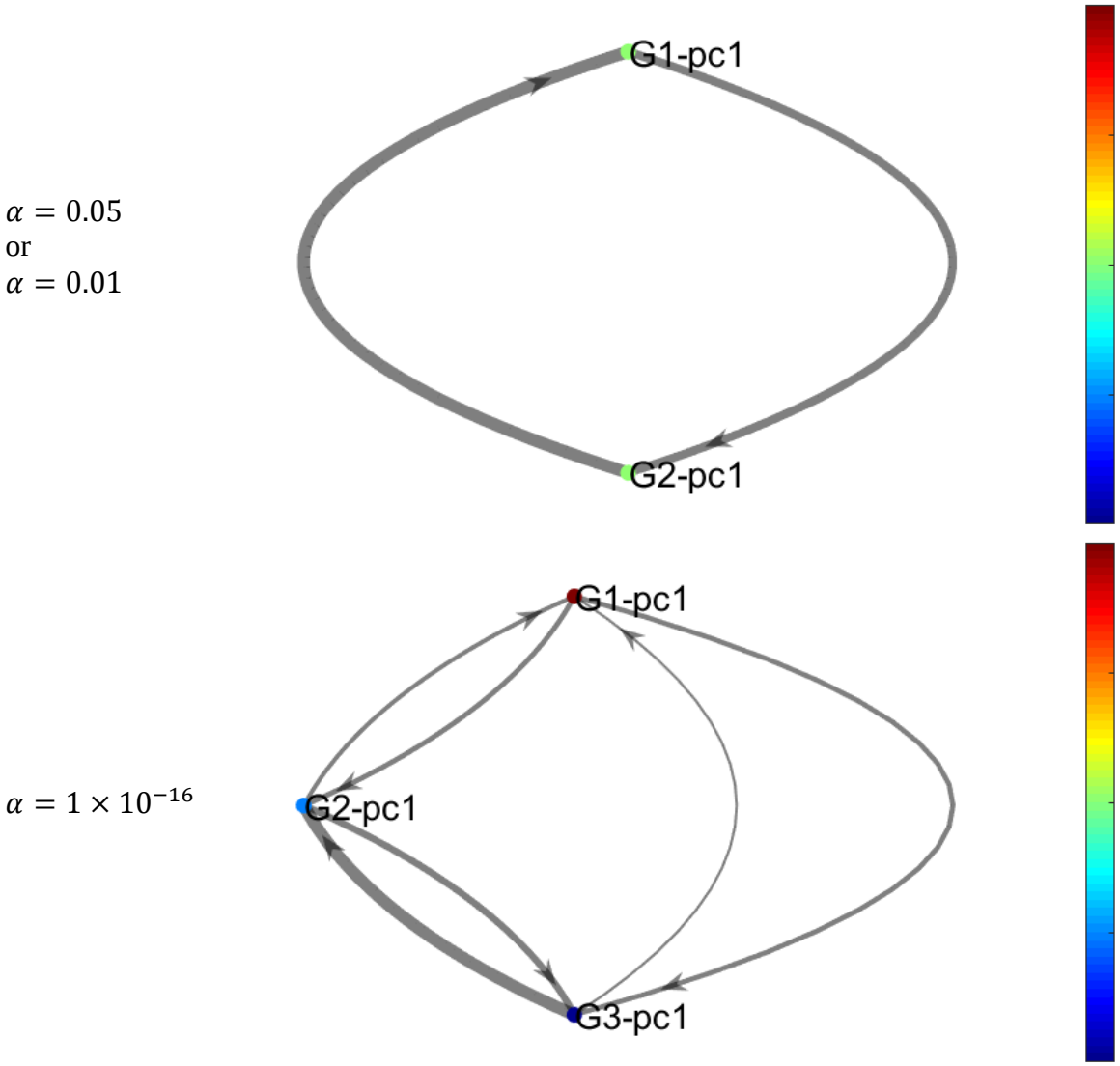


Figure 28: The representative map, where grouping is done according to modules identified in the standard causality map and PC1s used as representatives *(Mod-PC1 causality map)* for three significance levels. Note that the significance levels are not applied to the *Mod-PC1 causality map itself; but to* the standard causality map, from which the modules are identified. *The root cause is a faulty valve located in G1 and G3 respectively.*

Table 12: Variables assigned to each group identified in the standard causality map with a significance level of 0.05 or 0.01. *The root cause variable is indicated in bold.*

G1		G2
LI 003	LI 005	LI 001
LI 004	LI 006	LI 002
FI 004	FI 007	FI 001
FI 005	FI 008	FI 002
FI 006	FI 009	FI 003

Table 13: Variables assigned to each group identified in the standard causality map with a significance level of 1×10^{-16}. *The root cause variable is indicated in bold.*

G1		G2	G3
LI 001	LI 005	LI 004	LI 003
LI 002	FI 008	LI 006	FI 004
FI 001		FI 007	FI 005
FI 002			**FI 006**
FI 003			FI 009

6.6. Survey results

This section presents the aggregated results from the 20 survey responses, of which ten were obtained from SU final-year/postgraduate students, six were obtained from employees of an industrial IoT firm, and four were obtained from Anglo American Platinum (AAP) employees. The results consist of two main parts: first, the results from closed questions, regarding the role of tools/characteristics and sliders in decision-making for fault identification map user-friendliness ratings; second, the thematic analysis of the open-ended questions regarding practical application of the proposed tools and approaches for causality analysis in industry.

It should be noted that the survey was created before the heuristic approach for selecting SP and TW was developed (see Sections 4.1.2 and 6.2), so the causality maps in the survey look slightly different from the causality maps presented in the body of this report. However, this does not detract from the survey results – since no analysis was performed on which variable(s) the participants identified as the root cause, but rather on the results regarding the process of interpreting the causality maps.

6.6.1. Closed questions results

Characteristics/tools used to decide on the root cause

Participants were asked to rank the the extent to which characteristic/tools played a role in their decision-making while identifying a root cause from a causality map from 1 (largest role) to 4 (smallest role), and the results for the base case of transitive reduction (TR), the PS-PC1 approach, and the Mod-PC1 (see Section 4.5) are presented in Figure 29. Node degree incorporates the number of arrows pointing towards and away from node(s) (see Section 2.2); edges type refers to whether edges are solid or dashed to indicate validated or non-validated connections (see Section 6.3.1); node colour refers to node ranking (see Section 6.4.1); and edge weight refers to the strength of connection (see Section 3.2.3).

The ranking for node degree is observed to be evenly spread in all three approaches, so no comment can be made on it. Node colour is ranked first by a large portion of participants in all three approaches (PS-PC1 and Mod-PC1, and TR). This is likely because many of the maps are cyclic, and then node colour is the most noticeable difference between potential root causes.

Edge type (solid/dashed) is not ranked first by many respondents for any of the approaches. This is contrary to expectation for PS-PC1 and Mod-PC1, as it is the characteristic that distinguishes between certain (trustworthy) and uncertain (untrustworthy) connections. Since TR has only solid edges, it is not unexpected that edge type is ranked last for this approach.

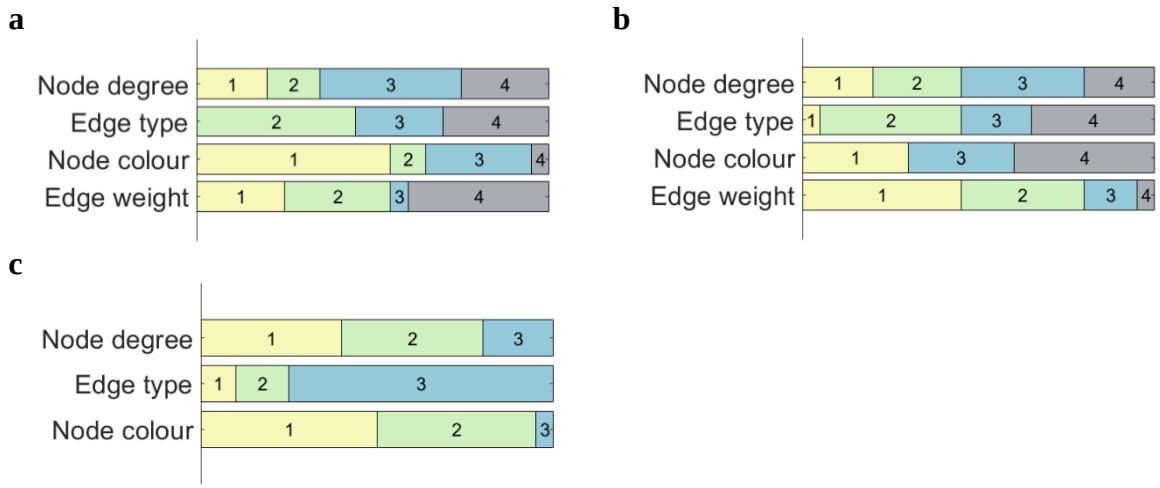


Figure 29: Rankings of the extent to which each characteristic/tools played a role in decision-making while identifying a root cause from a causality map. *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR*

However, when only the responses from participants with prior experience with causality analysis are considered (Figure 30), more emphasis is placed on edge type and edge weight than when all responses are included. This again touches on the importance of incorporating process knowledge in causality analysis – in this case to make the maps trustworthy by communicating whether a connection is validated by process knowledge. It should be noted that only six of the responses are from participants who had prior experience with causality analysis, which emphasizes the topic specialty of causality analysis.

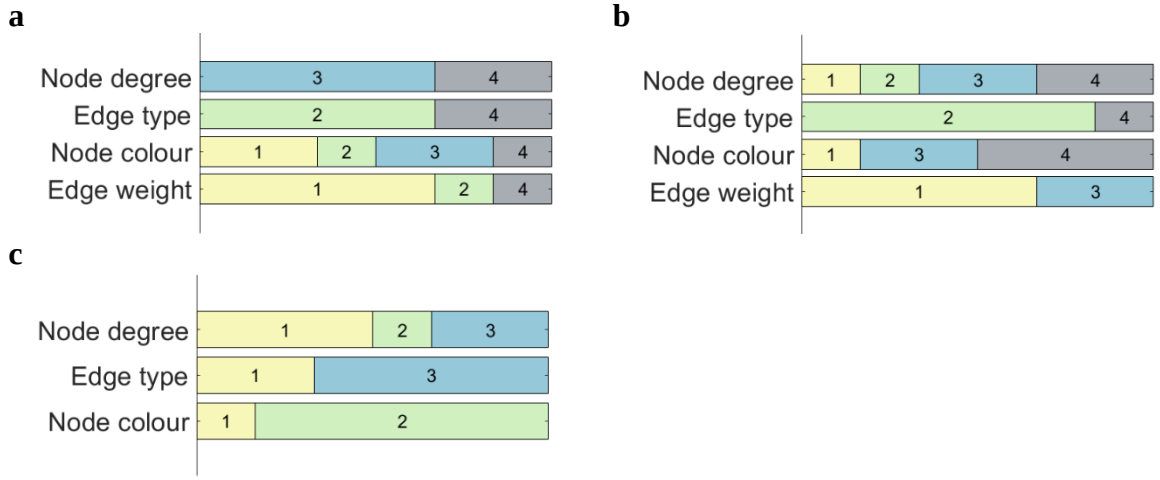


Figure 30: Rankings of the extent to which each characteristic/tools played a role in decision-making while identifying a root cause from a causality map **for participants with prior experience with causality analysis**. *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* **a.** PS-PC1. **b**. Mod-PC1. **c**. TR

The aim is to use these results to inform an importance-ranking for which characteristics/tools to look at when a causality map is used for fault identification. Since, there is no distinct trend for the role that these characteristics/tools play in the decision-making process that holds for all three approaches (PS-PC1, Mod-PC1, and TR in Figure 29 and/or Figure 30), the most value is placed on the results pertaining to PS-PC1, as this produced the most useful map (see Section 6.5). For this analysis, rankings of 1 and 2 are grouped together, and rankings of 3 and 4 are grouped together, to identify whether the characteristic/tool is used in the beginning of the

interpretation process or at the end of the interpretation process. Based on the aggregated result from all 20 responses for PS-PC1 (Figure 29a), the importance-ranking of the characteristics/tools is: (1) node colour, (2) edge weight, (3) edge type, (4) node degree. Based on the aggregated results from only the six responses from participants with prior experience with causality analysis for PS-PC1 (Figure 30a), the importance-ranking is: (1) edge weight, (2) edge type, (3) node colour, (4) node degree.

In both cases, node colour and edge weight are more important than node degree. This reinforces the idea that tools specifically aimed at aiding in causality map interpretation are useful. When no such tools are incorporated, node degree plays a large role in the interpretation process; but here it is clearly shown that when the tools to aid in interpretation are available, they are more important to the user than node degree. In addition, it is shown that participants with prior experience in causality analysis place value on process knowledge incorporated in causality maps. Plots that distinguish between university/company, level of education, and prior experience with causality analysis are available in Appendix H.

Sliders

Participants were also asked to indicate whether they used the sliders and which map (i.e. for which value of the slider(s)) they used to make their final decision w.r.t. the root cause. The results are presented in Figure 31, and they show that a significant majority of participants used the sliders. However, often-times, they end up on the default value(s) of the slider(s) again. This may indicate that the participants merely used the sliders because they were available, but that they were not user-friendly or particularly helpful. This is discussed further in the thematic analysis of the open-ended questions. Again, plots that distinguish between university/company, level of education, and prior experience with causality analysis are available in Appendix H.

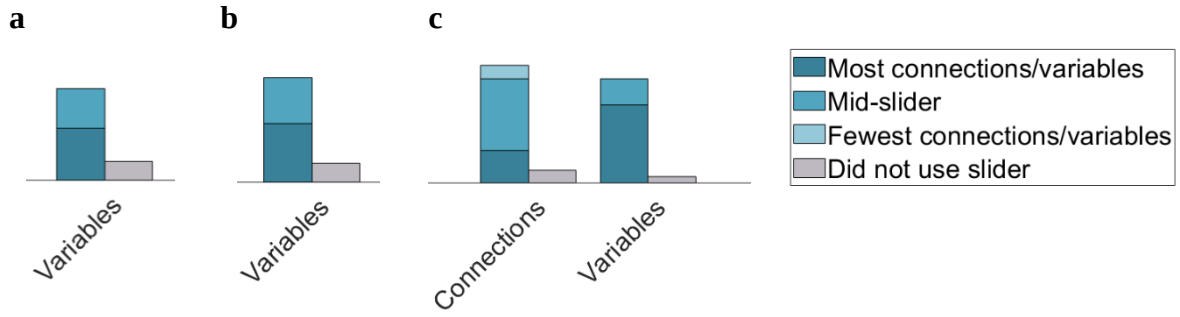


Figure 31: Comparison of the number of participants who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default.* **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

User-friendliness

Participants were asked to indicate the user-friendliness of the causality map in each approach (PS-PC1, Mod-PC1, TR), where 1 represents the least user-friendly and 4 represents the most user-friendly (Figure 32). The majority of participants indicated that all three approaches are user-friendly, as over half of the responses for each approach are a user-friendliness value on the top half of the scale (i.e. 3 or 4). This indicates that all three approaches show promise for use in industry, when looking from a purely interpretability point of view (i.e. with no consideration of accuracy). Furthermore, these results allow the approaches to be ranked from best to worst in terms of user-friendliness: (1) PS-PC1, (2) Mod-PC1, (3) TR; where it should also be noted that PS-PC1 is never given a 1 (i.e. least user-friendly).

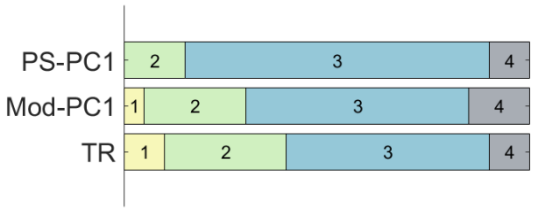


Figure 32: Comparison of the user-friendliness rankings assigned to the PS-PC1, Mod-PC1, and TR approaches, with 1 representing least user-friendly and 4 representing most user-friendly.

6.6.2. Thematic analysis of open-ended questions

The open-ended questions asked the participants to indicate and motivate whether the TR, PS-PC1, and Mod-PC1 could be practically applied in industry to provide actionable insights. In these open-ended questions, 13 (65 %) of the participants directly stated that TR could be applied in industry; 15 (75 %) stated that PS-PC1 could be applied, and 10 (50 %) stated that Mod-PC1 could be applied. Further insights were gathered from the participants' motivations, by grouping the data according to four themes: (1) Process knowledge and causality analysis experience; (2) grouping variables; (3) accuracy during simplification; and (4) sliders. This section presents the major findings in each of these themes, and Table 16 in Appendix H contains a summary of lines from participant responses grouped according to codes and subsequently the four themes.

Process knowledge and causality analysis experience

Process knowledge was once again highlighted as an important aspect of causality analysis, and specifically mentioned by 8 (40 %) of the participants. It was mentioned that it is important for process knowledge to be incorporated in the causality analysis procedure itself, with one participant stating that they would be "reserved to do this [Mod-PC1] as I believe a pure data driven causality map without any intrinsic process knowledge will lead to causal mistakes." In addition, the user(s) should also have their own process knowledge in addition to the causality map(s), with participants stating that, "having a background of the mineral processing plant further makes it easier to understand why a certain stage would be influenced by the previous stage as compared to just looking at the causality map alone," and that the user's process knowledge can serve as a "reality check". However, it should be noted that two (10 %) of the respondents mentioned a contrasting view, where a purely data-based approach (e.g. Mod-PC1) is preferable, because then "knowledge about the process will not bias the investigator's view."

Furthermore, causality analysis is a specialized technique and it was highlighted that users should be taught how to apply and interpret it effectively. Five (25 %) of the participants mentioned that, "some form of training / experience is needed on my part".

Grouping variables

In all responses regarding grouping variables, it is mentioned in a positive light, with four (25 %) of the respondents stating a variation of, "fewer variables allows for easier interpretation." Furthermore, the positive view of process knowledge is carried over to this theme by three (15 %) of the participants; stating that, "when grouping according to plant sections, a certain degree of plant knowledge is introduced," and, "grouping by plant layout just seemed more intuitive." This is complemented by the idea that grouping according to modules (Mod-PC1) adds confusion. One participant mentioned that, "the groupings did not quite make sense" in response to the Mod-PC1 approach. In addition, the data-based modules can end up

with numerous variables, with another participant mentioning that, "the sheer number of variables in certain modules can make identifying problem sections challenging."

Accuracy during simplification

Accuracy was highlighted as important for causality maps by three (15 %) of the participants. As a participant mentioned, "anything that simplifies the causality graphs is good (assuming it stays correct)." Concerns were raised regarding the use of only the first principal component to represent a group of variables, as, "the first PC captures bulk of variation - some important information contributing to a variation may be lost in subsequent PCs." However, using more than one PC per group of variables would introduce confusion by increasing the number of nodes in the causality map, as well as making it difficult to trace the root cause to a single group of variables when a single group of variables is represented by multiple nodes. In addition, a participant recommended that the uncertainty in the connection strengths should be quantified and communicated to aid in clarity of interpretation.

Sliders

Half of the participants directly mentioned that the sliders are confusing. The main reason is that the causality maps completely change according to the sliders, as a participant mentioned: "Interpretation of the maps changes with your sliders - seems clear on one and less on another." The confusion is amplified when there are two sliders (i.e. connections-slider and variables-slider), as, "there are too many things to consider as the two slider options create multiple permutations of possible setting. Further confusion was added in the Mod-PC1 approach, where a change in the variables-slider produces a different standard map with a different modular structure – so, "the changing of groups of variables, depending on the value selected in the slider, made this even more complex."

6.6.3. Summary of survey findings

The survey results are analysed to gain insight into which characteristics/tools are more useful in causality maps, and whether PS-PC1, Mod-PC1, or TR can be practically applied in industry. Regarding characteristics/tools to aid in causality analysis, node colour is found to be useful for cyclic maps in all three approaches (PS-PC1, Mod-PC1, TR), and node colour and edge weight are deemed as more important than node degree – emphasising the fact that tools specifically aimed at aiding in causality map interpretation are valued by users. Sliders are found to be confusing, as causality maps completely change according to the sliders. Regarding the three approaches for plant-wide causality maps (PS-PC1, Mod-PC1, TR), the majority of participants indicate that all three approaches are user-friendly, but PS-PC1 is found to be the most user-friendly, followed by Mod-PC1, and then TR.

Another main finding is that experienced causality analysis users place value on process knowledge incorporated in causality maps. This is further highlighted in the results from the open-ended questions, where it is mentioned that process knowledge should be incorporated in the causality map itself, and that the user should have their own process knowledge while interpreting a causality map. Furthermore, PS-PC1 is intuitive, as it incorporates process knowledge by grouping variables according to PSs.

In addition, it is important to maintain accuracy when simplifying a causality map, and using only the first PC to represent a group of variables may result in some important information to be lost. However, using more than one PC would increase the number of nodes in the causality map, and PS-PC1 did result in a useful and accurate plant-wide causality map in this study.

CHAPTER 7

Conclusions and recommendations

7.1. Conditional Granger causality for plant-wide fault identification

Recall that the first objective of this work, listed in Section 1.2, is concerned with the incorporation of process knowledge in causality maps. This was partially addressed by the development of a novel heuristic approach in Section 6.2, to select the sampling period (SP) and time window (TW) that provide the strongest connection strength for a known connection in conditional GC. However, it was found that the use of this heuristic approach does not sufficiently improve the use of conditional GC for plant-wide causality analysis, as the plant-wide causality map produced using this heuristic approach did not identify the true root cause. The main reason for this was postulated to be the variety of different residence times and time delays involved in plant-wide analysis.

In future studies, the heuristic approach developed in this work can be improved by removing emphasis from simply trying to capture true connections, and rather trying to capture true connections while also preventing spurious connections. Instead of comparing connection strength of a known connection in different scenarios, the connection strength ratio of a known spurious connection and a known true connection could be compared in the scenarios, where a smaller ratio is desired. The ratio should have the true connection as the denominator to prevent division by zero when the spurious connection is not identified by the causality analysis.

Three additional limitations to the application of standard conditional GC for plant-wide analysis were identified. Firstly, it was found that conditional GC requires large amounts of data where the fault is present, in the order of weeks and even months, for analysis with 15 variables – which are unlikely to be available in industry, and means that the fault needs to be present for weeks or months before it can be identified by standard conditional GC. This motivates the need to perform sequential conditional GC, as discussed further in Section 7.3. Secondly, it was noted that conditional GC can only be applied if the ratio between the fault oscillation period and the process dynamics is large enough to allow the process dynamics to be captured in the time frame by using a large enough SP without occurrence of the aliasing effect. Thirdly, it was found that AIC is not suitable for model order selection, as it does not penalise the number of parameters enough – but keeps selecting a larger model order if a larger TW is available, despite the time frame increasing past the true process dynamics.

GC could potentially be improved beyond its performance in this work by implementing alternatives to AIC for model order selection, such as Bayesian information criterion (BIC), cross-validation, or checking known connection strength at different model orders. However, GC remains an unsupervised approach with low interpretability in terms of the underlying

regression, where local optimal values are obtained for hyperparameters such as model order. This is typical of an engineering application, with the reason behind it being ease of automation. However, a more robust approach may be necessary, and a recommendation for future studies is to treat hyperparameter selection in GC or causality analysis in general as an optimisation problem. The question then becomes what the overall performance function would be, and since causality analysis is an explanatory method, there is no clear answer to this - yet.

7.2. Process knowledge, tools, and visualisation techniques for causality map interpretability

Recall that the first, second, and third objectives of this work, as listed in Section 1.2, are concerned with incorporating process knowledge in causality analysis; incorporating tools and visualisation techniques to aid in interpretation of plant-wide causality maps; and performing a usability study to gain insights on whether causality maps with the proposed tools and visualisation techniques can be applied in industry, respectively. The first objective was largely addressed by validating data-based connections in a causality map with process knowledge from a P&ID. The resulting causality map was displayed in two different ways: namely, excluding the non-validated connections completely, or indicating non-validated connection as uncertain by displaying them with dashed lines as opposed to the solid lines used for validated connections. Excluding non-validated connections decreased the graph density from 0.66 to 0.3, but simply identifying uncertain connections prevents any loss of information if the P&ID is outdated. Both versions were found to be an improvement over the standard causality map with no process knowledge, as all spurious connections were removed or identified as uncertain respectively. In addition, the survey in the usability study found that experienced causality analysis users place value on process knowledge in causality maps.

Process knowledge was also incorporated in Section 6.3.2 to decrease the pool of potential root cause variables, by displaying nodes that cannot reach all other nodes in the causality map in a faded grey. This was found to successfully exclude a node that was not the true root cause. Variable selection prior to causality analysis was advised, since this approach is only effective if all nodes in the map show an effect of the same fault.

The second objective was addressed by implementing tools such as node ranking and sliders that can be applied to a standard causality map. Firstly, node rankings determined according to the page rank algorithm were displayed visually on the causality map by colouring the nodes alongside a colour bar in Section 6.4.1. This was found to be especially helpful for cyclic maps, as node colour is then the main distinguishing factor between nodes. Furthermore, the survey in the usability study found that node colour and edge weight are important tools for causality map interpretation, more so than node degree. This provides substantial motivation for the use of tools specifically aimed at aiding in causality map interpretability, above simply looking at node degree for root cause identification.

Secondly, a connections-slider was implemented to vary the connection strength required for the connection to be displayed on the causality map in Section 6.4.2. A stricter threshold for connection strength decreased graph density, but also removed vital connections on the fault propagation path. However, that should not be the case in an accurate causality map. Two versions of a variables-slider based on the spectral envelope method were implemented in Section 6.4.3 to vary the variables included in the causality analysis. The first version was

based on how close the oscillation contribution (OC) of the variable is to the OC of an identified oscillating variable, and the second was based purely on the size of OC of the variable. Neither version of the variables-slider was found to be useful, as the first version depends heavily on which variable ends up as the identified oscillating variable, and it was found that the true root cause does not have the largest OC, so a stricter threshold removed it from the causality map. Furthermore, 50 % of the survey participants in the usability study directly stated that the sliders added confusion. Variable selection should therefore be performed using the spectral envelope method on a once-off basis prior to causality analysis. In this work, all the variables were found to oscillate significantly, but that is likely because exaggerated valve stiction was modelled in the case study, which will not be the case in industry. Variable selection using the spectral envelope method is useful, as it only involves one hyperparameter so it can be applied with confidence, and fewer variables requires a smaller dataset for conditional GC.

7.3. Hierarchical approach for plant-wide causality analysis

This work proposed a novel hierarchical approach for plant-wide causality analysis in Section 6.5, which further addresses the second objective listed in Section 1.2. This hierarchical approach consists of two stages: where a less-detailed, plant-wide map was constructed and used to localise the root cause to a specific group of variables in the first stage; and a causality map was constructed for only that group of variables and used to identify the root cause in the second stage. The plant-wide map was constructed in two different ways: (1) with variables grouped according to plant sections (PSs), and first principal components (PC1s) used to represent each group (PS-PC1); (2) with variables grouped according to modules identified in the standard causality map and PC1s used to represent each group (Mod-PC1).

Mod-PC1 is not recommended for plant-wide causality analysis, as it did not clearly identify the true root cause location in this study, and it requires the same amount of data as standard conditional GC – which means that the fault needs to be present for weeks or months before Mod-PC1 can identify it. However, PS-PC1 showed promising results, as its plant-wide causality map consisted of only three nodes and still clearly identified the true root cause location. This is an improvement above both the standard causality map and its transitive reduction, which both consisted of 15 variables and did not identify the true root cause. Furthermore, the PS-PC1 approach requires significantly fewer samples than both a standard plant-wide causality map and Mod-PC1, which motivates its application in industry where the root cause of a fault should be identified before a large amount of data with the fault present becomes available. A unique OLS solution for the full model in conditional GC was shown to require only 602 samples for the plant-wide map in PS-PC1, compared to the 13 562 samples required for the standard causality map. In addition, 75 % of the participants in the usability study survey directly stated that they believe the PS-PC1 approach can be practically applied in industry to provide actionable insights.

The application of PS-PC1 to an industrial dataset can therefore be investigated in future work. The main challenge is the large number of variables in industry, since one PC may not capture sufficient variance if a PS contains variables in the order of 100 instead of only the five variables per PS in this work. This can potentially be addressed by decreasing the number of variables represented by a single PC, by (1) using more than two stages in the hierarchical approach, or (2) including only specific types of variables (e.g. CVs, MVs, flowrates, levels) in the PCA in the first stage of the hierarchical approach.

References

Aho, A. V., Garey, M. R. and Ullman, J. D. (1972) 'The Transitive Reduction of a Directed Graph', *SIAM Journal on Computing*, 1(2), pp. 131–137. doi: 10.1137/0201008.

Akaike, H. (1974) 'A New Look at the Statistical Model Identification', *IEEE Transactions on Automatic Control*, 19(6), pp. 716–723. doi: 10.1093/ietfec/e90-a.12.2762.

Arroyo Esquivel, E. (2017) *Capturing and Exploiting Plant Topology and Process Information as a Basis to Support Engineering and Operational Activities in Process Plants*, *http://edoc.sub.uni-hamburg.de/hsu/volltexte/2017/3162/pdf/PhDDissertation_Arroyo_Final.pdf*. Helmut Schmidt University, Hamburg.

Barnett, L., Barret, A. B. and Seth, A. K. (2009) 'Granger causality and transfer entropy are equivalent for Gaussian variables', *Physical Review Letters*, 103(23), p. 238701.

Barnett, L. and Bossomaier, T. (2013) 'Transfer entropy as a log-likelihood ratio', *Physical Review Letters*, 109(13), p. 138105.

Barnett, L. and Seth, A. K. (2014) 'The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference', *Journal of Neuroscience Methods*. Elsevier B.V., 223, pp. 50–68. doi: 10.1016/j.jneumeth.2013.10.018.

Barnett, L. and Seth, A. K. (2017) 'Detectability of Granger causality for subsampled continuous-time neurophysiological processes', *Journal of Neuroscience Methods*. Elsevier B.V., 275, pp. 93–121. doi: 10.1016/j.jneumeth.2016.10.016.

Barrett, A. B., Barnett, L. and Seth, A. K. (2010) 'Multivariate Granger causality and generalized variance', *Physical Review E*, 81. doi: 10.1103/PhysRevE.81.041907.

Bauer, M., Cox, J. W., Caveness, M. H., Downs, J. J. and Thornhill, N. F. (2007) 'Finding the Direction of Disturbance Propagation in a Chemical Process Using Transfer Entropy', *IEEE Transactions on Control Systems Technology*, 15(1), pp. 12–21. doi: 10.1109/TCST.2006.883234.

Bauer, M. and Thornhill, N. F. (2008) 'A practical method for identifying the propagation path of plant-wide disturbances', *Journal of Process Control*, 18, pp. 707–719. doi: 10.1016/j.jprocont.2007.11.007.

Blondel, V. D., Guillaume, J. L., Lambiotte, R. and Lefebvre, E. (2008) 'Fast unfolding of communities in large networks', *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), p. P10008. doi: 10.1088/1742-5468/2008/10/P10008.

Blondel, V. D., Guillaume, J., Lambiotte, R. and Lefebvre, E. (2004) 'Fast unfolding of communities in large networks', *Journal of Statistical Mechanics: Theory and Experiment*.

Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3, pp. 77–101. doi: 10.1191/1478088706qp063oa.

Bressler, S. L. and Seth, A. K. (2011) 'Wiener-Granger Causality: A well established methodology', *NeuroImage*. Elsevier Inc., 58(2), pp. 323–329. doi: 10.1016/j.neuroimage.2010.02.059.

Bryan, K. and Leise, T. (2006) 'The $25,000,000,000 Eigenvector: The Linear Algebra behind Google', *SIAM Review*, 48(3), pp. 569–581. doi: 10.1137/050623280.

Bullmore, E. and Sporns, O. (2009) 'Complex brain networks: graph theoretical analysis of structural and functional systems.', *Nature reviews. Neuroscience*, 10, pp. 186–198. doi: 10.1038/nrn2575.

Choudhury, M. A. A. S., Thornhill, N. F. and Shah, S. L. (2005) 'Modelling valve stiction', *Control Engineering Practice*, 13, pp. 641–658. doi: 10.1016/j.conengprac.2004.05.005.

Duan, P. (2014) *Information Theory-based Approaches for Causality Analysis with Industrial Applications, Diss. University of Alberta*. University of Alberta.

Duan, P., Chen, T., Shah, S. L. and Yang, F. (2014) 'Methods for Root Cause Diagnosis of Plant-Wide Oscillations', *AIChE Journal*, 60(6), pp. 2019–2034. doi: 10.1002/aic.

Duan, P., Yang, F., Shah, S. L. and Chen, T. (2015) 'Transfer Zero-Entropy and Its Application for Capturing Cause and Effect Relationship Between Variables', *IEEE Transactions on Control Systems Technology*, 23(3), pp. 855–867. doi: 10.1109/TCST.2014.2345095.

Geweke, J. (1982) 'Measurement of linear dependence and feedback between multiple time series', *Journal of the American Statistical Association*, 77(378), pp. 304–313. doi: 10.1080/01621459.1982.10477803.

Geweke, J. F. (1984) 'Measures of Conditional Linear Dependence and Feedback Between Time Series', *Journal of the American Statistical Association*, 79(388), pp. 907–915.

Gibbons, A. A. M. (1985) *Algorithmic graph theory*. Cambridge, England: Cambridge, England : Cambridge University Press,.

Goebel, R., Roebroeck, A., Kim, D. S. and Formisano, E. (2003) 'Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping', *Magnetic Resonance Imaging*, 21, pp. 1251–1261. doi: 10.1016/j.mri.2003.08.026.

Granger, C. W. J. (1969) 'Investigating Causal Relations by Econometric Models and Cross-spectral Methods', *Econometrica*, 37(3), pp. 424–438.

Hill, R. (2011) *Principles of econometrics*. 4th edn. NJ: NJ: Wiley, Hoboken.

Hurvich, C. M. and Tsai, C. L. (1989) 'Regression and time series model selection in small samples', *Biometrika*, 76(2), pp. 297–307. doi: 10.1093/biomet/76.2.297.

Jenkins, G. M. and Watts, D. G. (1968) *Spectral Analysis and Its Applications*. Holden Day.

Jiang, H., Choudhury, M. A. A. S. and Shah, S. L. (2007) 'Detection and Diagnosis of Plant-wide Oscillations From Industrial Data using the Spectral Envelope Method', *Journal of Process Control*, 17(2), pp. 143–155.

Jiang, H., Patwardhan, R. and Shah, S. L. (2008) 'Root Cause Diagnosis of Plant-wide Oscillations using the Adjacency Matrix', in *The 17th World Conference, The International Federation of Automatic Control*. Seol, Korea, pp. 13893–13900.

Van den Kerkhof, P., Vanlaer, J., Gins, G. and Van Impe, J. F. M. (2013) 'Analysis of smearing-out in contribution plot based fault isolation for Statistical Process Control', *Chemical Engineering Science*, 104, pp. 285–293.

Kühnert, C. and Beyerer, J. (2014) 'Data-Driven Methods for the Detection of Causal Structures in Process Technology', *Machines*, 2(4), pp. 255–274. doi: 10.3390/machines2040255.

Landman, R. (2019) *Data-based causality analysis by exploiting process connectivity information*. Aalto University.

Landman, R. and Jämsä-Jounela, S. L. (2016) 'Hybrid approach to casual analysis on a complex industrial system based on transfer entropy in conjunction with process connectivity information', *Control Engineering Practice*, 53, pp. 14–23. doi: 10.1016/j.conengprac.2016.04.010.

Landman, R., Kortela, J., Sun, Q. and Jämsä-Jounela, S. L. (2014) 'Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity', *Computers and Chemical Engineering*, 71, pp. 446–456. doi: 10.1016/j.compchemeng.2014.09.017.

Leicht, E. A. and Newman, M. E. J. (2008) 'Community structure in directed networks', *Physical Review Letters*, 100(11).

Leicht, E. A. and Newman, M. E. J. (2008) 'Community structure in directed networks', *Physical Review Letters*, 100(11), p. 118703. doi: 10.1103/PhysRevLett.100.118703.

Lidner, B. S. (2014) *Exploiting Process Topology for Optimal Process Monitoring*. Stellenbosch University.

Lindner, B. and Auret, L. (2015) 'Application of data-based process topology and feature extraction for fault diagnosis of an industrial platinum group metals concentrator plant', *IFAC-PapersOnLine*, 28(17), pp. 102–107. doi: 10.1016/j.ifacol.2015.10.086.

Lindner, B., Auret, L. and Bauer, M. (2017) 'Investigating the Impact of Perturbations in Chemical Processes on Data-Based Causality Analysis. Part 1: Defining Desired Performance of Causality Analysis Techniques', *IFAC-PapersOnLine*. Elsevier B.V., 50(1), pp. 3269–3274. doi: 10.1016/j.ifacol.2017.08.463.

Lindner, B., Chioua, M., Groenewald, J. W. D., Auret, L. and Bauer, M. (2018) 'Diagnosis of Oscillations in an Industrial Mineral Process Using Transfer Entropy and Nonlinearity Index', *IFAC-PapersOnLine*. Elsevier B.V., 51(24), pp. 1409–1416. doi: 10.1016/j.ifacol.2018.09.539.

Lindner, B. S. (2019) *Improving the performance of causality analysis techniques for automated fault diagnosis in mineral processing plants*.

Lindner, B. S. and Auret, L. (2014) *Data-driven fault detection with process topology for fault identification, IFAC Proceedings Volumes*. IFAC. doi: 10.3182/20140824-6-ZA-1003.02139.

Marlin, T. E. (1995) 'Level and inventory control', in Clark, B. J. and Morriss, J. M. (eds) *Process Control: Designing Processes and Control Systems for Dynamic Performance*. McGraw-Hill Inc., pp. 581–603.

MATLAB (2018) *9.7.0.1190202 (R2019b)*. Natick, Massachusetts: The MathWorks Inc.

Maurya, M. R., Rengaswamy, R. and Venkatasubramanian, V. (2003) 'A systematic framework for the development and analysis of signed digraphs for chemical processes. 1. Algorithms and analysis', *Industrial and Engineering Chemistry Research*, 42(20), pp. 4789–4810. doi: 10.1021/ie020644a.

Miao, T. and Seborg, D. E. (1999) 'Automatic detection of excessively oscillatory feedback control loops', in *Proceedings of the 1999 IEEE International Conference on Control Applications*. Kohala Coast-Island of Hawai'i, Hawai'i, USA, pp. 359–364. doi: 10.1109/cca.1999.806659.

Miskin, J. J. (2016) *Control performance assessment for a high pressure leaching process by means of fault database creation and simulation, Stellenbosch University, South Africa*. Stellenbosch University.

Newman, M. E. J. (2003) 'The Structure and Function of Complex Networks', *SIAM Review*, 45(2), pp. 167–256. doi: 10.1137/s003614450342480.

Newman, M. E. J. (2004) 'Analysis of weighted networks', *Physical Review E - Statistical,*

nonlinear, and soft matter physics, 70(5), p. 056131. doi: 10.1103/PhysRevE.70.056131.

Newman, M. E. J. (2006) 'Modularity and community structure in networks', *Proc. Natl. Acad. Sci. USA*, 103, pp. 8577–8582.

Newman, M. E. J. and Girvan, M. (2004) 'Finding and evaluating community structure in networks', *Physical Review E*, 69.

Palla, G., Derényi, I., Farkas, I. and Vicsek, T. (2005) 'Uncovering the overlapping community structure of complex networks in nature and society', *Nature*, 435(9), pp. 814–818. doi: 10.1038/nature03607.

Pearl, J. (2009) 'Causal inference in statistics: An overview', *Statistics Surveys*, 3, pp. 96–146. doi: 10.1214/09-SS057.

Pollonini, L., Patidar, U., Situ, N., Rezaie, R., Papanicolaou, A. C. and Zouridakis, G. (2010) 'Functional connectivity networks in the autistic and healthy brain assessed using Granger causality', *32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, pp. 1730–1733. doi: 10.1109/IEMBS.2010.5626702.

Price, W. L. (1971) *Graphs and Networks: An Introduction*. Edited by K. B. Haley. London Butterworths.

Reijneveld, J. C., Ponten, S. C., Berendse, H. W. and Stam, C. J. (2007) 'The application of graph theoretical analysis to complex networks in the brain', *Clinical Neurophysiology*, 118, pp. 2317–2331. doi: 10.1016/j.clinph.2007.08.010.

Reis, M. S. and Gins, G. (2017) 'Industrial Process Monitoring in the Big Data/Industry 4.0 Era: From Detection, to Diagnosis, to Prognosis', *Processes*, 5(4), p. 35. doi: 10.3390/pr5030035.

Shiozaki, J., Matsuyama, H., O'Shima, E. and Iri, M. (1979) 'An algorithm for diagnosis of system failures in the chemical process', *Computers and Chemical Engineering*, 3(14), pp. 489–493. doi: 10.1016/0098-1354(85)80006-5.

Shiozaki, J., Matsuyama, H., O'Shima, E. and Iri, M. (1985) 'An Improved Algorithm for Diagnosis of System Failures in the Chemical Process', *Compuers and Chemical Engineering*, 9(3), pp. 285–293.

Shu, Y. and Zhao, J. (2013) 'Data-driven causal inference based on a modified transfer entropy', *Computers and Chemical Engineering*. Elsevier Ltd, 57, pp. 173–180. doi: 10.1016/j.compchemeng.2013.05.011.

Streicher, S. J., Wilken, S. E. and Sandrock, C. (2014) 'Eigenvector Analysis for the Ranking of Control Loop Importance', *Computer Aided Chemical Engineering*. Elsevier, 33, pp. 835–840.

Sugiyama, K., Tagawa, S. and Toda, M. (1981) 'Methods for Visual Understanding of Hierarchical System Structures', *IEEE Transactions on Systems, Man and Cybernetics*, 11(2), pp. 109–125. doi: 10.1109/TSMC.1981.4308636.

Suresh, R., Sivaram, A. and Venkatasubramanian, V. (2019) 'A hierarchical approach for causal modeling of process systems', *Computers and Chemical Engineering*. Elsevier Ltd, 123, pp. 170–183. doi: 10.1016/j.compchemeng.2018.12.017.

Tarjan, R. (1972) 'Depth-First Search and Linear Graph Algorithms', *IAM Journal on Computing*, 1(2), p. 15.

Thambirajah, J., Benabbas, L., Bauer, M. and Thornhill, N. F. (2009) 'Cause-and-effect analysis in chemical processes utilizing XML, plant connectivity and quantitative process history', *Computers and Chemical Engineering*, 33, pp. 503–512. doi: 10.1016/j.compchemeng.2008.10.002.

Thornhill, N. F. (2005) 'Finding the Source of Nonlinearity in a Process With Plant-Wide Oscillation', *IEEE Transactions on Control Systems Technology*, 13(3), pp. 434–443.

Wakefield, B. J., Lindner, B. S., McCoy, J. T. and Auret, L. (2018) 'Monitoring of a simulated milling circuit: Fault diagnosis and economic impact', *Minerals Engineering*, 120, pp. 132–151. doi: 10.1016/j.mineng.2018.02.007.

Wiener, N. (1956) 'The theory of prediction', *Modern Mathematics for Engineers*, 1, pp. 125–139.

Yang, F., Duan, P., Shah, S. L. and Chen, T. (2014) *Capturing connectivity and causality in complex industrial processes*. 1st edn. Springer Internationl Publishing. doi: 10.1007/978-3-319-05380-6.

Yang, F., Sirish, L. S. and Xiao, D. (2010) 'Signed Directed Graph modeling of industrial processes and their validation by data-based methods', in *2010 Conference on Control and Fault-Tolerant Systems (SysTol)*, pp. 387–392.

Yang, F. and Xiao, D. (2012) 'Progress in Root Cause and Fault Propagation Analysis of Large-Scale Industrial Processes', *Journal of Control Science and Engineering*, 2012, pp. 1–10. doi: 10.1155/2012/478373.

Yim, S. Y., Ananthakumar, H. G., Benabbas, L., Horch, A., Drath, R. and Thornhill, N. F. (2006) 'Using process topology in plant-wide control loop performance assessment', *Computers & Chemical Engineering*, 31(2), pp. 86–99.

Yuan, T. and Qin, S. J. (2014) 'Root cause diagnosis of plant-wide oscillations using Granger causality', *Journal of Process Control*. Elsevier Ltd, 24(2), pp. 450–459. doi: 10.1016/j.jprocont.2013.11.009.

Zhou, D., Xiao, Y., Zhang, Y., Xu, Z. and Cai, D. (2014) 'Granger causality network reconstruction of conductance-based integrate-and-fire neuronal systems', *PLOS ONE*, 9(2). doi: 10.1371/journal.pone.0087636.

Zhou, Z., Ding, M., Chen, Y., Wright, P., Lu, Z. and Liu, Y. (2009) 'Detecting directional influence in fMRI connectivity analysis using PCA based Granger causality', *Brain Research*. Elsevier B.V., 1289, pp. 22–29. doi: 10.1016/j.brainres.2009.06.096.

APPENDIX A

Partial derivation of the spectral optimisation method

A.1. Partial derivation of spectral optimisation method – Part 1

This section provides a step-by-step derivation from eqn 31 to eqn 32 in Section 3.4:

$$Q = \frac{1}{2m}\sum_i\sum_j \left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right](s_i s_j + 1) \quad [31]$$

$$Q = \frac{1}{2m}\sum_i\sum_j \left\{\left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right]s_i s_j + \left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right]\right\} \quad [A1]$$

$$Q = \frac{1}{2m}\sum_i\sum_j \left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right]s_i s_j + \frac{1}{2m}\sum_i\sum_j A_{ij} - \frac{1}{2m}\sum_i\sum_j \frac{k_i^{in}k_j^{out}}{m} \quad [A2]$$

$$Q = \frac{1}{2m}\sum_i\sum_j \left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right]s_i s_j + \frac{1}{2m}\sum_i\sum_j A_{ij} - \frac{1}{2m}\sum_i \frac{k_i^{in}}{m}\sum_j k_j^{out} \quad [A3]$$

Using the fact that:

$$\sum_i\sum_j A_{ij} = \sum_i k_i^{in} = \sum_j k_j^{out} = m \quad [A4]$$

Eqn A3 is rewritten as:

$$Q = \frac{1}{2m}\sum_i\sum_j \left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right]s_i s_j + \frac{1}{2m}\sum_i\sum_j A_{ij} - \frac{1}{2m}\sum_j k_j^{out} \quad [A5]$$

$$Q = \frac{1}{2m}\sum_i\sum_j \left[A_{ij} - \frac{k_i^{in}k_j^{out}}{m}\right]s_i s_j \quad [A6]$$

Remember that the modularity matrix $\boldsymbol{B}$ is defined to have entries:

$$\beta_{ij} = A_{ij} - \frac{k_i^{in}k_j^{out}}{m} \quad [A7]$$

Then eqn A6 is rewritten as:

$$Q = \frac{1}{2m}\sum_i\sum_j s_i \beta_{ij} s_j \quad [32]$$

A.2. Partial derivation of spectral optimisation method – Part 2

This section provides a step-by-step derivation from eqn [37] to eqn [38] in Section 3.4:

$$Q^* = \sum_i a_i \boldsymbol{v}_i^T (\boldsymbol{B} + \boldsymbol{B}^T) \sum_j a_j \boldsymbol{v}_j^T \quad [37]$$

Let $\boldsymbol{C} = \boldsymbol{B} + \boldsymbol{B}^T$, so that eqn [37] is rewritten as:

$$Q^* = \sum_i a_i \boldsymbol{v}_i^T \boldsymbol{C} \sum_j a_j \boldsymbol{v}_j^T \quad [A8]$$

Using the fact that $\boldsymbol{C}$ is symmetric (i.e. $\boldsymbol{C} = \boldsymbol{C}^T$), and remembering that $\boldsymbol{s} = \sum_i a_i \boldsymbol{v}_i$, eqn A8 can be rewritten as:

$$Q^* = \sum_i a_i (\boldsymbol{C}\boldsymbol{v}_i)^T \boldsymbol{s} \quad [A9]$$

Using eigenvalue decomposition, $\boldsymbol{C}\boldsymbol{v}_i = \gamma_i \boldsymbol{v}_i$, where γ_i is the eigenvalue that corresponds to the eigenvector $\boldsymbol{v}_i$. Eqn [A9] is then rewritten as:

$$Q^* = \sum_i a_i (\gamma_i \boldsymbol{v}_i)^T \boldsymbol{s} \quad [A10]$$

Since $\boldsymbol{s} = \sum a_i \boldsymbol{v_i}$ (see eqn 37 in Section 3.5), and all the eigenvectors of a symmetric matrix are orthogonal (i.e. $\boldsymbol{v}_i^T \boldsymbol{v}_j = \boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$), it follows that:

$$\boldsymbol{v_j^T s} = \sum a_i \boldsymbol{v_j^T v_i} = a_j \quad [A11]$$

Substituting eqn A11 into eqn A10:

$$Q^* = \sum_i \boldsymbol{v_i^T s} (\gamma_i \boldsymbol{v}_i)^T \boldsymbol{s} \quad [A12]$$

$$Q^* = \sum_i \gamma_i (\boldsymbol{v_i^T s})^2 \quad [38]$$

APPENDIX B

Author's code

This appendix presents the author's code for conditional GC (which is validated against the GCCA mode in the MVGC toolbox), as well as the author's code for validating data-based causal connections with a connectivity matrix, for both a standard causality matrix and a causality matrix where nodes represent groups of variables.

B.1. Conditional Granger causality

The author's conditional GC code is presented below, and the raw results for the causality matrices (containing the G-causality indices) produced by the author's code and the GCCA mode in the MVGC toolbox (Barnett and Seth, 2014) are presented in Table 14 and Table 15 respectively. The causality matrix produced by the author's code is identical to the one produced by the toolbox, and both the author's code and the toolbox selected a model order of 5 using AIC; thus the author's code is validated.

```
%% Conditional Granger causality

%% PURPOSE
% Performs conditional GC on a dataset X, where the rows are variables and
% the columns are observations.
% Selects model order using AICc, and checks that there are more
% observations than parameters for the VAR.

%% INPUTS
% X - data to perform GC on, where the rows are variables and the columns
% are observations
% pmax - maximum model order to calculate AIC for in model order selection

%% OUTPUTS
% cm - matrix containing magnitudes of G-causality between variables,
% where the columns are source variables and rows are sink variables
% p - model order used for regression

%%
function [cm, p] = Conditional_GC_updated(X, pmax)
X = X';
X = zscore(X);
X = X';
[n, m] = size(X); % n = no. variables; m = total no. observations
% Pre-allocate space for output:
cm = NaN(n);
% Determine model order using AICc:
AIC = NaN(pmax, 1);
% C_per_model_order = cell(n, pmax);
for q = 1:pmax
    % Full regression:
    X0 = X(:, q+1:end);
```

APPENDIX B

```
    XL = NaN(q*n, m-q);
    for i_XL = 0:(q-1)
        XL(i_XL*n+1:(i_XL+1)*n, :) = X(:, (q-i_XL):(m-1-i_XL)); % Lagged
                                                                % observations
    end
    A = X0/XL; % Solve system of equations with OLS
    E = X0 - A*XL;
    SIG = E*E'/(m-q-1); % Covariance matrix (denominator is to correct for
                                                                    % bias)
    DSIG = det(SIG); % Determinant of covariance matrix (generalised
                                                                % variance)
    % AIC:
    L = -(m/2)*log(DSIG);
    k = q*n*n;
    if m-k-1 <= 0 % Are there sufficient observations for a unique solution
                                                        % to the regression?
        AIC(q) = NaN;
    else
        AIC(q) = -2*L + 2*k*(m/(m-k-1)); % Note AIC without correction =
                                                          % -2*L + 2*k
    end
end
p = find( AIC==min(AIC) ); % Choose model order with smallest value for AIC
% Full regression:
X0 = X(:, p+1:end);
XL = NaN(p*n, m-p);
for i_XL = 0:(p-1)
    XL(i_XL*n+1:(i_XL+1)*n, :) = X(:, (p-i_XL):(m-1-i_XL)); % Lagged
                                                            % observations
end
A = X0/XL; % Solve system of epuations with OLS
E = X0 - A*XL;
SIG = E*E'/(m-p-1); % Covariance matrix (denominator is to corrected for
                                                                % bias)
VAR = diag(SIG); % Vector of variances (because variances are on diagonal
                                                    % of covariance matrix)
% Reduced regression & G-causality index:
for i = 1:n
    Xr = X;
    Xr(i, :) = []; % Remove potential source variable

    X0r = Xr(:, p+1:end);
    XLr = NaN(p*(n-1), m-p);
    for i_XL = 0:(p-1)
        XLr(i_XL*(n-1)+1:(i_XL+1)*(n-1), :) = Xr(:, (p-i_XL):(m-1-i_XL));
                                                     % Lagged observations
    end
    Ar = X0r/XLr;
    Er = X0r-Ar*XLr;
    SIGr = Er*Er'/(m-p-1);
    VARr = diag(SIGr);
    iii = 1;
    for ii = 1:n
        if ii == i
            cm(ii, i) = 0;
        else
            cm(ii, i) = log(VARr(iii)/VAR(ii));
            iii = iii +1;
        end
    end
end
end
```

APPENDIX B

Table 14: Raw results for the causality matrix produced using the author's conditional GC. *Columns are source variables and rows are sink variables.*

0	0.07474	0.145231	0.081042	0.074155	0.008497	0.008675	0.016913	0.018496	0.004254	0.092569	0.07528	0.140678	0.091651	0.07209
0.05200565	0	0.101657	0.051835	0.063769	0.012408	0.009339	0.01696	0.022674	0.002082	0.070479	0.045097	0.093793	0.0645	0.049023
0.05087999	0.060952	0	0.049902	0.057648	0.004996	0.051719	0.003486	0.006337	0.016119	0.053268	0.071549	0.081884	0.054408	0.077813
0.08951239	0.079529	0.164366	0	0.077384	0.009968	0.009484	0.022875	0.02254	0.003575	0.099358	0.081524	0.159715	0.095833	0.077187
0.1145506	0.103883	0.240022	0.116469	0	0.026022	0.017611	0.047952	0.049588	0.002154	0.140501	0.09968	0.226192	0.121791	0.098719
0.00371087	0.003469	0.002058	0.004012	0.003869	0	0.011964	0.097606	0.02911	0.010102	0.001541	0.002741	0.002154	0.001615	0.004276
0.00996687	0.011817	0.012896	0.010109	0.011771	0.032017	0	0.027135	0.034623	0.00239	0.009274	0.012254	0.008229	0.008003	0.012406
0.00311163	0.002027	0.001661	0.004557	0.001475	0.019352	0.011194	0	0.019861	0.012392	0.00214	0.002148	0.001696	0.003236	0.001616
0.00319376	0.003794	0.003478	0.005468	0.003029	0.023016	0.016468	0.152793	0	0.018891	0.003809	0.006	0.003487	0.004653	0.005506
0.00057672	0.001001	0.001146	0.000362	0.000586	0.024761	0.012033	0.069402	0.020624	0	0.000479	0.000794	0.001292	0.000448	0.000551
0.00486078	0.006388	0.006288	0.004718	0.008828	0.014151	0.013464	0.019563	0.023577	0.001012	0	0.007132	0.006935	0.013935	0.003851
0.00317511	0.018906	0.004122	0.003656	0.036671	0.019541	0.0129	0.029993	0.035442	0.000511	0.030465	0	0.004543	0.015624	0.020721
0.00156135	0.015278	0.000991	0.005609	0.016879	0.0215	0.099895	0.010575	0.022242	0.020362	0.021791	0.060404	0	0.027339	0.099419
0.00882078	0.009426	0.011347	0.008513	0.010902	0.01657	0.014301	0.027554	0.029587	0.001706	0.021547	0.011206	0.012196	0	0.006717
0.01452752	0.023947	0.018654	0.015064	0.035664	0.031521	0.018071	0.056915	0.05764	0.002353	0.041275	0.015602	0.02048	0.015509	0

Table 15: Raw results for the causality matrix produced using the GCCA mode in the MVGC toolbox *(Barnett and Seth, 2014)*. *Columns are source variables and rows are sink variables.*

NaN	0.07474	0.145231	0.081042	0.074155	0.008497	0.008675	0.016913	0.018496	0.004254	0.092569	0.07528	0.140678	0.091651	0.07209
0.052006	NaN	0.101657	0.051835	0.063769	0.012408	0.009339	0.01696	0.022674	0.002082	0.070479	0.045097	0.093793	0.0645	0.049023
0.05088	0.060952	NaN	0.049902	0.057648	0.004996	0.051719	0.003486	0.006337	0.016119	0.053268	0.071549	0.081884	0.054408	0.077813
0.089512	0.079529	0.164366	NaN	0.077384	0.009968	0.009484	0.022875	0.02254	0.003575	0.099358	0.081524	0.159715	0.095833	0.077187
0.114551	0.103883	0.240022	0.116469	NaN	0.026022	0.017611	0.047952	0.049588	0.002154	0.140501	0.09968	0.226192	0.121791	0.098719
0.003711	0.003469	0.002058	0.004012	0.003869	NaN	0.011964	0.097606	0.02911	0.010102	0.001541	0.002741	0.002154	0.001615	0.004276
0.009967	0.011817	0.012896	0.010109	0.011771	0.032017	NaN	0.027135	0.034623	0.00239	0.009274	0.012254	0.008229	0.008003	0.012406
0.003112	0.002027	0.001661	0.004557	0.001475	0.019352	0.011194	NaN	0.019861	0.012392	0.00214	0.002148	0.001696	0.003236	0.001616
0.003194	0.003794	0.003478	0.005468	0.003029	0.023016	0.016468	0.152793	NaN	0.018891	0.003809	0.006	0.003487	0.004653	0.005506
0.000577	0.001001	0.001146	0.000362	0.000586	0.024761	0.012033	0.069402	0.020624	NaN	0.000479	0.000794	0.001292	0.000448	0.000551
0.004861	0.006388	0.006288	0.004718	0.008828	0.014151	0.013464	0.019563	0.023577	0.001012	NaN	0.007132	0.006935	0.013935	0.003851
0.003175	0.018906	0.004122	0.003656	0.036671	0.019541	0.0129	0.029993	0.035442	0.000511	0.030465	NaN	0.004543	0.015624	0.020721
0.001561	0.015278	0.000991	0.005609	0.016879	0.0215	0.099895	0.010575	0.022242	0.020362	0.021791	0.060404	NaN	0.027339	0.099419
0.008821	0.009426	0.011347	0.008513	0.010902	0.01657	0.014301	0.027554	0.029587	0.001706	0.021547	0.011206	0.012196	NaN	0.006717
0.014528	0.023947	0.018654	0.015064	0.035664	0.031521	0.018071	0.056915	0.05764	0.002353	0.041275	0.015602	0.02048	0.015509	NaN

B.2. Validating data-based causality with a connectivity matrix

B.2.1. Standard causality map

The author's code for validating data-based causal connections with a connectivity matrix for a standard causality matrix is presented below:

```
%% Validate causal connections with connectivity matrix - for a STANDARD
% CAUSALITY MATRIX

%% PURPOSE
% Checks whether the causal connections in the causality matrix are
% validated by the connectivity matrix, and returns a refined causality
% matrix, containing only the validated connections.

%% INPUTS
% causalM_original - causality matrix, where rows are source variables and
% columns are sink variables
% connectM - connectivity matrix, where rows are source variables and
% columns are sink variables
% IDs_causal - Sensor tags that correspond to the measured variables in the
% causality matrix
% IDs_connect - Tag names corresponding to the connectivity matrix entries

%% OUTPUTS:
% causalM_refined - refined causality matrix, containing only connections
% validated by the connectivity matrix

%%
function causalM_refined = validate_propPath(causalM_original, connectM,...
    IDs_causal, IDs_connect)
    [causal_rows, causal_cols] = size(causalM_original);
    causalM_refined = zeros(causal_rows, causal_cols);
    % Step through causal matrix to check each connection:
    for i_causal_row = 1:causal_rows
        for i_causal_col = 1:causal_cols
            if i_causal_row == i_causal_col % exclude diagnoal to omit self
                    % loops (self loops are terrible for interpretability)
              causalM_refined(i_causal_row, i_causal_col) = 0;
            else
                % ***
                % Find indices of entries where there are causal
                % connections:
                if causalM_original(i_causal_row, i_causal_col) ~= 0
                    % ***
                    ID_sn = IDs_causal(i_causal_row); % sn = start node for
                    % DFS (depth-first search); sn gets overwritten for
                    % each causal connection that is checked
                    idx = find(strcmp(IDs_connect, ID_sn));
                    % ***

                    % What is the instrument connected to? (Check row & col
                    % corresponding to this instrument for 1's):
                    idx_row = find(connectM(idx, :)==1); % Row containing
                                                % indices where there is a 1
                    idx_col = find(connectM(:, idx)==1); % Col containing
                                               % indicies where there is a 1

                    % Save indices as start nodes (doesn't matter if they
                    % were row or col indices, because variables are in the
```

```
                    % same order in the rows & cols, so can use the
                    % obtained idices directly as row indices for the start
                    % nodes):
                    snVect = zeros(size(idx_row, 2)+size(idx_col, 1), 1);
                                                % snVect is a col vector
                    snVect(1:size(idx_row, 2)) = idx_row;
                    snVect(size(idx_row, 2)+1:size(idx_row, 2)+...
                        size(idx_col, 1)) = idx_col;

                    % DFS:
                    % Create digraph out of connectivity matrix
                    connectM(isnan(connectM))=0; % Replace all NaN values
                    % with zero (because the DFS sees NaN values as entries
                    % & therefore connections in the conectivity matrix
                    % (which would be wrong)
                    connectM_dg = digraph(connectM, IDs_connect);
                                                        % dg = digraph

                    % Pre-allocate space to store the paths
                    connectM_paths = cell(1, size(snVect, 1));

                    % DFS to get all nodes that can be reached from the
                    % specified start nodes
                    for i_DFS = 1:size(snVect, 1)
                       connectM_paths{1, i_DFS} = dfsearch(connectM_dg,...
                           IDs_connect{1, snVect(i_DFS)}); % 1st entry
                       % contains nodes reachable from 1st specified start
                       % node, and so on..
                    end

                    % Is the target node in the list of nodes that can be
                    % reached from the start nodes? If yes, causal
                    % connection is validated. If no, causal connection is
                    % deemed spurious and removed.
                    for i_targetReachable = 1:size(connectM_paths, 2)
                        a = strcmp(IDs_causal(i_causal_col), ...
                            connectM_paths{1, i_targetReachable});
                        if sum(a) > 0
                            causalM_refined(i_causal_row, i_causal_col)...
                                = causalM_original(i_causal_row,...
                                i_causal_col);
                            break;
                        else
                            causalM_refined(i_causal_row, i_causal_col)...
                                = 0;
                        end
                    end
                end
            end
        end
    end
end
```

B.2.2. Causality map with nodes representing groups of variables

The author's code for validating data-based causal connections with a connectivity matrix for a causality matrix where nodes represent groups of variables is presented below.

```
%% Validate causal connections with connectivity matrix - for a CAUSALITY
% MATRIX WHERE NODES REPRESENT GROUPS OF VARIABLES

%% PURPOSE
% Checks whether the causal connections in the causality matrix are
```

```
% validated by the connectivity matrix, and returns a refined causality
% matrix, containing only the validated connections.

%% INPUTS
% causalM_original - causality matrix, where rows are source variables and
% columns are sink variables
% connectM - connectivity matrix, where rows are source variables and
% columns are sink variables
% IDs_causal - Sensor tags that correspond to the measured variables in the
% causality matrix
% IDs_connect - Tag names corresponding to the connectivity matrix entries

%% OUTPUTS:
% causalM_refined - refined causality matrix, containing only connections
% validated by the connectivity matrix

%%
function causalM_refined = validate_propPath(causalM_original, connectM,...
    all_IDs_causal, IDs_connect)
    [numGroups, ~] = size(causalM_original);
    numVars = NaN(1, numGroups); % Row vector that contains the no. of vars
                                                          % in each group
    for i = 1:numGroups
        [~, numVars(i)] = size(all_IDs_causal{1, i}); % Okay for now,
        % because there are an equal number of variables in each section
    end
    causalM_refined = zeros(numGroups);
    % Step through causal matrix to check each connection:
    for i_sourceSection = 1:numGroups
        for i_sinkSection = 1:numGroups
            if i_sourceSection == i_sinkSection % exclude diagnoal to omit
                % self loops (self loops are terrible for interpretability)
              causalM_refined(i_sourceSection, i_sinkSection) = 0;
            else
                % ***
                % Find indices of entries where there are causal
                % connections:
                if causalM_original(i_sourceSection, i_sinkSection) ~= 0

                    % ***
                    % Instead of stepping through all the variables in the
                    % source section, just include all those variables in
                    % the list of start nodes. sn has to be a scalar, so
                    % will need to step throuhg all the variables after
                    % all:
                    for i_sourceSectionVars = 1:numVars(i_sourceSection)
                    ID_sn = all_IDs_causal{1, i_sourceSection}...
                        {i_sourceSectionVars}; % numVars must equal the
                    % number of variables in the current source section;
                    % generic for now, because all the sections have an
                    % equal number of variables
                    idx = find(strcmp(IDs_connect, ID_sn));
                    % ***

                    % What is the instrument connected to? (Check row & col
                    % corresponding to this instrument for 1's):
                    idx_row = find(connectM(idx, :)==1); % Row containing
                                                  % indices where there is a 1
                    idx_col = find(connectM(:, idx)==1); % Col containing
                                                 % indicies where there is a 1

                    % Save indices as start nodes (doesn't matter if they
```

```
                    % were row or col indices, because variables are in the
                    % same order in the rows & cols, so can use the
                    % obtained idices directly as row indices for the start
                    % nodes):
                    snVect = zeros(size(idx_row, 2)+size(idx_col, 1), 1);
                                                % snVect is a col vector
                    snVect(1:size(idx_row, 2)) = idx_row;
                    snVect(size(idx_row, 2)+1:size(idx_row, 2)+...
                        size(idx_col, 1)) = idx_col;

                    % DFS:
                    % Create digraph out of connectivity matrix
                    connectM_dg = digraph(connectM, IDs_connect);
                                                        % dg = digraph

                    % Pre-allocate space to store the paths
                    connectM_paths = cell(1, size(snVect, 1));

                    % DFS to get all nodes that can be reached from the
                    % specified start nodes
                    for i_DFS = 1:size(snVect, 1)
                       connectM_paths{1, i_DFS} = dfsearch(connectM_dg,...
                           IDs_connect{1, snVect(i_DFS)}); % 1st entry
                       % contains nodes reachable from 1st specified start
                       % node, and so on..
                    end

                    % Is the target node in the list of nodes that can be
                    % reached from the start nodes? If yes, causal
                    % connection is validated. If no, causal connection is
                    % deemed spurious and removed.

                    % IDs_causal(i_causal_col) is the target node -
                    % Make it all nodes in the sink section

                    % The for loop for the variables in the sink section
                    % must be here:
                    b = 0;
                    for i_sinkSectionVars = 1:numVars(i_sinkSection)
                        for i_targetReachable = 1:size(connectM_paths, 2)
                            a = strcmp(all_IDs_causal{1, i_sinkSection}...
                                (i_sinkSectionVars), connectM_paths{1,...
                                i_targetReachable});
                            b = b+sum(a);
                            if b > 0
                              causalM_refined(i_sourceSection,...
                                  i_sinkSection) = causalM_original...
                                  (i_sourceSection, i_sinkSection);
                              break;
                            end
                        end
                    end
                    end
                end
            end
        end
    end
end
```

APPENDIX C

Tank network model development and implementation

The tank network model was developed by performing a mass balance around a tank, and extending that to all the tanks in the network. The symbols used in the model development correspond to those in Figure 33.

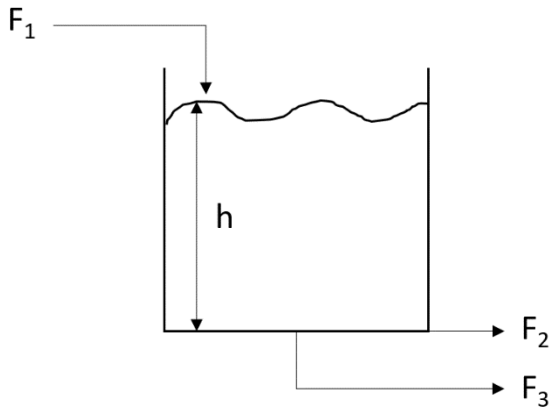


Figure 33: Diagram with symbols corresponding to one-tank model development

$$\frac{dm}{dt} = \rho(F_1 - F_2 - F_3) \quad [C1]$$

Assume the fluid is incompressible, so the density is constant.

$$\frac{dV}{dt} = F_1 - F_2 - F_3 \quad [C2]$$

$$\frac{d(Ah)}{dt} = F_1 - F_2 - F_3 \quad [C3]$$

Since the cross-sectional area A remains constant:

$$\frac{dh}{dt} = \frac{F_1 - F_2 - F_3}{A} \quad [C4]$$

Divide by the maximum height possible in the tank H_{max} to obtain the tank height as a fraction of its total:

$$\frac{d\left(\frac{h}{H_{max}}\right)}{dt} = \frac{F_1 - F_2 - F_3}{AH_{max}} \quad [C5]$$

$$\frac{dh_{frac}}{dt} = \frac{F_1 - F_2 - F_3}{V_{max}} \quad [C6]$$

Since F_2 is the MV, it is calculated using the PI controller equation in the Laplace domain, where K_c and τ_I are determined from the controller tuning relations derived in Appendix E.

$$F_2 = K_c\left[1 + \frac{1}{s\tau_I}\right] \quad [C7]$$

F_3 is the underflow, which is proportional to the squareroot of the tank level:

$$F_3 = k\sqrt{h_{frac}} \quad [C8]$$

Eqns C6, C7, and C8 were implemented in Simulink, as shown in Figure 34.

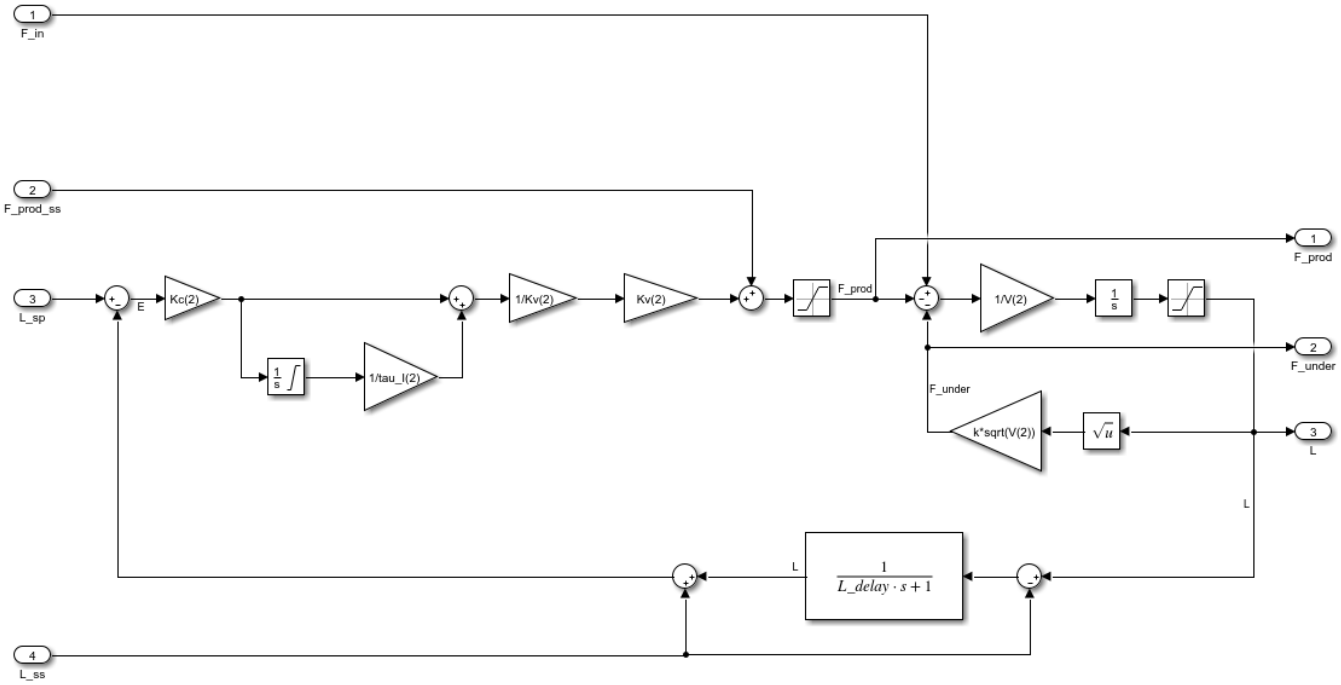


Figure 34: Simulink implementation of the one-tank model

The exogenous disturbances fed to each plant section were modelled as random walks by integrating a series of random numbers multiplied by a gain, as shown in Figure 35.

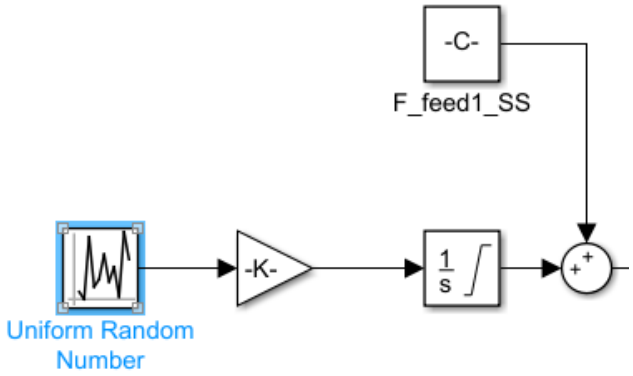


Figure 35: Random walk implementation in Simulink

APPENDIX D

Modelling valve stiction

Valve stiction can be described as the resistance exhibited by a valve to start moving as a result of large static friction (Miskin, 2016). It consists of four components, namely: deadband, stickband, slipjump and the moving phase, as shown in Figure 36. The deadband and stickband represent the behaviour of the valve when the controller wants it to move, but it is not moving. The slipjump represents the abrupt jump in valve position that occurs when the valve overcomes the static friction that had been keeping it stuck.

For this project, valve stiction was modelled using the empirical model described in Choudhury, Thornhill and Shah, (2005). The two parameters that govern this model are J and S, which represent the slipjump and deadband + stickband respectively. S was set as 60 %, so that the valve takes long to respond to the controller output and so causes a low frequency oscillation. J was set as 65 %, so that J > S, resulting in overshoot stiction to provide a larger fault amplitude.

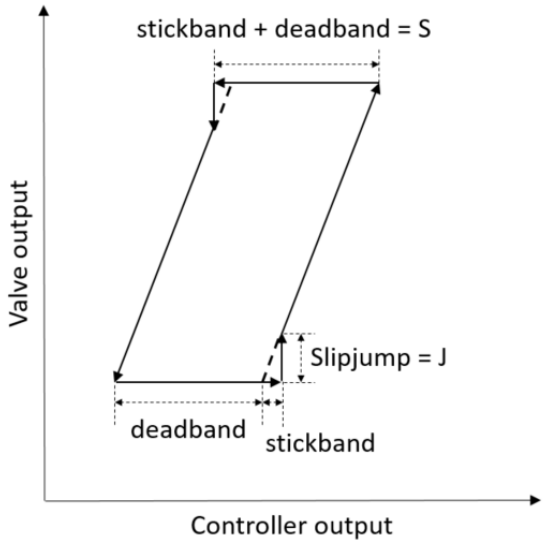


Figure 36: Typical input-output behaviour of a sticky valve, redrawn from *Choudhury, Thornhill and Shah, (2005)*

The model was implemented in Simulink (MATLAB, 2018), as shown in Figure 38, where x(k) is the controller output for the valve position and y(k) is the actual valve position. The "switch" block on the RHS of Figure 38 outputs either the moving valve position or the stuck valve position, depending on its input. If the input is one, the valve is stuck; if the input is zero, the valve is moving. The switch block input is initiated at zero, so that the valve starts off moving.

The valve becomes stuck every time it attempts to change direction (i.e. the sign of the controller output slope changes) or there is no change in the controller output for valve position. Either of these conditions activate the "sample and hold" block, which outputs the current stuck

valve position until it is activated again. The valve remains stuck while eqn D1 and eqn D2 remain true, where x_{SS} is the current stuck valve position. If the valve is currently stuck, eqn D1 is used to determine if it overcomes the slipjump; if the valve is currently moving, eqn D2 is used to determine if it moves further than the deadband + stickband.

$$|x(k) - x_{SS}| \leq J \quad [D1]$$
$$|x(k) - x_{SS}| \leq S \quad [D2]$$

The behaviour of the valve stiction fault is displayed by feeding an artificial sinusoidal input to the valve stiction model, as shown in Figure 37. Figure 37a shows the difference between the controller output (i.e. what the control valve would do if valve stiction was not present) and the actual valve output against time; and Figure 37b shows the input-output behaviour of the sticky valve, which corresponds partially to the behaviour explained in Figure 33.

The valve was initialised to start off in the moving phase, so the valve output trend in Figure 37a follows the controller output trend in the beginning of the first oscillation. However, once the valve becomes stuck for the first time, it exhibits the slipjump for the first time, which takes it directly to another stuck position, and this behaviour continues for the rest of the time window. This is because the slipjump J is a large value, so the valve never overcomes it (i.e. eqn D1 remains true the entire time after the initial movement phase). This initial movement phase followed by a continuous jumping from one stuck position to another is also seen in Figure 37b, where the slanted line represents the initial movement phase, and the rectangle represents the jumping from one stuck position to another.

The overshoot due to J > S is visible in both Figure 37a and Figure 37b. In Figure 37a, the valve output peaks and troughs above and below the controller output respectively; and in Figure 37b, a valve output larger than 90 % corresponds to a controller output of only 90 %.

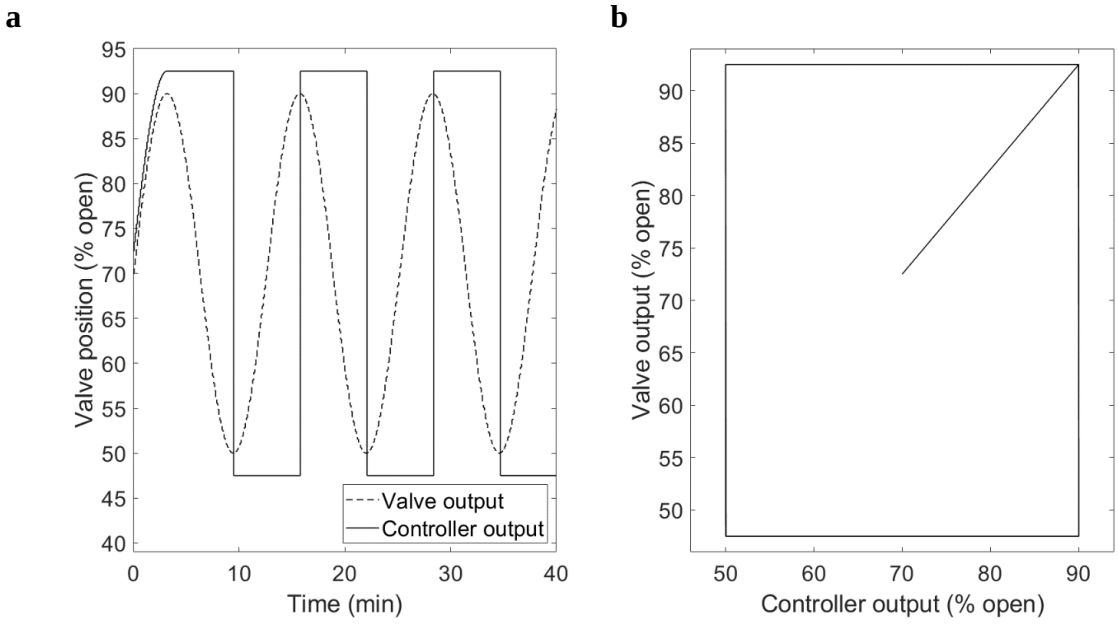


Figure 37: Valve stiction fault behaviour: **a**. The difference between controller output and valve output as a function of time. **b**. Input-output behaviour of the sticky valve modelled for this project.

This valve stiction model adjusts the controller output value to the actual valve output in terms of valve position. Linear installed valve characteristics were assumed to convert the controller output in flowrate units to valve position for the valve stiction model and back again afterwards.

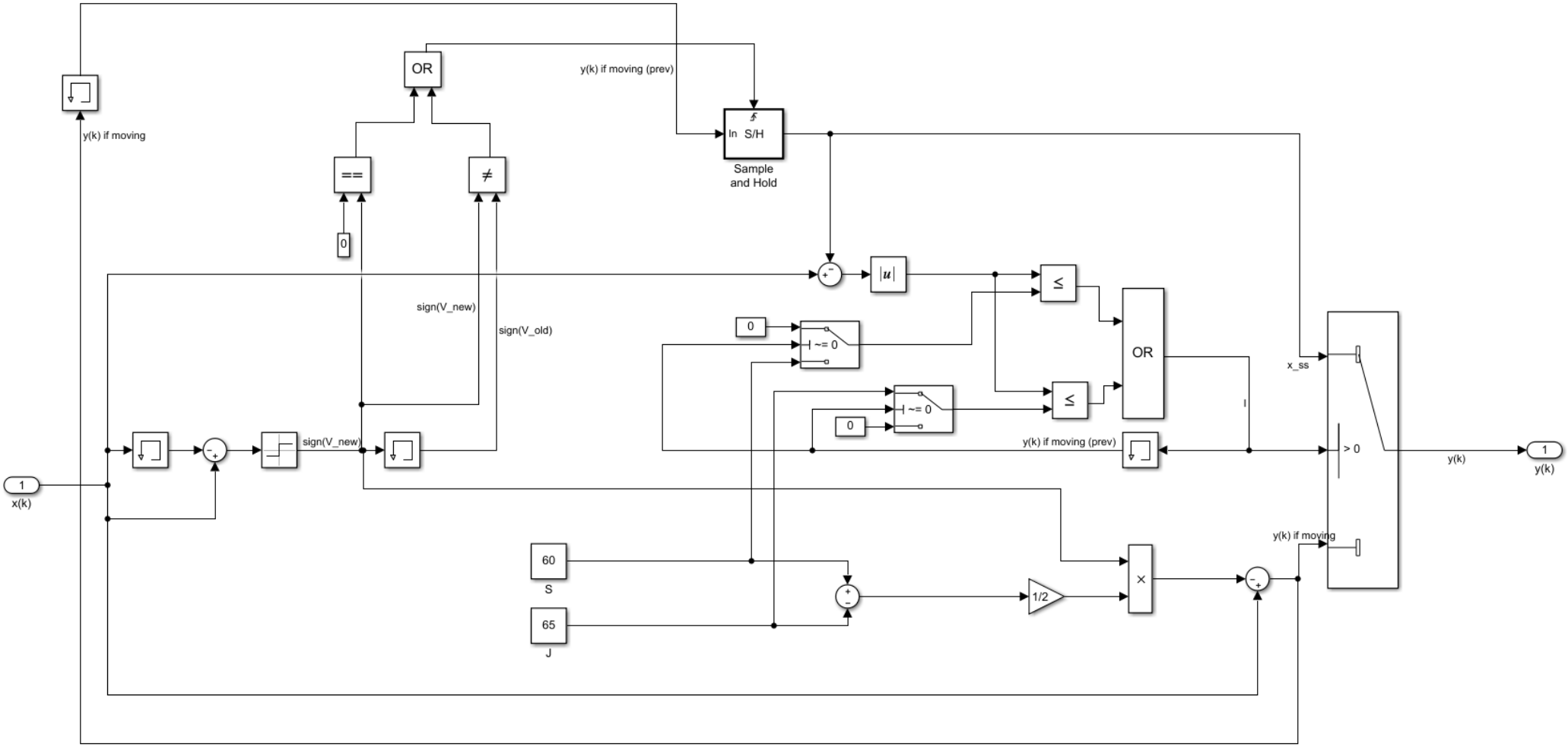


Figure 38: Simulink implementation of valve stiction model described in *Choudhury, Thornhill and Shah, (2005)*

APPENDIX E

Derivation of controller tuning relations

The linear averaging controller relation was derived for the system described in Chapter 5 and Appendix C, following the description provided in Marlin, (1995).

Writing eqn E1 in deviation variables, and ignoring the underflow (since it is negligibly small compared to the product flow, and is not involved in controlling the system):

$$\frac{dh'_{frac}}{dt} = \frac{F'_{in} - F'_{prod}}{V_{max}} \qquad [E1]$$

Taking the PI controller equation in deviation variables:

$$F'_{prod} = -K_c \left[h'_{frac} + \frac{1}{\tau_I} \int_0^t h'_{frac} \right] dt \qquad [E2]$$

Substituting eqn E2 into eqn E1:

$$\frac{V'_{max} dh'_{frac}}{dt} = F'_{in} + K_c \left[h'_{frac} + \frac{1}{\tau_I} \int_0^t h'_{frac} \right] dt \qquad [E3]$$

Converting to the Laplace domain:

$$V_{max} s h'_{frac}(s) = F'_{in}(s) + K_c \left[h'_{frac}(s) + \frac{1}{\tau_I} \frac{h'_{frac}(s)}{s} \right] \qquad [E4]$$

$$V_{max} s h'_{frac}(s) - K_c h'_{frac}(s) - \frac{K_c}{\tau_I} \frac{h'_{frac}(s)}{s} = F'_{in}(s) \qquad [E5]$$

Multiplying by s:

$$V_{max} s^2 h'_{frac}(s) - s K_c h'_{frac}(s) - \frac{K_c}{\tau_I} h'_{frac}(s) = s F'_{in}(s) \qquad [E6]$$

Factorising $h'_{frac}(s)$ out of the LHS:

$$\left[V_{max} s^2 - s K_c - \frac{K_c}{\tau_I} \right] h'_{frac}(s) = s F'_{in}(s) \qquad [E7]$$

Multiplying by $\left(-\frac{\tau_I}{K_c}\right)$:

$$\left[\frac{-V_{max} \tau_I s^2}{K_c} + \tau_I s + 1 \right] h'_{frac}(s) = \left(-\frac{\tau_I}{K_c} \right) s F'_{in}(s) \qquad [E8]$$

$$\frac{h'_{frac}(s)}{F'_{in}(s)} = \frac{\left(-\frac{\tau_I}{K_c}\right)s}{\frac{-V_{max}\tau_I s^2}{K_c} + \tau_I s + 1} \quad [E9]$$

To write eqn E9 in the form:

$$\frac{h'_{frac}(s)}{F'_{in}(s)} = \frac{\left(-\frac{\tau_I}{K_c}\right)s}{\tau^2 s^2 + 2\epsilon\tau s + 1} \quad [E10]$$

Let:

$$\tau^2 = -\frac{V_{max}\tau_I}{K_c} \quad [E11]$$

$$\tau = \sqrt{-\frac{V_{max}\tau_I}{K_c}} \quad [E12]$$

So that:

$$\epsilon = \frac{1}{2}\sqrt{\frac{-\tau_I K_c}{V_{max}}} \quad [E13]$$

Taking a step change in the inlet flowrate (i.e. the disturbance variable) in the time domain [eqn E14] and in the Laplace domain [eqn E15]; and letting the system be critically damped (i.e. $\epsilon = 1$).

$$\Delta F_{in} = \frac{dF'_{in}}{dt} \quad [E14]$$

$$F'_{in} = \frac{\Delta F}{s} \quad [E15]$$

Then substituting that information in eqn E10:

$$h'_{frac}(s) = \frac{\left(-\frac{\tau_I}{K_c}\right)s}{\tau^2 s^2 + 2\tau s + 1}\frac{\Delta F}{s} \quad [E16]$$

$$h'_{frac}(s) = \frac{\left(-\frac{\tau_I}{K_c}\right)\Delta F}{\tau^2 s + 2\tau + 1} \quad [E17]$$

$$h'_{frac}(s) = \frac{\left(-\frac{\tau_I}{K_c}\right)\Delta F}{(\tau s + 1)^2} \quad [E18]$$

Converting back to the time domain:

$$h'_{frac}(t) = \left(-\frac{\tau_I}{K_c}\right)\Delta F\frac{1}{\tau^2}te^{-\frac{t}{\tau}} \quad [E19]$$

Substituting eqn E11 in eqn E19:

$$h'_{frac}(t) = \left(-\frac{\tau_I}{K_c}\right)\Delta F(-\frac{K_c}{V_{max}\tau_I})te^{-\frac{t}{\tau}} \quad [E20]$$

$$h'_{frac}(t) = \frac{\Delta F}{V_{max}}te^{-\frac{t}{\tau}} \quad [E21]$$

Taking advantage of the fact that $\epsilon = 1$ with eqn E13:

$$\epsilon = 1 = \frac{1}{2}\sqrt{\frac{-\tau_I K_c}{V_{max}}} = 2\sqrt{\frac{V_{max}}{-\tau_I K_c}} \quad [E22]$$

Then multiplying eqn E12 by eqn E13 (i.e. multiplying by 1):

$$\tau = \sqrt{-\frac{V_{max}\tau_I}{K_c}} \times 2\sqrt{\frac{V_{max}}{-\tau_I K_c}} \quad [E23]$$

$$\tau = \frac{2V_{max}}{-K_c} \quad [E24]$$

Substituting eqn E24 into eqn E21:

$$h'_{frac}(t) = \frac{\Delta F}{V_{max}} t e^{-t\left(-\frac{K_c}{2V_{max}}\right)} \quad [E25]$$

Taking the derivative of eqn E25 using the chain rule, and setting it to zero to find its maximum:

$$\frac{dh'_{frac}(t)}{dt} = \frac{\Delta F}{V_{max}}\left[t\frac{-K_c}{2V_{max}} e^{-t\left(-\frac{K_c}{2V_{max}}\right)} + e^{-t\left(-\frac{K_c}{2V_{max}}\right)}\right] = 0 \quad [E26]$$

The solution to eqn E26 is:

$$t_{max} = \frac{2V_{max}}{-K_c} \quad [E27]$$

Substituting the solution into eqn E25:

$$\Delta h_{frac_{max}} = \frac{\Delta F_{max}}{V_{max}} \frac{2V_{max}}{-K_c} e^{-\left(\frac{2V_{max}}{-K_c}\right)\left(-\frac{K_c}{2V_{max}}\right)} \quad [E28]$$

$$\Delta h_{frac_{max}} = \frac{2\Delta F_{max}}{-K_c} e^{-1} \quad [E29]$$

Simplifying to give the controller tuning relation provided in Marlin, (1995):

$$\Delta h_{frac_{max}} = 0.736 \frac{\Delta F_{max}}{-K_c} \quad [E30]$$

$$K_c = -\frac{0.736\Delta F_{max}}{\Delta h_{frac_{max}}} \quad [E31]$$

K_c was therefore determined using eqn E31, and τ_I by rearranging eqn E13 .

APPENDIX F

Tank network connectivity matrix

	Pipe 001	FI 001	TK 001	LI 001	LC 001	V 001	Pipe 002	P 001	FI 002	Pipe 003	TK 002	Pipe 015	LI 002	LC 002	V 002	Pipe 004	P 002	FI 003	Pipe 005	BTK 001	Pipe 006
Pipe 001	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 001	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 001	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 001	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 001	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 001	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Pipe 002	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
P 001	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 002	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Pipe 003	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
TK 002	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	1
Pipe 015	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 002	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
LC 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
V 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Pipe 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0
P 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
FI 003	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
Pipe 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0

APPENDIX F

	Pipe 001	FI 001	TK 001	LI 001	LC 001	V 001	Pipe 002	P 001	FI 002	Pipe 003	TK 002	Pipe 015	LI 002	LC 002	V 002	Pipe 004	P 002	FI 003	Pipe 005	BTK 001	Pipe 006
BTK 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Pipe 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 018	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 007	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 008	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 009	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 010	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BTK 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 011	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX F

	Pipe 001	FI 001	TK 001	LI 001	LC 001	V 001	Pipe 002	P 001	FI 002	Pipe 003	TK 002	Pipe 015	LI 002	LC 002	V 002	Pipe 004	P 002	FI 003	Pipe 005	BTK 001	Pipe 006
Pipe 020	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 019	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 007	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 007	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 012	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 008	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 013	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 008	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 017	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 014	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 009	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 021	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX F

	Pipe 018	FI 004	TK 004	LI 003	LC 003	V 003	Pipe 007	P 003	FI 005	Pipe 008	TK 005	Pipe 016	LI 004	LC 004	V 004	Pipe 009	P 004	FI 006	Pipe 010	BTK 002	Pipe 011
Pipe 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 015	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BTK 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 006	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX F

	Pipe 018	FI 004	TK 004	LI 003	LC 003	V 003	Pipe 007	P 003	FI 005	Pipe 008	TK 005	Pipe 016	LI 004	LC 004	V 004	Pipe 009	P 004	FI 006	Pipe 010	BTK 002	Pipe 011
Pipe 018	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 004	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 004	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 003	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 003	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 003	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Pipe 007	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
P 003	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 005	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 008	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
TK 005	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	1
Pipe 016	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 004	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
LC 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
V 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Pipe 009	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0
P 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
FI 006	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
Pipe 010	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
BTK 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Pipe 011	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	Pipe 018	FI 004	TK 004	LI 003	LC 003	V 003	Pipe 007	P 003	FI 005	Pipe 008	TK 005	Pipe 016	LI 004	LC 004	V 004	Pipe 009	P 004	FI 006	Pipe 010	BTK 002	Pipe 011
Pipe 020	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 019	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 007	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 007	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 012	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 008	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 013	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 008	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 017	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 014	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 009	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 021	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX F

	Pipe 020	Pipe 019	FI 007	TK 007	LI 005	LC 005	V 005	Pipe 012	P 005	FI 008	Pipe 013	TK 008	Pipe 017	LI 006	LC 006	V 006	Pipe 014	P 006	FI 009	Pipe 021
Pipe 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 015	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BTK 001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	Pipe 020	Pipe 019	FI 007	TK 007	LI 005	LC 005	V 005	Pipe 012	P 005	FI 008	Pipe 013	TK 008	Pipe 017	LI 006	LC 006	V 006	Pipe 014	P 006	FI 009	Pipe 021
Pipe 018	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 007	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 003	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 008	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 009	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P 004	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 010	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BTK 002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 011	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX F

	Pipe 020	Pipe 019	FI 007	TK 007	LI 005	LC 005	V 005	Pipe 012	P 005	FI 008	Pipe 013	TK 008	Pipe 017	LI 006	LC 006	V 006	Pipe 014	P 006	FI 009	Pipe 021
Pipe 020	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 019	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
FI 007	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TK 007	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
LI 005	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LC 005	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
V 005	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Pipe 012	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0
P 005	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
FI 008	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Pipe 013	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
TK 008	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0
Pipe 017	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LI 006	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
LC 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
V 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Pipe 014	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0
P 006	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
FI 009	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0
Pipe 021	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

APPENDIX G

Sampling period and time window plots

G.1. Plots for a standard causality map

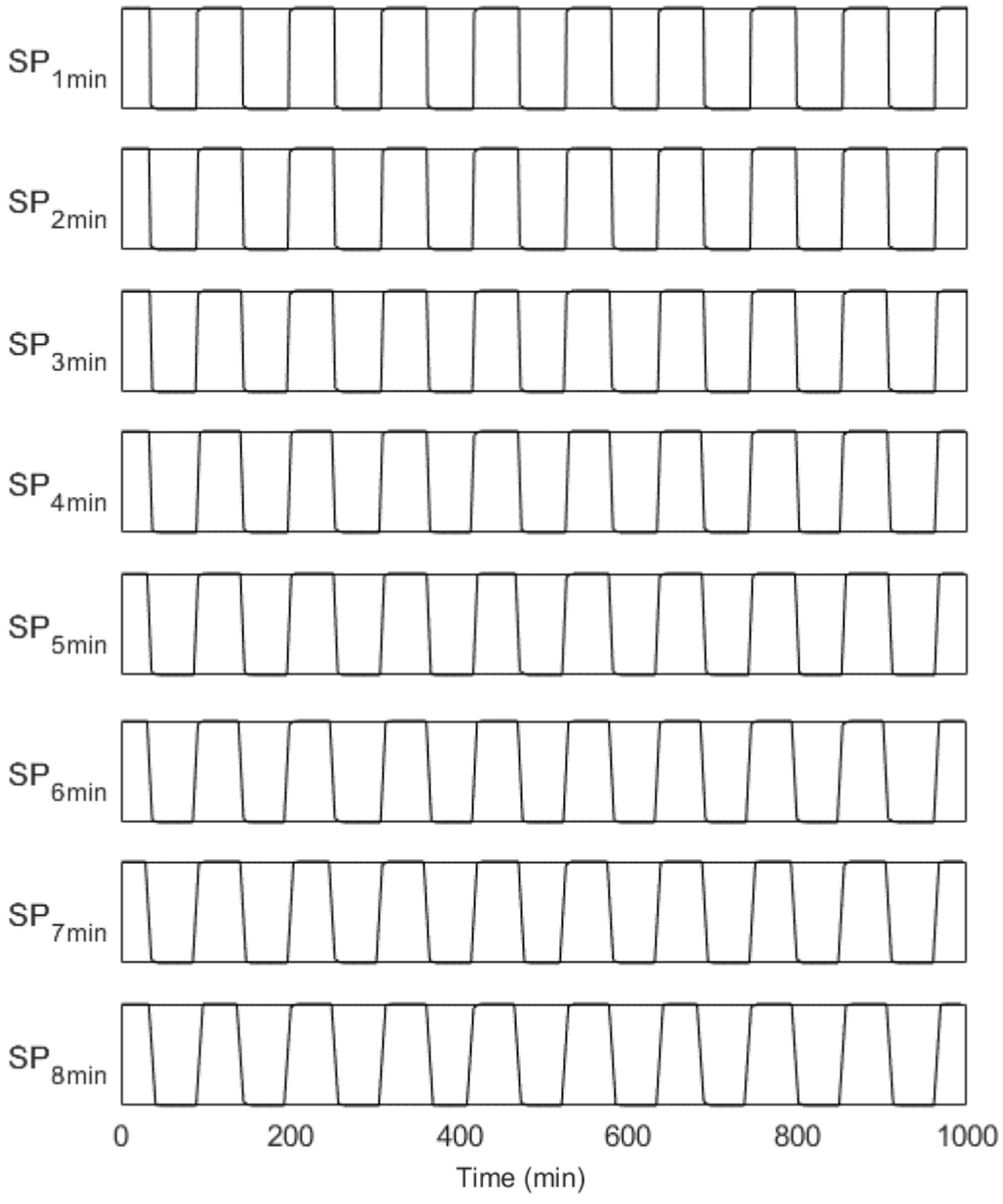


Figure 39: FI 006 time series with sampling periods 1 – 8 min.

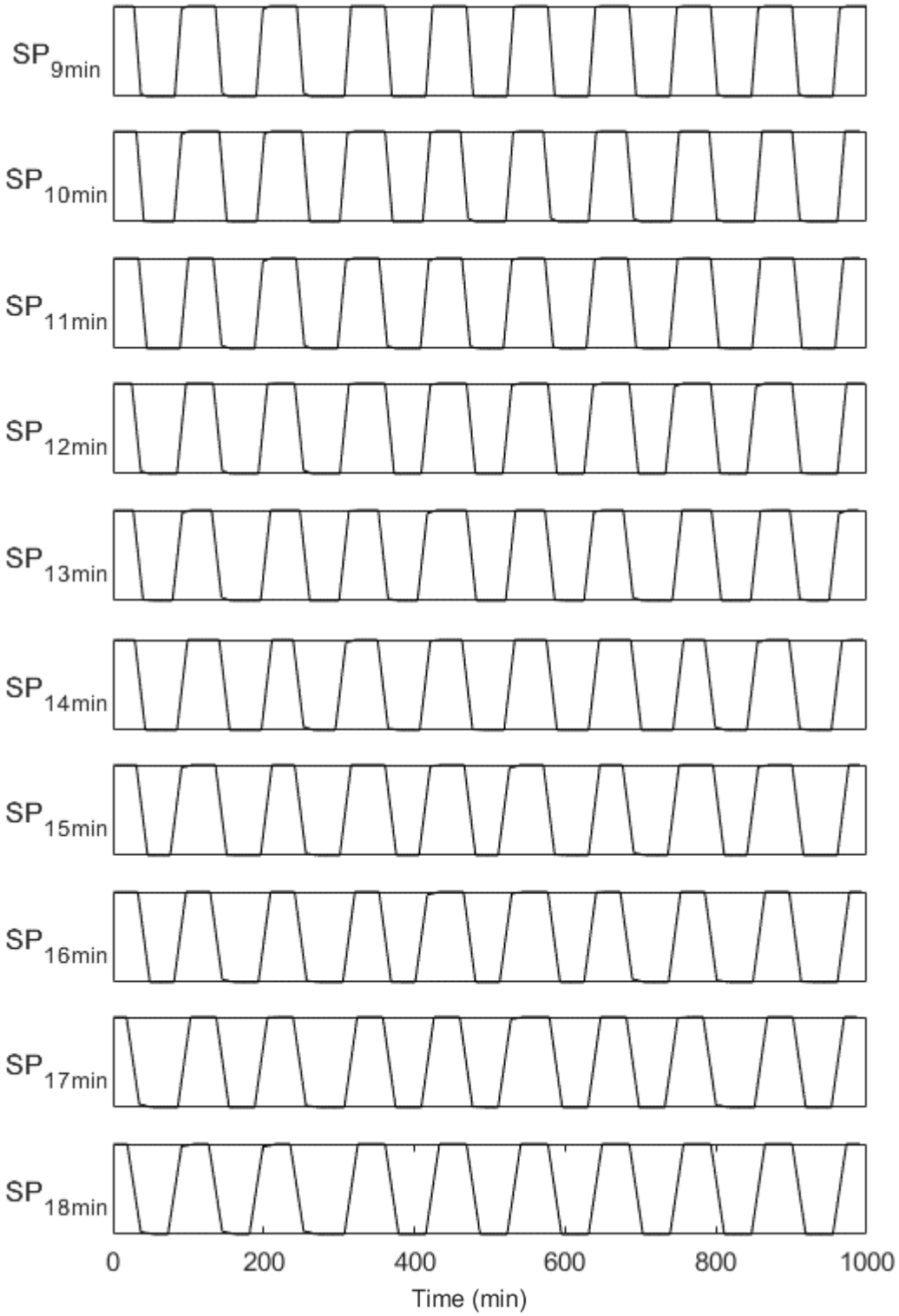


Figure 40: FI 006 time series with sampling periods 9 – 18 min.

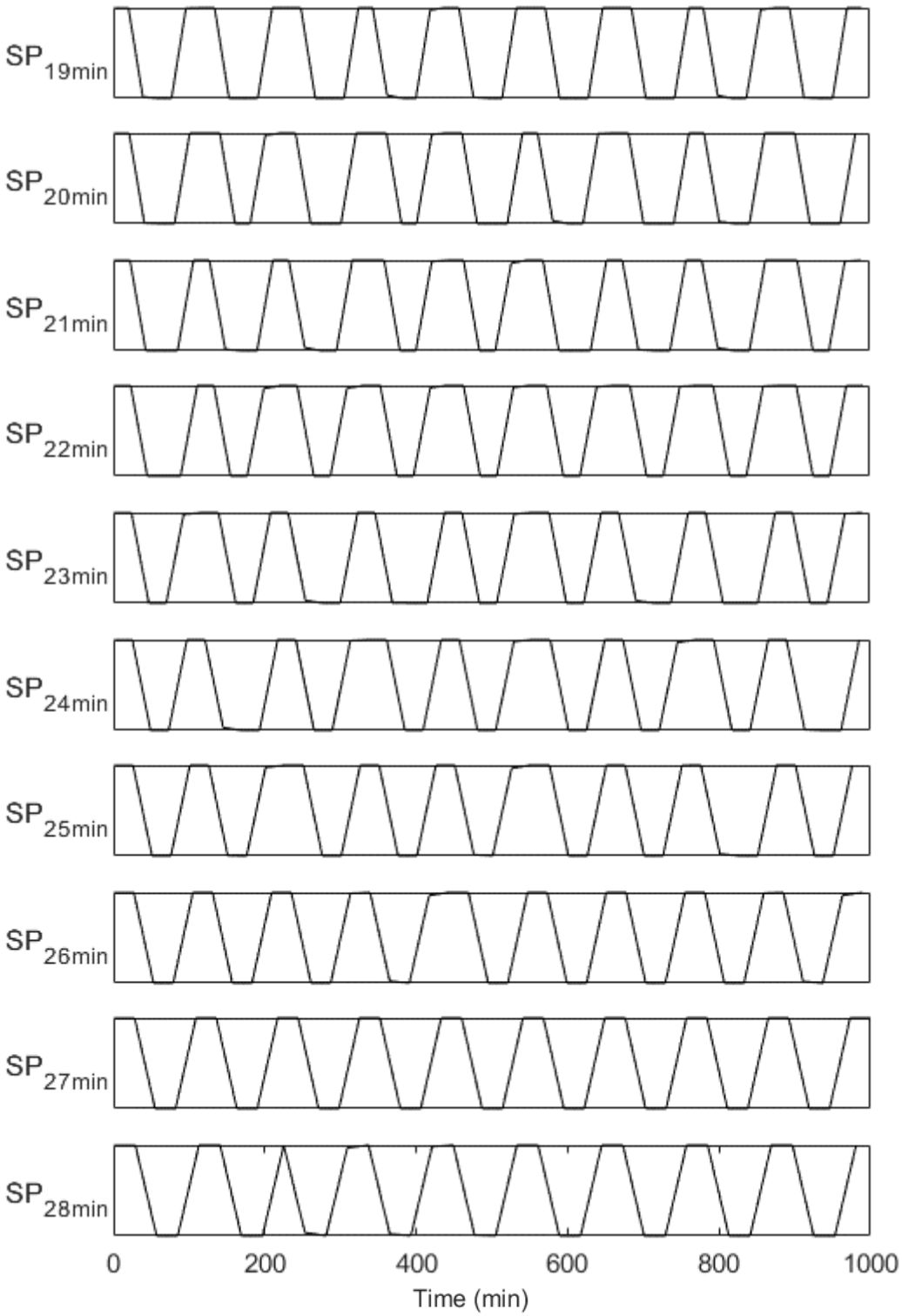


Figure 41: FI 006 time series with sampling periods 19 – 28 min.

G.2. Plots for the proposed hierarchical approach

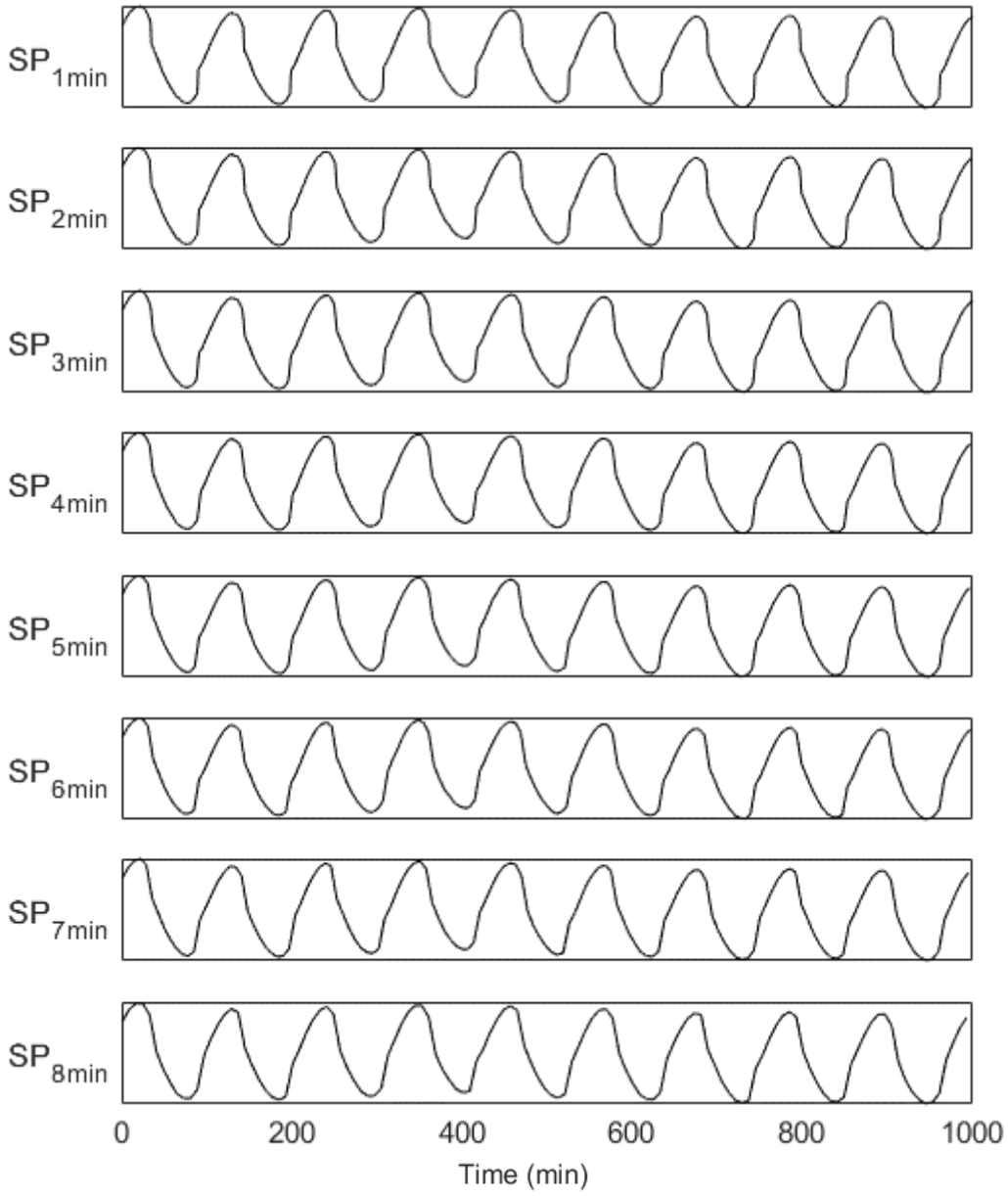


Figure 42: Plant section 2 time series scores with sampling periods 1 – 8 min.

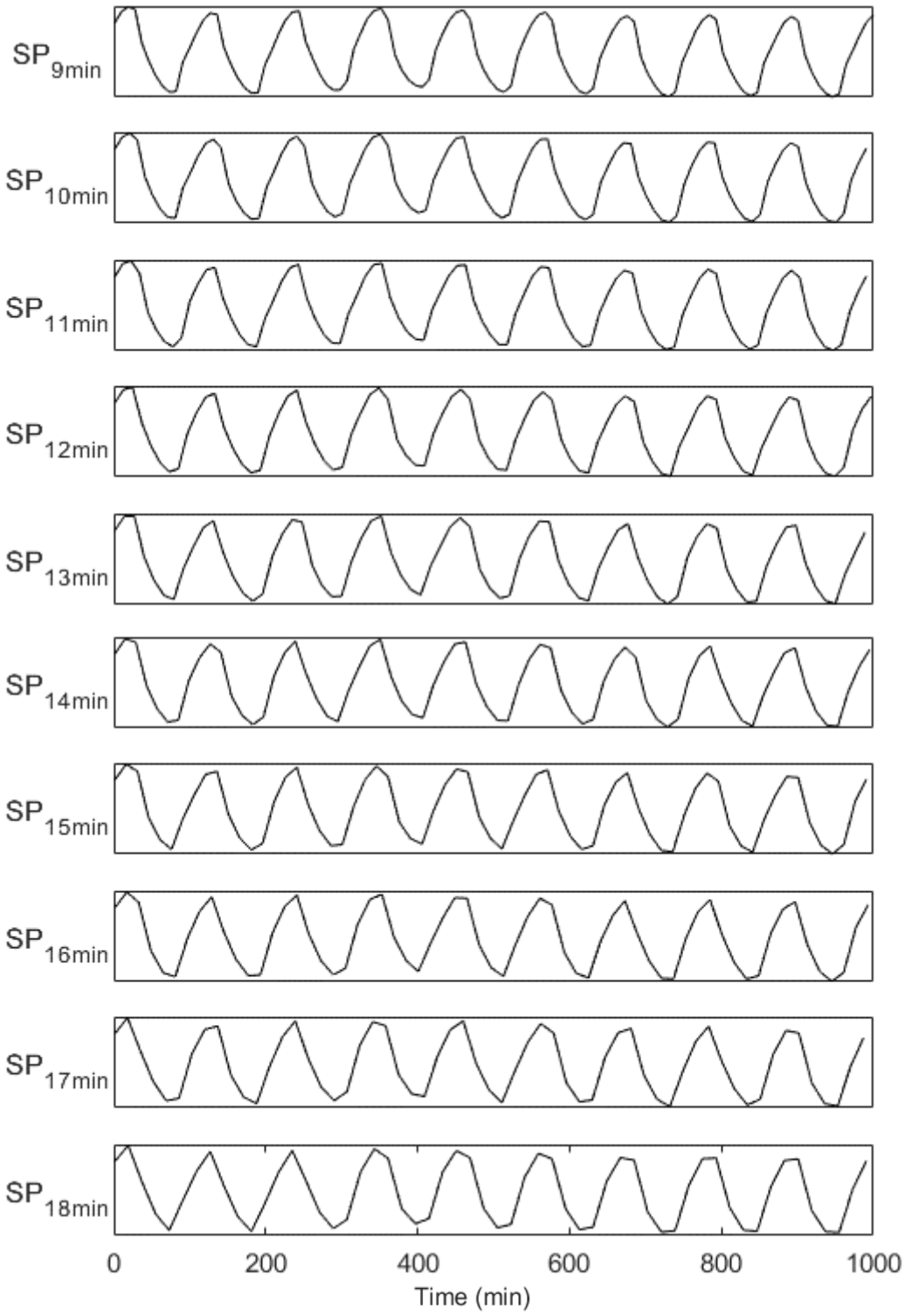


Figure 43: Plant section 2 time series scores with sampling periods 9 – 18 min.

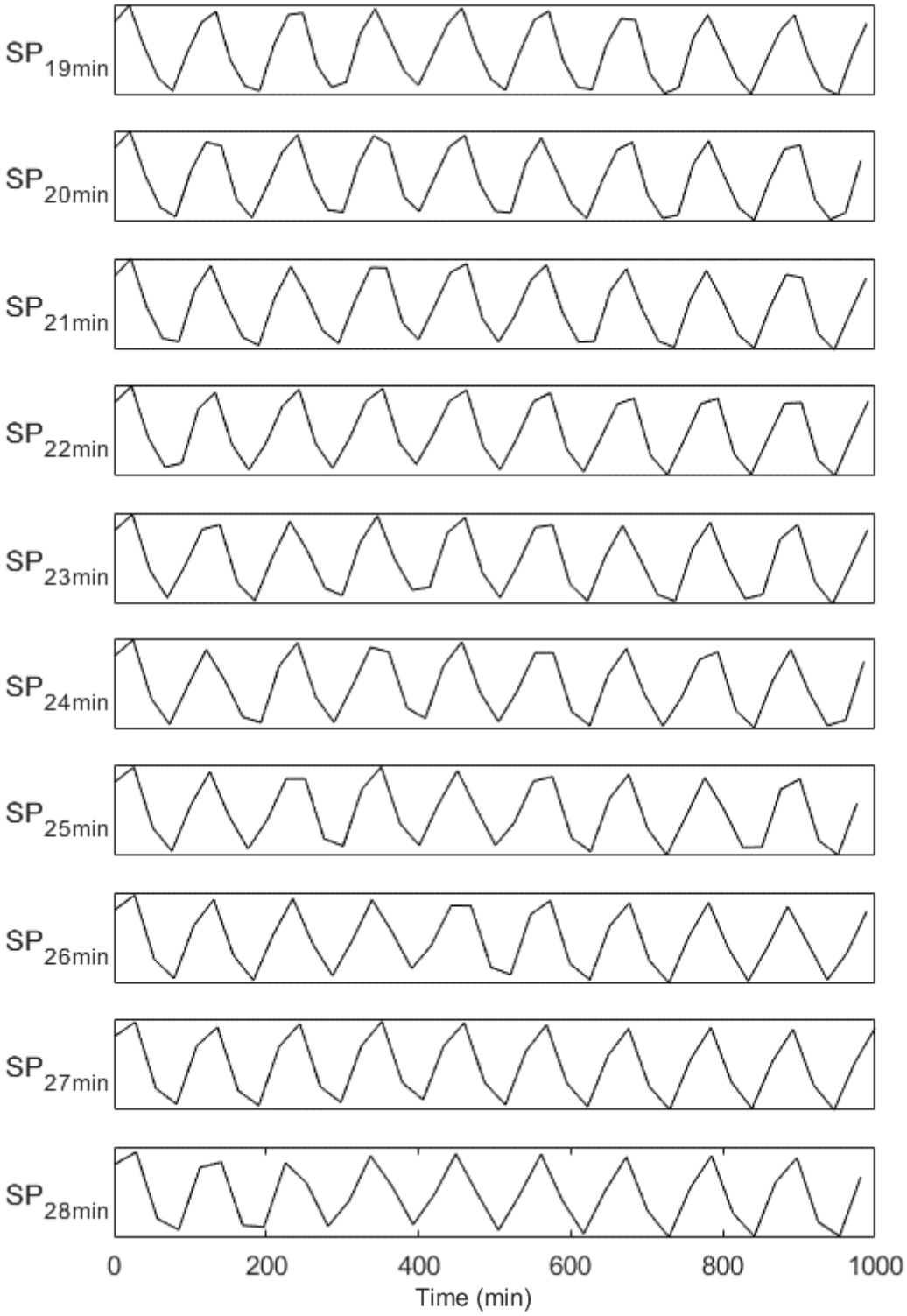


Figure 44: Plant section 2 time series scores with sampling periods 19 – 28 min.

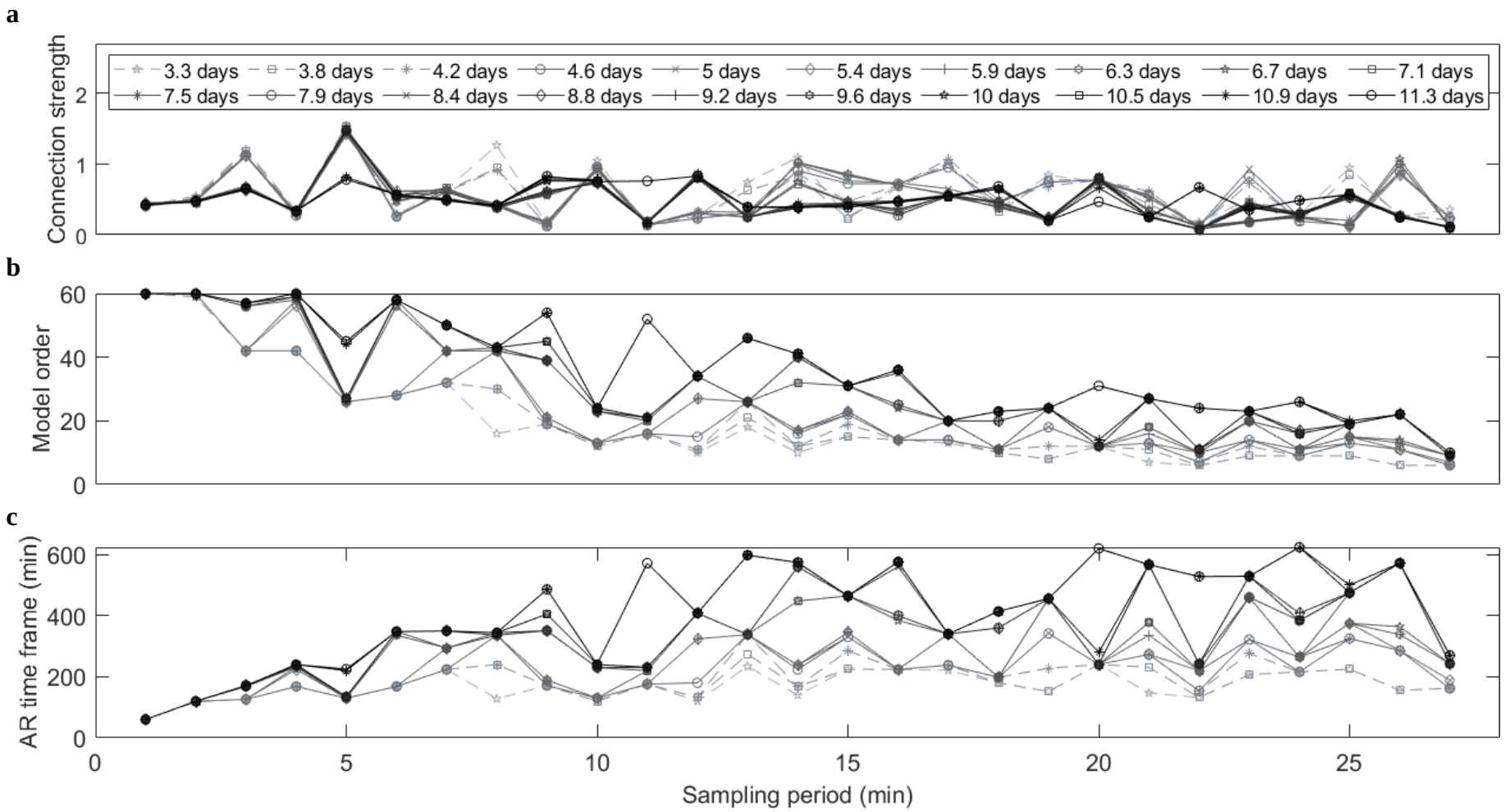


Figure 45: The effects of time window and sampling period on **a**: the strength of the known causal connection PS2→PS1, **b**: the model order selected by AIC, and **c**: the resulting autoregressive (AR) time frame for time windows required for at least **602 samples** at each SP. *Darker lines indicate larger TWs.*

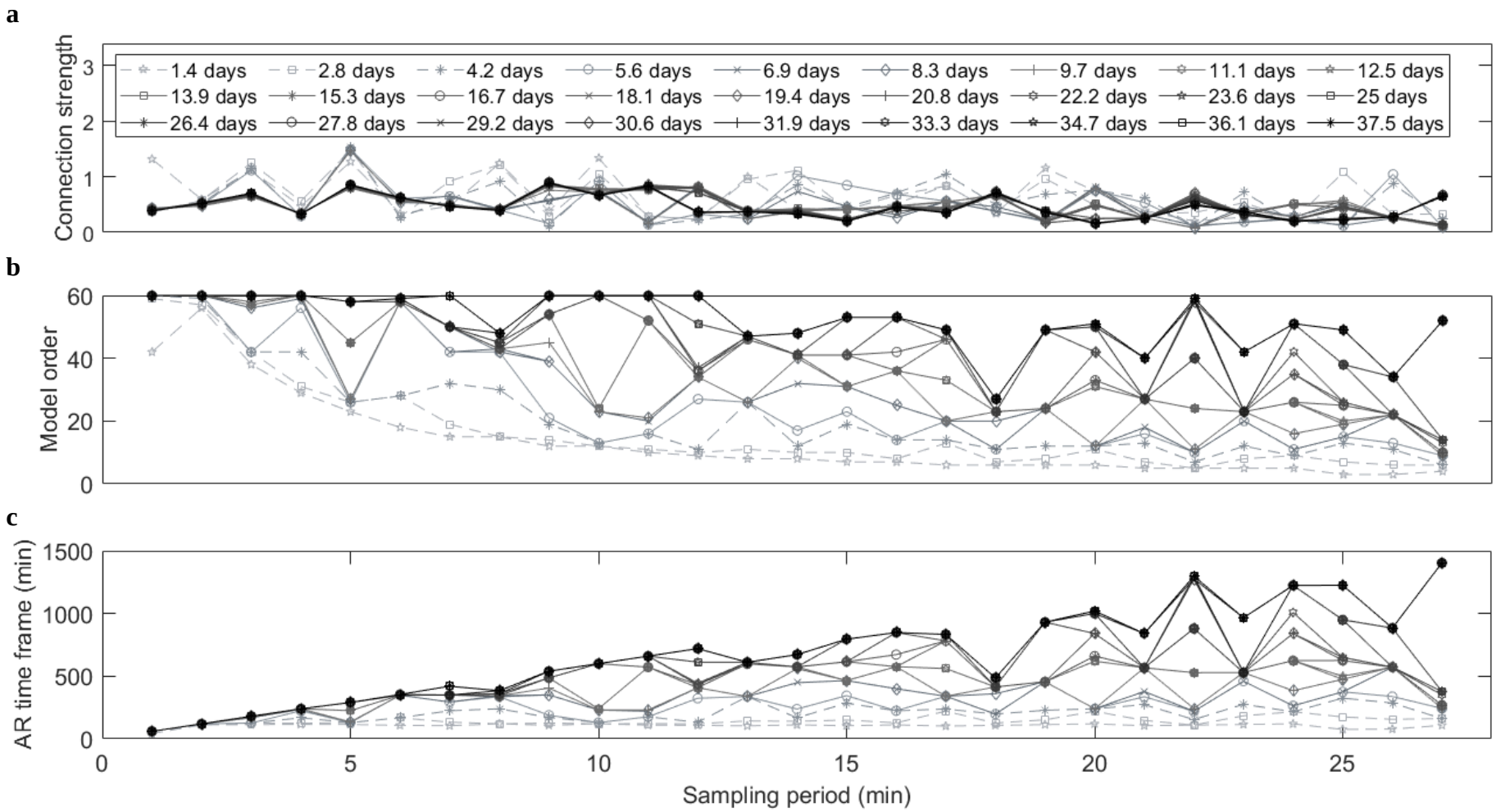


Figure 46: The effects of time window and sampling period on **a**: the strength of the known causal connection PS2→PS1, **b**: the model order selected by AIC, and **c**: the resulting autoregressive (AR) time frame for the time windows required for at least **2000 samples** at each SP. *Darker lines indicate larger TWs.*

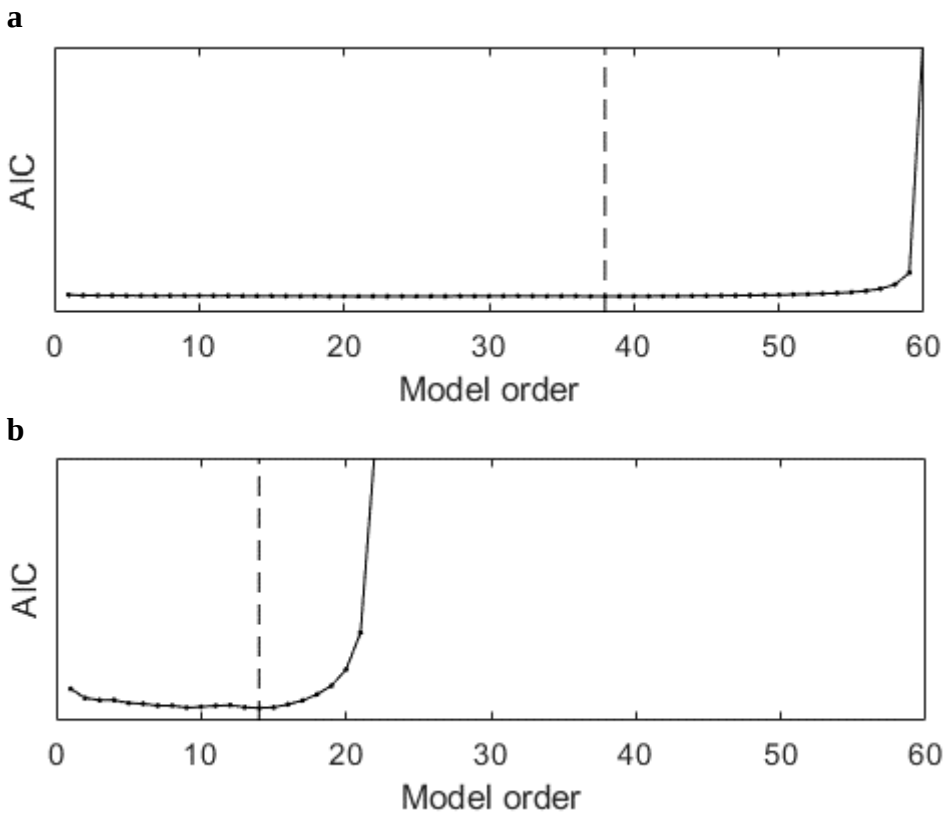


Figure 47: AIC plots for the range of allowable mode orders, with vertical dashed lines indicating the selected model order. **a**. AIC plot for SP3 min and TW1.3 days. **b**. AIC plot for SP6 min and TW2.9 days. *For this plot, AIC is only calculated for a model order up to 22, because too few samples are available for a unique OLS solution for the full model at higher model orders for this TW.*

APPENDIX H

Survey

H.1. Survey questionnaire

Improving the interpretability of causality maps for fault identification survey[1]

INFORMED CONSENT

TITLE OF RESEARCH PROJECT:	Improving the interpretability of causality maps for fault identification
REFERENCE NUMBER:	14385
PRINCIPAL INVESTIGATOR:	Natali van Zijl
ADDRESS:	Department of Process Engineering, Stellenbosch University, Banghoek Road, Stellenbosch Central, 7600, South Africa
CONTACT NUMBER:	082 824 3670
E-MAIL:	19010370@sun.ac.za

Dear prospective participant

Kindly note that I am a MEng student at the Department of Process Engineering at Stellenbosch University, and I would like to invite you to participate in a research project entitled "Improving the interpretability of causality maps for fault identification".

Please take some time to read the information presented here, which will explain the details of this project and contact me if you require further explanation or clarification of any aspect of the study. This study has been approved by the Research Ethics Committee (REC) at Stellenbosch University and will be conducted according to accepted and applicable national and international ethical guidelines and principles.

1. **INTRODUCTION:** Modern chemical and mineral processing plants continuously aim to increase their productivity to address stiff global competition. Rather than introduce costly new initiatives, process monitoring and optimisation are used to ensure optimal use of existing processes and infrastructure. A current bottleneck in process monitoring is the identification of the root cause of a fault. Fortunately, research from the last decade has shown data-based causality analysis techniques to be useful for fault diagnosis. Causality analysis uses the cause-effect relationships between variables to construct a causality map, where the propagation path of a fault can be traced back to its root cause. However, due to poor causality map interpretability, engineers struggle to understand and use causality analysis results for fault diagnosis, and it has yet to become widely accepted in industry.
2. **PURPOSE:** The purpose of this project is to improve the interpretability of causality maps for fault identification, and so to make this technique usable in industry. A hierarchical approach

[1] This questionnaire was reformatted for use within SUN Surveys format.

has been proposed to construct causality maps for fault identification, and the purpose of this survey is to investigate whether this approach can be used to successfully identify a fault in an engineering process, and subsequently advise on whether the interpretability of causality maps can be improved to the extent where they can be used for fault identification in industry.

3. **PROCEDURES:** If you voluntarily choose to participate in this study, you will be asked to complete a once-off online survey.
4. **TIME:** The once-off online survey will take approximately 30-40 minutes to complete.
5. **RISKS:** No potential risk or discomfort is foreseen as a result of participating in this study, other than the time and effort invested in completing the online survey.
6. **BENEFITS:** If you volunteer to participate in this study, you may choose to enter into a lucky draw, where you stand the chance to win a Woolworths voucher valued at R500. Furthermore, your participation will give us insight into whether the proposed approach to improve the interpretability of causality maps is a usable tool for fault identification by users who can reasonably be expected to understand the concept of a causality map. This will help us to make a recommendation as to whether causality maps should be widely used in industry.
7. **PARTICIPATION & WITHDRAWAL:** You may choose whether you want to participate in this study or not. There are no consequences for you if you withdraw from this study or your data is withdrawn. If you want to withdraw, all data gathered from you will be deleted from all platforms where it was stored as soon as the primary investigator is informed of your desire to withdraw.
8. **CONFIDENTIALITY:** All information obtained in this study will remain confidential and disclosed only with your permission or as required by law. At no point in this survey will you be asked for identifiable information, beyond background information regarding the institution you are affiliated with, your highest level of qualification, and whether you have any prior experience with causality analysis. All participants will complete the survey anonymously and the results will be processed in such a manner that individual respondents cannot be identified. All data will be stored confidentially at the password-protected Stellenbosch University SUN Survey server, the primary investigator's password-protected personal computer, and an access-restricted SharePoint, where access is restricted to the primary investigator, Natali van Zijl, her Stellenbosch University-affiliated supervisors, Dr TM Louw, Prof SM Bradshaw, and Dr L Auret, and their collaborator, Dr De Villiers Groenewald, the lead control engineer at Anglo American Platinum. The primary investigator, Natali van Zijl, is responsible for data collection and analysis, and she is committed to governing the data securely and confidentially. Aggregated results will be published in the form of a thesis for a master's degree in Extractive Metallurgical Engineering. Please note that aggregated results may be linked to the aforementioned background information and published as such. Results could also be presented in conference presentations and journal articles.
9. **RECORDINGS:** This study does not involve any interviews. Therefore, no voice or video recordings will be made use of.
10. **DATA STORAGE:** The data from this survey will be anonymous and stored on the primary investigator, Natali van Zijl's, password protected personal computer, as well as on an access-restricted SharePoint, where access is restricted to the primary investigator, Natali van Zijl, her SU-affiliated supervisors, Dr TM Louw, Prof SM Bradshaw, and Dr L Auret, and their collaborator, Dr De Villiers Groenewald, the lead control engineer at Anglo American Platinum. This data could be used in further investigations into the interpretability of causality maps by any of the primary investigator's SU-affiliated supervisors or the collaborator, as listed above. All email addresses obtained from participants who choose to participate in the lucky draw will be temporarily stored on the primary investigator's password protected Outlook account and password protected personal computer until the winner has been determined and contacted. Thereafter, the email addresses will be immediately deleted from the primary investigator's Outlook account and personal computer, and permanently deleted from the deleted folder and the recycling bin respectively.

If you have any questions or concerns about this research project, please feel free to contact me, Natali van Zijl, at 19010370@sun.ac.za or 082 824 3670, or my supervisor, Tobi Louw, at tmlouw@sun.ac.za.

RIGHTS OF RESEARCH PARTICPANTS: You may withdraw your consent at any time and discontinue participation without penalty. You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you have questions regarding your rights as a research subject, contact Ms Maléne Fouché (mfouche@sun.ac.za / 021 808 4622) at the Division for Research Development. You have the right to receive a copy of this Consent form.

If you click the "yes" button below, you as the participant confirm that:

- I have read the above information and it is written in a language with which I am fluent and comfortable.
- I have had a chance to ask questions and all my questions have been adequately answered.
- I understand that taking part in this study is voluntary and I have not been pressurised to take part.
- I may choose to leave the study at any time and will not be penalised or prejudiced in any way.
- If the principal investigator feels that it is in my best interest, or if I do not follow the study plan as agreed to, then I may be asked to leave the study before it has finished.
- All issues related to privacy, and the confidentiality and use of the information I provide, have been explained to my satiSPaction.

Thank you for your time and effort. Your input is highlight valued.

Yours sincerely

Natali van Zijl

*Informed consent

I hereby consent to participate in the study.

Yes ☐
No ☐

--- PAGE BREAK ---

ORIENTATION TO SURVEY

What is fault identification?

Modern chemical and mineral processing plants are highly interconnected via process units, equipment, material flow, energy flow, and information flow; and this is complicated even further by recycle streams, complex control strategies and control interaction. This interconnectivity creates the opportunity for something called the smearing effect, which is when a fault originates somewhere in a plant, and then propagates throughout the plant, so that

numerous variables end up showing an effect of the fault. Since so many variables show an effect of the fault, it is difficult to identify where the fault originated, or in other words, to find its root cause – and that is exactly what fault identification entails.

What is a causality map?

Causality refers to the cause-effect relationships between variables, where change in one variable *causes* change in another. The causal structure of a process (i.e. cause-effect relationship between all the variables in the process) can be determined using data-driven techniques, or from process knowledge in the form of piping and instrumentation diagrams (P&IDs), where each approach has its own advantages and disadvantages. In short, the data-driven techniques can give quantitative causal connections in the form of causal strengths, but could give some spurious (incorrect) results; while causal structures obtained from process knowledge can only be qualitative, but can be trusted to be correct.

Once the causal structure of a process has been determined, it can be presented in a causality map, with a simple example of three variables (X, Y, Z) provided in Figure . The nodes are the variables, and the edges are the causal connections between them (so X causes Y, which causes Z).

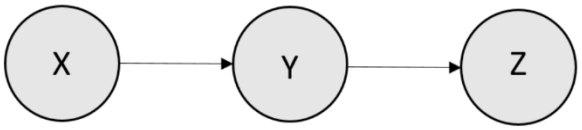


Figure 1: Example of causality map with three variables. *Nodes represent variables and edges represent casual connections.*

How can you use causality analysis for fault identification?

Causality maps can be used for fault identification by tracing the fault propagation path back to its root cause. This is easy if the causality map resembles the one in Figure , where there are no cycles, and no variables *cause* X, so it is identified as the root cause. However, it would be difficult to identify the root cause if there was also a causal connection from Z to X in Figure , especially if the causal connections were determined using a data-driven technique, so we cannot be completely sure if we can trust every connection. To address these issues, some tools to aid in the interpretation of causality maps have been included in the causality maps presented in this study:

(1) Data-driven techniques are used to determine the causal structures presented in the causality maps in this study, but these connections are validated using process knowledge. **Solid lines** are used to represent causal connections where the data-driven technique and the process knowledge agree, and **dashed lines** are used to represent causal connections identified by the data-driven technique but are not validated by process knowledge.
(2) **Causal strength** is indicated with **line thickness**.
(3) The root cause of a fault is assumed to have a large influence on the causal network, so nodes (variables) are ranked according to their influence on the causal network. This is represented by **colouring the nodes according to their rank** with a colour bar, where the colour at the top of the colour bar indicates the largest influence on the causal

network, and the colour at the bottom of the colour bar indicates the smallest influence on the causal network.

(4) In this study, the type of fault addressed is one that causes oscillations to occur. Only variables that show an effect of the fault are included in the causality map, and this is determined as variables that exhibit oscillations at the same frequency. In order for a variable to be the root cause of the fault, it therefore needs to be connected (directly or indirectly) to all the other variables in the causality map, because otherwise the fault could not have propagated from it to all the other variables. **Nodes (variables) that are not connected to all the other variables in the causality map** are therefore **faded** in this study, as they should likely not be considered as potential root causes. However, if all the nodes should be faded, then none of them are – to prevent information from being lost.

(5) There are two sliders for each causality map presented in this study, namely a connections-slider and a variables-slider (Figure 2). The **connections-slider** acts as a threshold for how strong a causal connection must be for it to be shown on the causality map (i.e. a **lower value** on the connections-slider means only the stronger connections will be displayed – so **fewer connections will be displayed**); and the **variables-slider** acts as a threshold for how much a variable must oscillate at the common frequency for it to be included in the causality analysis and shown on the causality map (i.e. a **lower value** on the variables-slider means only the variables that oscillate more will be included – so **fewer variables will be displayed**).

Connections-slider		
Lowest value (fewest connections)		Highest value (most connections)
Variables-slider		
Lowest value (fewest variables)		Highest value (most variables)

Figure 2: Diagram of the connections-slider and variables-slider

In this study, a hierarchical approach has been proposed to identify the root cause of a fault in a process. Two causality maps are constructed, where the first causality map consists of nodes that represent groups of variables, and should be used to trace the fault to one of those groups of variables; and the second causality map consists of nodes that represent each of the variables in the group identified in the first causality map, and should be used to trace the root cause to specific variable(s).

In this survey, you will be asked to determine the root cause of a fault from a base case, which is a causality map where transitive reduction has been applied. Thereafter, two variations of the proposed hierarchical approach will be presented, which you will also use to determine the root cause of a fault. The base case and two variations will be further explained in their relevant sections.

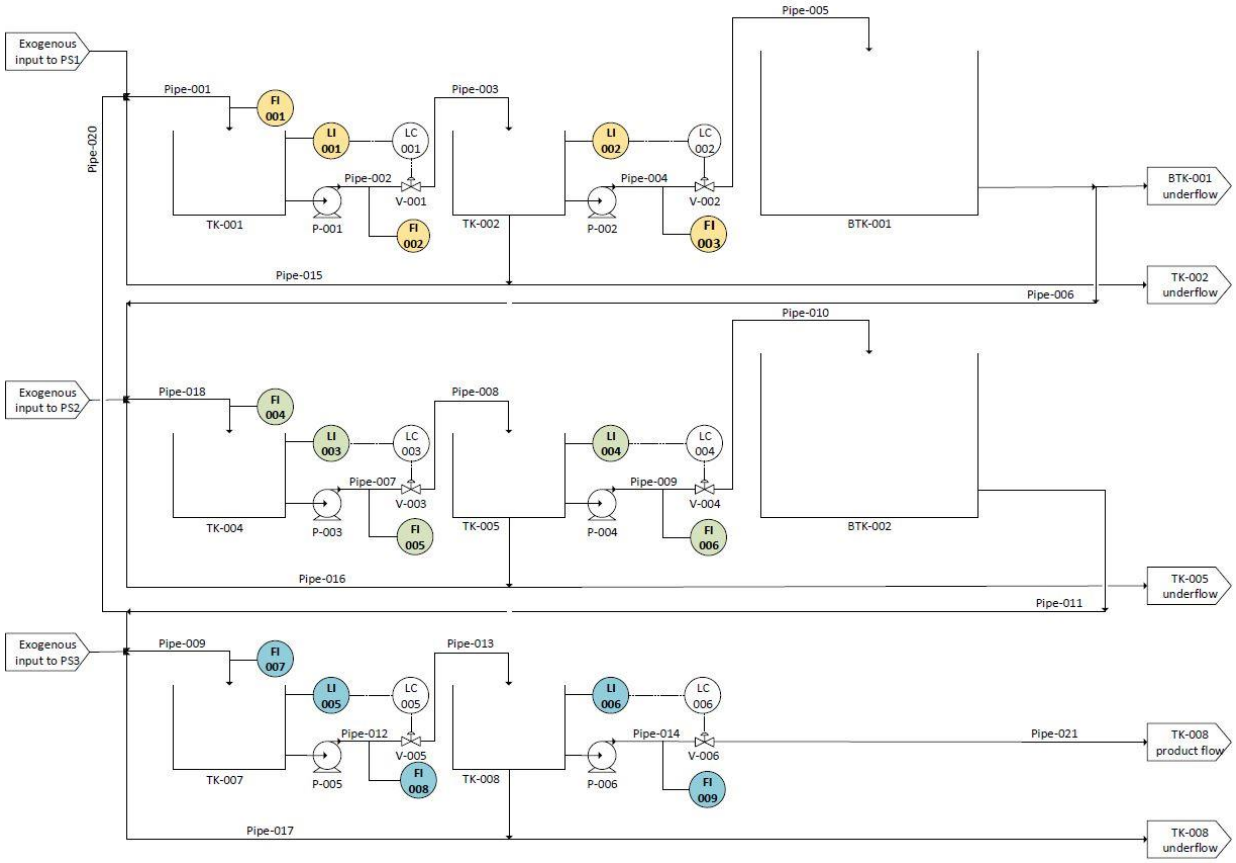
--- PAGE BREAK ---

BASE CASE: PLANT-WIDE CAUSALITY MAP WITH TRANSITIVE REDUCTION

The causality maps presented in this survey are constructed for the process shown below. The plant consists of three plant sections (PS1, PS2, PS3), which each have:

- Two tanks in series, where each tank has feedback level control.
- An exogenous input to the first tank.
- A recycle stream from the second tank's underflow to the first tank.

The plant sections are separated by buffer tanks, with a portion of each buffer tank's underflow fed to the following plant section. In addition, there is a recycle stream from the buffer tank following PS2 back to PS1.



3.1.*Click on the following link to open a presentation of the causality map(s)[2]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

This causality map is the transitive reduction of the causality map that consists of all the variables of in the process that show an effect of the fault. A transitive reduction is a map where all shortcut connections have been removed. For example, in Figure 3, X is shown to *cause* Z, and Z to *cause* Y, but X is also shown to *cause* Y directly. In the transitive reduction, the $X \rightarrow Y$ connection would be removed, as that causal connection is already captured in the $X \rightarrow Z \rightarrow Y$ connections (making it a shortcut connection).

[2] Refer to Section H.2 of this appendix for the causality maps.

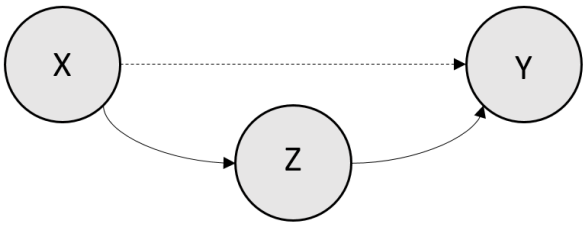


Figure 3: Example of shortcut connection from $X \rightarrow Y$

https://docs.google.com/presentation/d/e/2PACX-1vQTfr-3a3vdQC82o44nuBTzmlgDR04JNp6t_cFIqhPP2jnp75xGAk8Hlw_aO1tgXb1vg1uMscsdKNYR/pub?start=false&loop=false&delayms=3000&slide=id.p1

3.2.*Which variable is the root cause?

LI-001	
LI-002	
FI-001	
FI-002	
FI-003	
LI-003	
LI-004	
FI-004	
FI-005	
FI-006	
LI-005	
LI-006	
FI-007	
FI-008	
FI-009	
I can't decide between multiple variables.	

Routed question: If participant indicates *I can't decide between multiple variables.*

3.3.Which variables can you not decide between?

LI-001	
LI-002	
FI-001	
FI-002	
FI-003	
LI-003	

LI-004	
FI-004	
FI-005	
FI-006	
LI-005	
LI-006	
FI-007	
FI-008	
FI-009	

3.4.*Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 4 indicating the smallest role.

Node colour	
Edge weight	
Edge lines (solid/dashed)	
Number of arrows pointing towards/away from node(s)	

3.5.*Did you use the connections-slider for this causality map?

Yes	No

Routed question: If participant indicates *Yes*

3.6.*What value on the connections-slider did you end on?

1e-16	
0.01 (default)	
0.05	

3.7.*Did you use the variables-slider for this causality map?

Yes	No

Routed question: If participant indicates *Yes*

3.8.*What value on the variables-slider did you end on?

0.015	
0.9	
1 (default)	

3.9.*Indicate the user-friendliness of this causality map, with 1 representing least user-friendly and 4 representing most user-friendly.

1	2	3	4

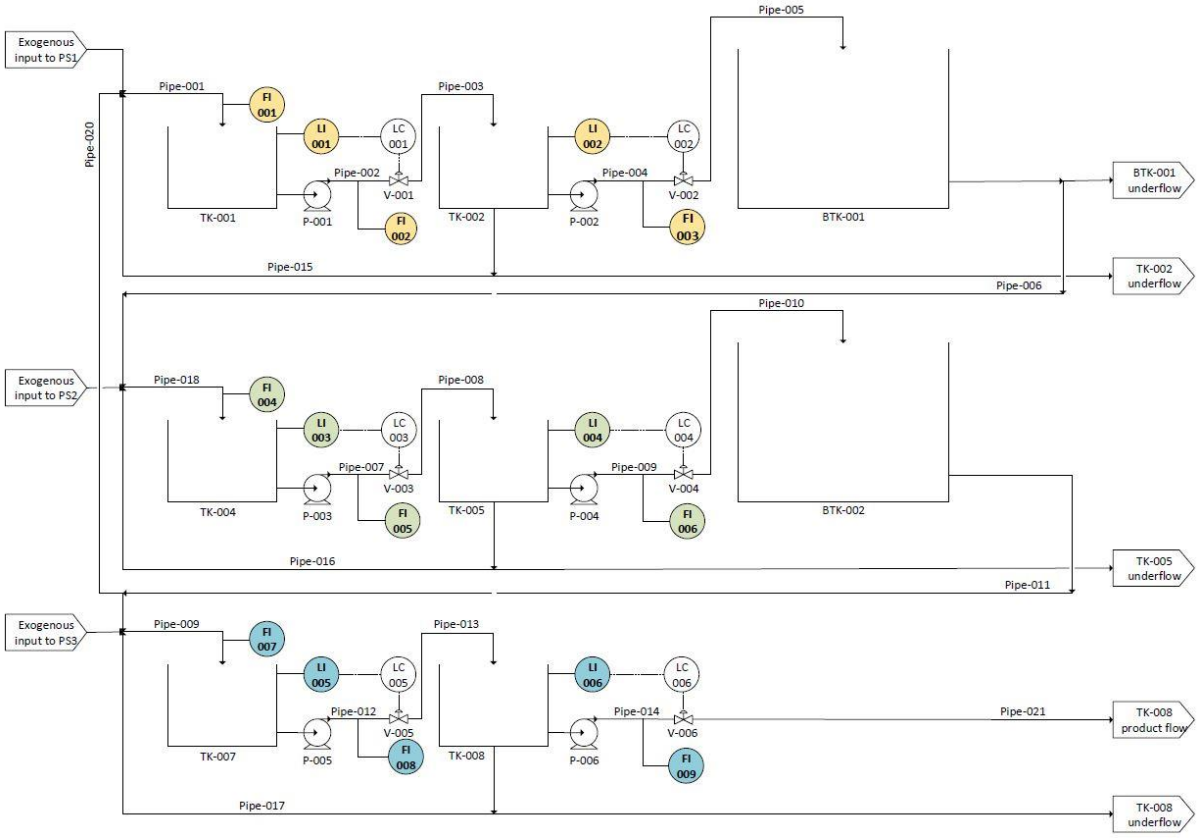
--- PAGE BREAK ---

HIERARCHICAL APPROACH 1: Variables grouped according to plant sections, with principal components as representatives

The causality maps presented in this survey are constructed for the process shown below. The plant consists of three plant sections (PS1, PS2, PS3), which each have:

- Two tanks in series, where each tank has feedback level control.
- An exogenous input to the first tank.
- A recycle stream from the second tank's underflow to the first tank.

The plant sections are separated by buffer tanks, with a portion of each buffer tank's underflow fed to the following plant section. In addition, there is a recycle stream from the buffer tank following PS2 back to PS1.



2.1.*Click on the following link to open a presentation of the causality map(s)[3]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

This causality map consists of nodes that each represent groups of variables, and it should be used to trace the fault to one of those groups of variables. In this case, the groups of variables correspond to variables belonging to the same plant sections, and the representative for each group of variables is the first principal component, as

[3] Refer to Section H.2 of this appendix for the causality maps.

determined by applying principal component analysis (PCA) to the time series data of those variables.

https://docs.google.com/presentation/d/e/2PACX-1vT7uhJaJ9VzO6A1ylif7lyrJaaguz5NQeTWCdvgd0sAdDsS6T6mnCxlh5Wt77fyXEJacKBXJaJtq0wq/pub?start=false&loop=false&delayms=3000

2.2.*In which group of variables is the root cause located?

G1 (PS1)	
G2 (PS2)	
G3 (PS3)	

2.3.*Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 4 indicating the smallest role.

Node colour	
Edge weight	
Edge lines (solid/dashed)	
Number of arrows pointing towards/away from node(s)	

2.4.*Did you use the variables-slider for this causality map?

Yes	No

Routed question: If participant indicates *Yes*

2.5.*What value on the variables-slider did you end on?

1 (default)	
0.9	
0.015	

2.6.*Indicate the user-friendliness of this causality map, with 1 representing least user-friendly and 4 representing most user-friendly.

1	2	3	4

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *1 (default)* in 2.5

2.7.*Click on the following link to open a presentation of the causality map(s)[4]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

[4] Refer to Section H.2 of this appendix for the causality maps.

This causality map consists of nodes that represent each of the variables in the group you identified in the previous causality map, and it should be used to trace the root cause to a specific variable.

https://docs.google.com/presentation/d/e/2PACX-1vTEWfTuCXgalXvYC5uFe8oEFKm3VZMslGQ4o9HsRg-lY5IXDMIm3M_p4GDLNw-qHKfVMlxw5AyUMLmu/pub?start=false&loop=false&delayms=3000

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *1 (default)* in 2.5

2.8.*Which variable is the root cause?

LI-003	
LI-004	
FI-004	
FI-005	
FI-006	
I can't decide between multiple variables.	

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *1 (default)* in 2.5

Routed question: If participant indicates *I can't decide between multiple variables.*

2.9.Which variables can you not decide between?

LI-003	
LI-004	
FI-004	
FI-005	
FI-006	

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *1 (default)* in 2.5

2.10. *Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 3 indicating the smallest role.

Node colour	
Edge weight	
Number of arrows pointing towards/away from node(s)	

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *0.9* in 2.5

2.11. *Click on the following link to open a presentation of the causality map(s)[5]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

[5] Refer to Section H.2 of this appendix for the causality maps.

This causality map consists of nodes that represent each of the variables in the group you identified in the previous causality map, and it should be used to trace the root cause to a specific variable.

https://docs.google.com/presentation/d/e/2PACX-1vQit5O2sr-Xt8xXWWChSBdhjyJ0bd-zbUnI1P_hMnqZTYj6fCOgTrbOK0OxSUh7ec7DbmnamFtS5LQ3/pub?start=false&loop=false&delayms=3000

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *0.9* in 2.5

2.12. *Which variable is the root cause?

LI-003	
LI-004	
FI-006	
I can't decide between multiple variables.	

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *0.9* in 2.5

Routed question: If participant indicates *I can't decide between multiple variables.*

2.13. Which variables can you not decide between?

LI-003	
LI-004	
FI-006	

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *0.9* in 2.5

2.14. *Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 3 indicating the smallest role.

Node colour	
Edge weight	
Number/weight of arrows pointing towards/away from node(s)	

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *0.015* in 2.5

2.15. *Click on the following link to open a presentation of the causality map(s)[6]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

This causality map consists of nodes that represent each of the variables in the group you identified in the previous causality map, and it should be used to trace the root cause to a specific variable.

[6] Refer to Section H.2 of this appendix for the causality maps.

https://docs.google.com/presentation/d/e/2PACX-1vQit5O2sr-Xt8xXWWChSBdhjyJ0bd-zbUnI1P_hMnqZTYj6fCOgTrbOK0OxSUh7ec7DbmnamFtS5LQ3/pub?start=false&loop=false&delayms=3000

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *0.015* in 2.5

2.16. *Which variable is the root cause?

LI-003	
LI-004	
FI-006	
I can't decide between multiple variables.	

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *0.015* in 2.5

Routed question: If participant indicates *I can't decide between multiple variables.*

2.17. Which variables can you not decide between?

LI-003	
LI-004	
FI-006	

Routed question: If participant indicates *G2 (PS2)* in 2.2 and *0.015* in 2.5

2.18. *Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 3 indicating the smallest role.

Node colour	
Edge weight	
Number/weight of arrows pointing towards/away from node(s)	

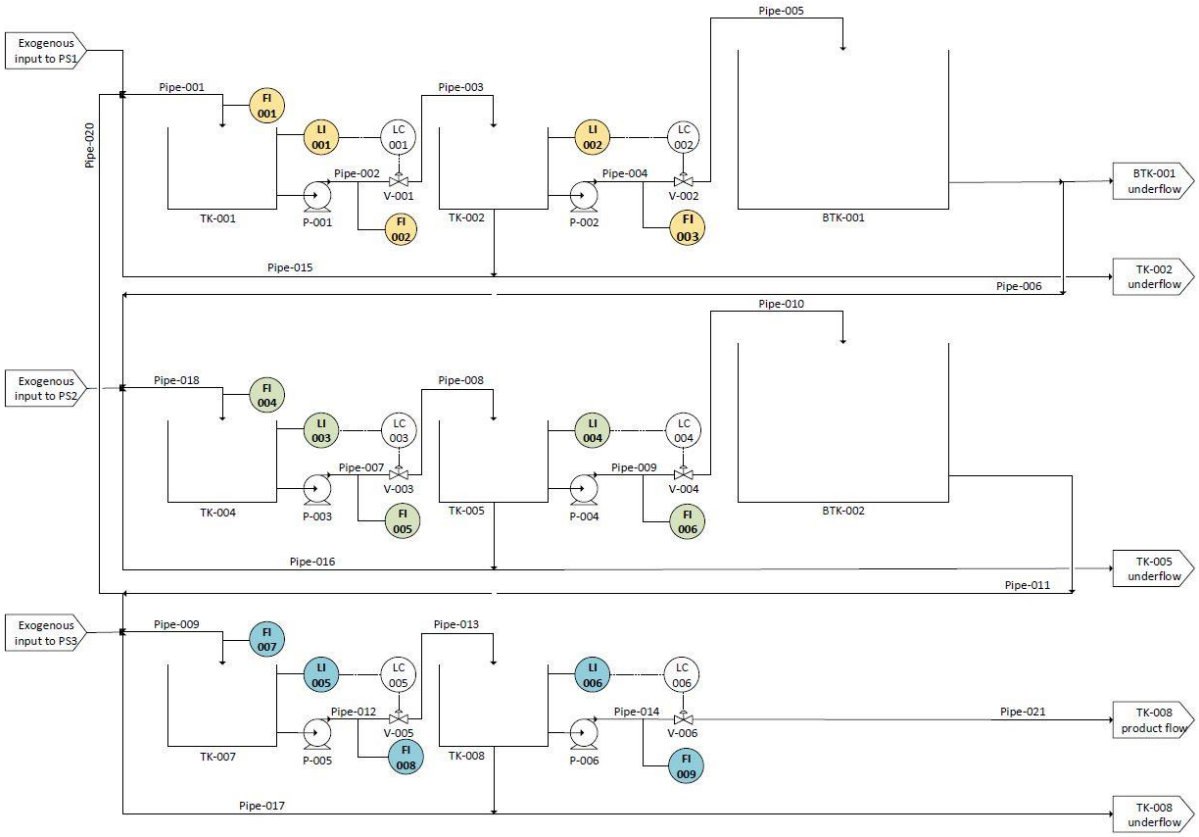
--- PAGE BREAK ---

HIERARCHICAL APPROACH 2: Variables grouped according to modules in the data, with principal components as representatives

The causality maps presented in this survey are constructed for the process shown below. The plant consists of three plant sections (PS1, PS2, PS3), which each have:

- Two tanks in series, where each tank has feedback level control.
- An exogenous input to the first tank.
- A recycle stream from the second tank's underflow to the first tank.

The plant sections are separated by buffer tanks, with a portion of each buffer tank's underflow fed to the following plant section. In addition, there is a recycle stream from the buffer tank following PS2 back to PS1.



3.1.*Click on the following link to open a presentation of the causality map(s)[7]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

This causality map consists of nodes that each represent groups of variables, and it should be used to trace the fault to one of those groups of variables. In this case, the groups of variables correspond to variables belonging to a module identified in the data by applying a modularisation optimisation technique to the causality map obtained from all the variables; and the representative for each group of variables is the first principal component, as determined by applying principal component analysis (PCA) to the time series data of those variables.

Take note: If you use the variables-slider, the variables belonging to each group change.

https://docs.google.com/presentation/d/e/2PACX-1vQUEgiRE2MzzO3wl81Vm2AzSEHx2QYmk0VTIhXV8P_vgkineOzzywBFckynXwrjIsB-mM58Wl4Yrqay/pub?start=false&loop=false&delayms=3000

3.2.*In which group of variables is the root cause located?

Take note: The group you select corresponds to a specific value for the variables-slider.

G1	
G2	

[7] Refer to Section H.2 of this appendix for the causality maps.

G3	

3.3.*Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 4 indicating the smallest role.

Node colour	
Edge weight	
Edge lines (solid/dashed)	
Number of arrows pointing towards/away from node(s)	

3.4.*Did you use the variables-slider for this causality map?

Yes	No

Routed question: If participant indicates *Yes*

3.5.*What value did you end on?

1 (default)	
0.9	
0.015	

3.6.*Indicate the user-friendliness of this causality map, with 1 representing least user-friendly and 4 representing most user-friendly.

1	2	3	4

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.7.*Click on the following link to open a presentation of the causality map(s)[8]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

This causality map consists of nodes that represent each of the variables in the group you identified in the previous causality map, and it should be used to trace the root cause to a specific variable.

https://docs.google.com/presentation/d/e/2PACX-1vRhkodco-sp8SgXqXwV2d5_LrfAKdvyf7Wxlp3QEQvW0gwOcPlLhwcBhIEf0C3J5Ygr0ed_FPDwz5Qr/pub?start=false&loop=false&delayms=3000

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.8.*Which variable is the root cause?

LI-001	
FI-001	
FI-002	

[8] Refer to Section H.2 of this appendix for the causality maps.

FI-003	
LI-004	
FI-006	
LI-005	
FI-007	
FI-008	
I can't decide between multiple variables.	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

Routed question: If participant indicates *I can't decide between multiple variables.*

3.9.*Which variables can you not decide between?

LI-001	
FI-001	
FI-002	
FI-003	
LI-004	
FI-006	
LI-005	
FI-007	
FI-008	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.10. *Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 3 indicating the smallest role.

Node colour	
Edge weight	
Number of arrows pointing towards/away from node(s)	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.11. *Did you use the connections-slider for this causality map?

Yes	No

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

Routed question: If participant indicates *Yes*

3.12. *What value did you end on?

1e-16	
0.01 (default)	
0.05	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.13. *Click on the following link to open a presentation of the causality map(s)[9]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

This causality map consists of nodes that represent each of the variables in the group you identified in the previous causality map, and it should be used to trace the root cause to a specific variable.

https://docs.google.com/presentation/d/e/2PACX-1vSkgW4sk18hE5TIylMUTEjDP1R9kltd3DKXTyHihjGmXhktAurJfiC10p990jHMv4YL1KXd50w7bjcj/pub?start=false&loop=false&delayms=3000

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.14. *Which variable is the root cause?

LI-002	
FI-003	
LI-003	
LI-004	
FI-006	
I can't decide between multiple variables.	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

Routed question: If participant indicates *I can't decide between multiple variables.*

3.15. Which variables can you not decide between?

LI-002	
FI-003	
LI-003	
LI-004	
FI-006	
I can't decide between multiple variables.	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

[9] Refer to Section H.2 of this appendix for the causality maps.

3.16. *Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 3 indicating the smallest role.

Node colour	
Edge weight	
Number of arrows pointing towards/away from node(s)	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.17. *Did you use the connections-slider for this causality map?

Yes	No

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

Routed question: If participant indicates *Yes*

3.18. *What value did you end on?

1e-16	
0.01 (default)	
0.05	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.19. *Click on the following link to open a presentation of the causality map(s)[10]. There are hyperlinks to represent the sliders – you may interact with them if you wish to do so.

This causality map consists of nodes that represent each of the variables in the group you identified in the previous causality map, and it should be used to trace the root cause to a specific variable.

https://docs.google.com/presentation/d/e/2PACX-1vRrEm-kYKJOQQ-M9BRnQj793NWitWNcXMdHVGjcjo-M96R6HAZ030qVKWp6mx8w54W22DAYQMtocABB/pub?start=false&loop=false&delayms=3000

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.20. *Which variable is the root cause?

LI-001	
LI-002	
LI-003	
LI-004	
FI-006	

[10] Refer to Section H.2 of this appendix for the causality maps.

I can't decide between multiple variables.	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

Routed question: If participant indicates *I can't decide between multiple variables.*

3.21. Which variables can you not decide between?

LI-001	
LI-002	
LI-003	
LI-004	
FI-006	

Routed question: If participant indicates *G2* in 3.2 and *1 (default)* in 3.5

3.22. *Rank the extent to which the following features played a role in your decision-making while identifying the root cause, with 1 indicating the largest role and 4 indicating the smallest role.

Node colour	
Edge weight	
Number of arrows pointing towards/away from node(s)	

--- PAGE BREAK ---

BACKGROUND INFORMATION

4.1.*Indicate which of the following is most applicable to you.

Final-year/Postgraduate Chemical Engineering student at Stellenbosch University	
Employee of Anglo American Platinum	
Employee of an IoT company in the mineral processing industry in the Western Cape	

4.2.*Indicate your highest level of education

National Senior Certificate (Matric) or equivalent	
Bachelor's degree	
Master's degree	
PhD	

4.3.*Have you had any experience with causality analysis prior to taking this survey?

Yes, extensive	
Yes, limited	

No	

4.4.*Do you think the base case approach (plant-wide causality map with transitive reduction) can be practically applied in industry to provide actionable insights? Motivate your answer.

4.5.*Do you think the first hierarchical approach (variables grouped according to plant sections, with principal components as representatives) can be practically applied in industry to provide actionable insights? Motivate your answer.

4.6.*Do you think the second hierarchical approach (Variables grouped according to modules in the data, with principal components as representatives) can be practically applied in industry to provide actionable insights? Motivate your answer.

--- PAGE BREAK ---

LUCKY DRAW

5.1.If you wish to participate in a lucky draw to win a Woolworths voucher valued at R 500, provide your email address below. Note that your email address will only be used to determine and contact the winner.

H.2. Survey causality maps

In the survey sent out to participants, the participants were directed to a PowerPoint presentation, where hyperlinks were used to represent the connections-slider and variables-slider. This section provides all versions of the causality maps according to the sliders, corresponding to the question numbers in the questionnaire (see Section H.1 of this appendix). If the connections-slider or variables-slider does not have an effect on the causality map, the respective slider was not included for that specific causality map.

H.2.1. Causality maps in Question 1.1

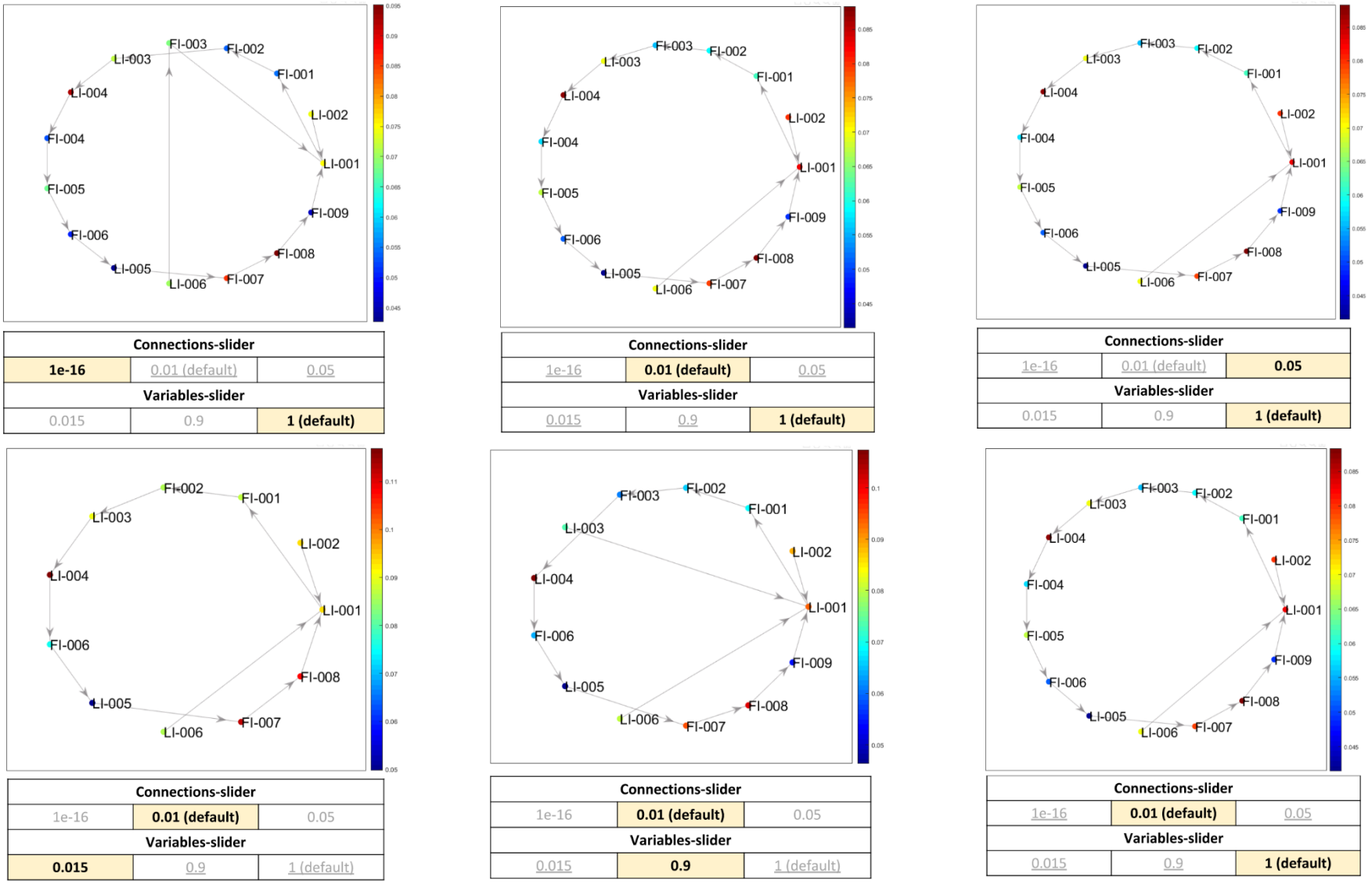
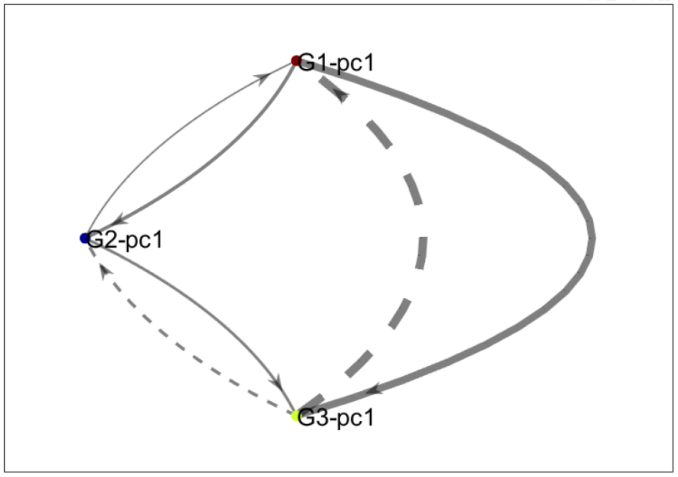


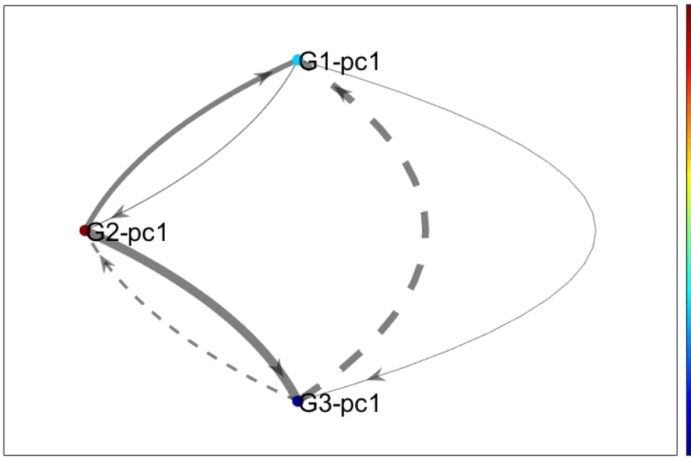
Figure 49: Causality maps for question 1.1 in the questionnaire. Top row, Connections-slider. Bottom row, Variables-slider.

H.2.2. Causality maps in Question 2.1



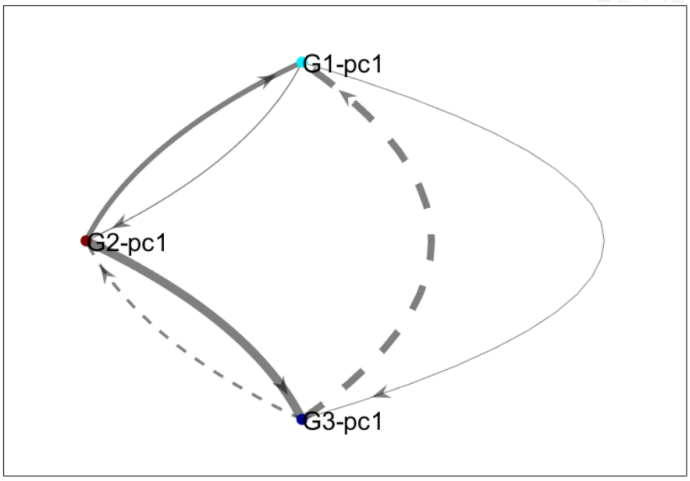
G1	LI-001	LI-002
G2	LI-003 LI-004	FI-006
G3	LI-005	LI-006

Variables-slider		
0.015	0.9	1 (default)



G1	LI-001 LI-002 FI-001	FI-002 FI-003
G2	LI-003 LI-004	FI-006
G3	LI-005 LI-006 FI-007	FI-008 FI-009

Variables-slider		
0.015	**0.9**	1 (default)



G1	LI-001 LI-002 FI-001	FI-002 FI-003
G2	LI-003 LI-004 FI-004	FI-005 FI-006
G3	LI-005 LI-006 FI-007	FI-008 FI-009

Variables-slider		
0.015	0.9	**1 (default)**

Figure 50: Causality maps for question 2.1 in the questionnaire.

H.2.3. Causality map in Question 2.7

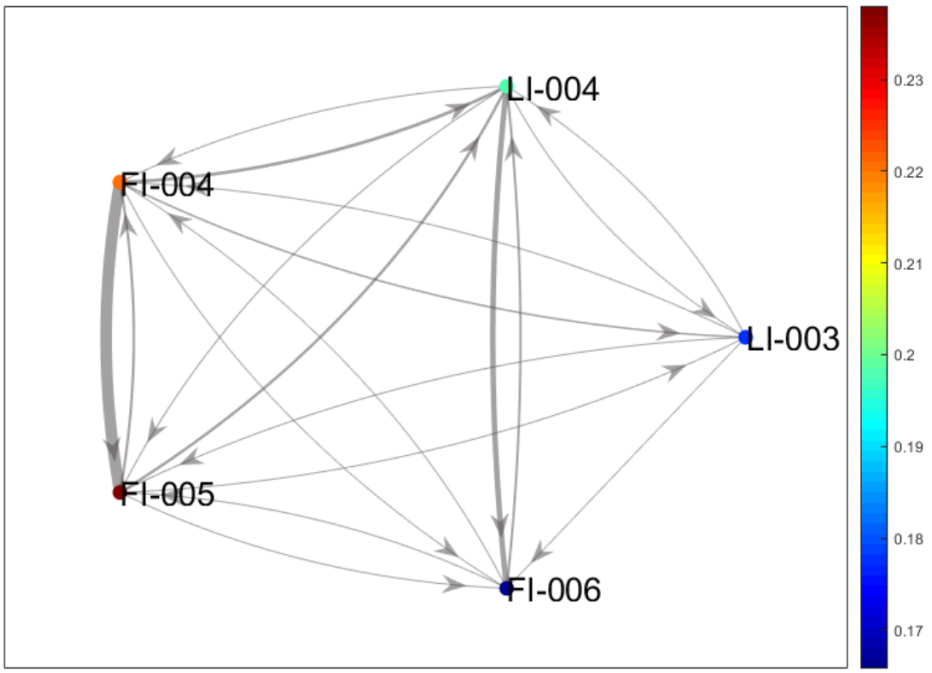


Figure 51: Causality map for question 2.7 in the questionnaire

H.2.4. Causality map in Questions 2.11 and 2.15

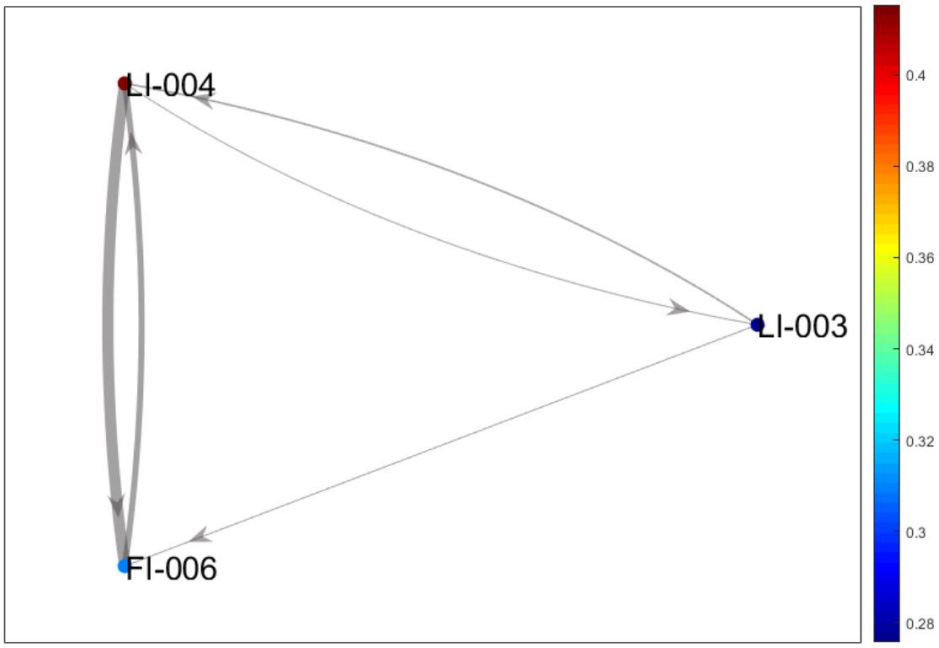
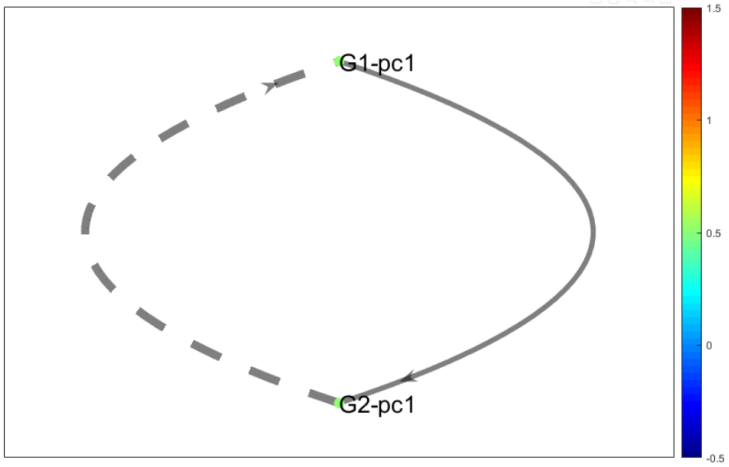


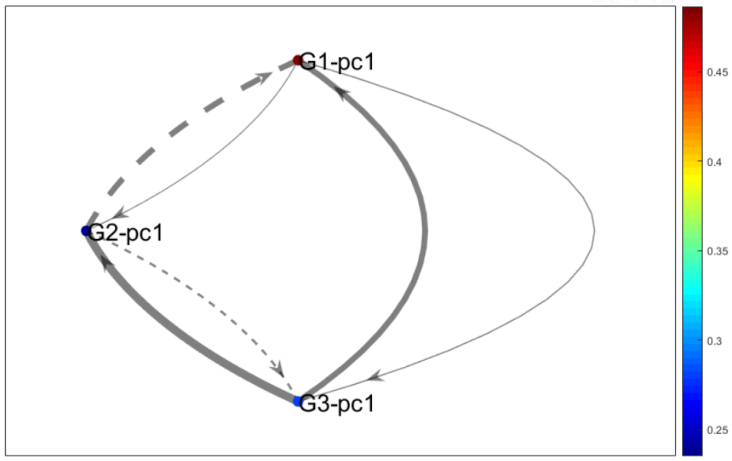
Figure 52: Causality map for questions 2.11 and 2.15 in the questionnaire

H.2.5. Causality maps in Question 31.



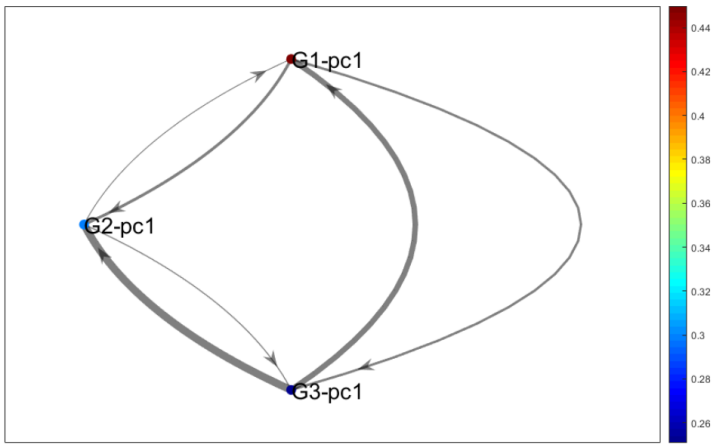
G1	LI-001 LI-002 LI-003	LI-004 FI-006
G2	LI-005	LI-006

Variables-slider		
0.015	0.9	1 (default)



G1	LI-001 FI-001	FI-002
G2	LI-005 LI-006 FI-007	FI-008 FI-009
G3	LI-002 FI-003 LI-003	LI-004 FI-006

Variables-slider		
0.015	**0.9**	1 (default)



G1	LI-003	FI-004	FI-005
G2	LI-001 FI-001 FI-002	FI-003 LI-004 FI-006	LI-005 FI-007 FI-008
G3	LI-002	LI-006	FI-009

Variables-slider		
0.015	0.9	**1 (default)**

Figure 53: Causality maps for question 3.1 in the questionnaire.

H.2.6. Causality maps in Questions 3.7 and 3.13

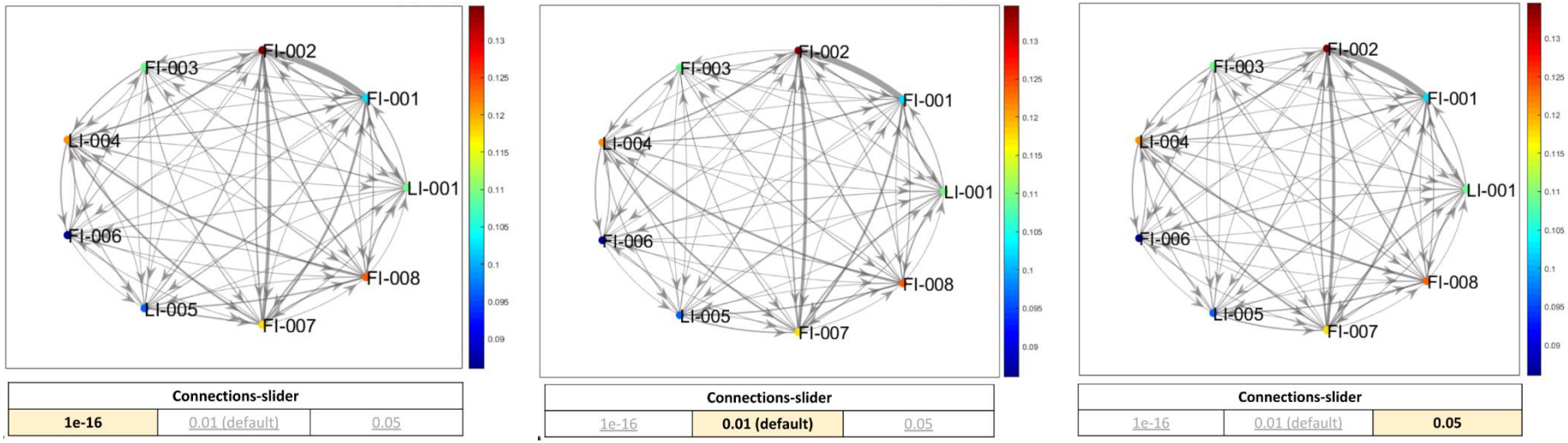


Figure 54: Causality maps for question 3.7 in the questionnaire

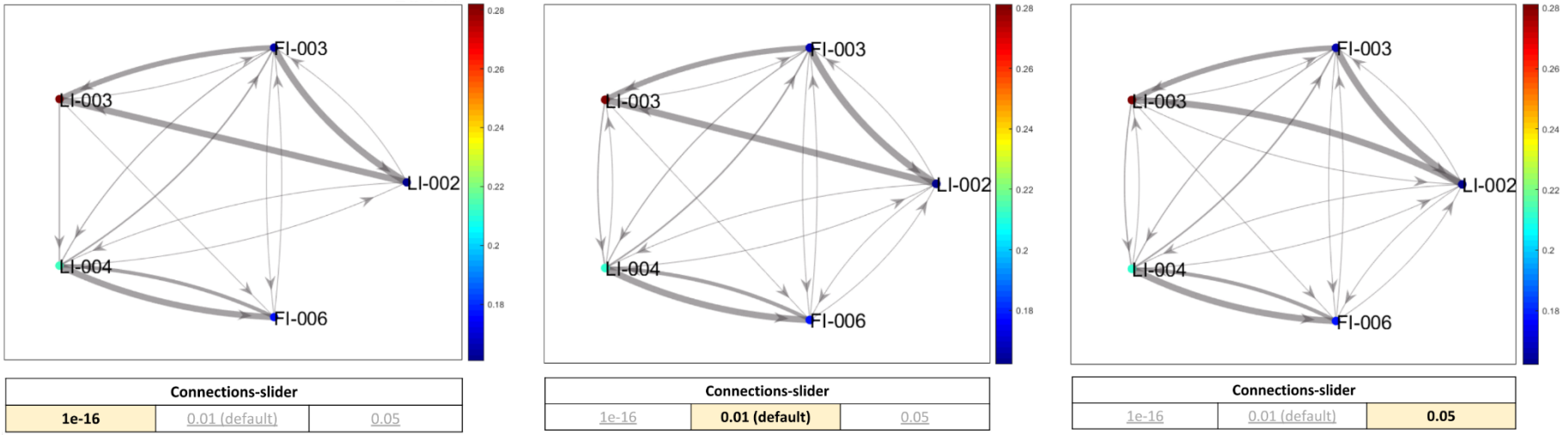


Figure 55: Causality maps for question 3.13 in the questionnaire

H.2.7. Causality maps in Question 3.19

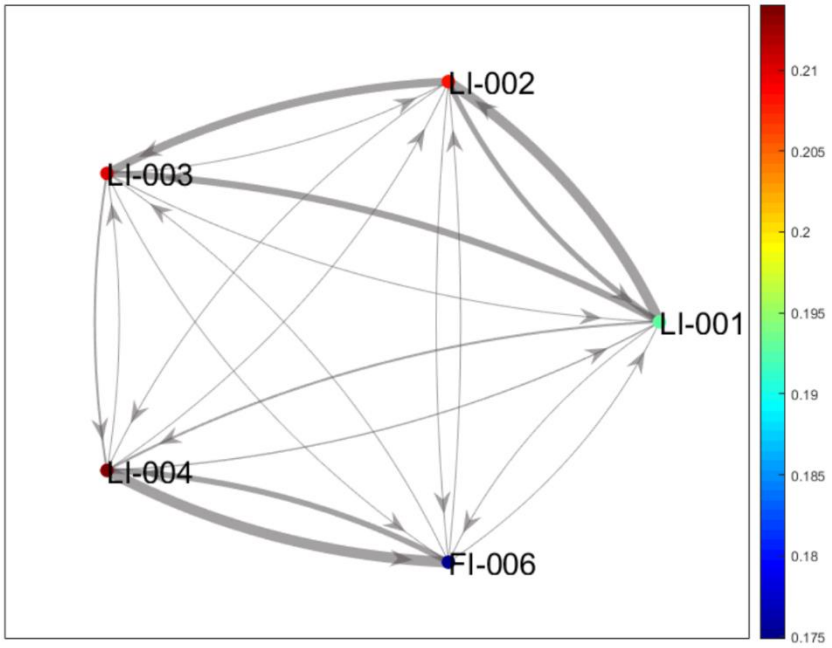


Figure 56: Causality map for question 3.19 in the questionnaire

H.3. Additional survey results

H.3.1.Sliders results

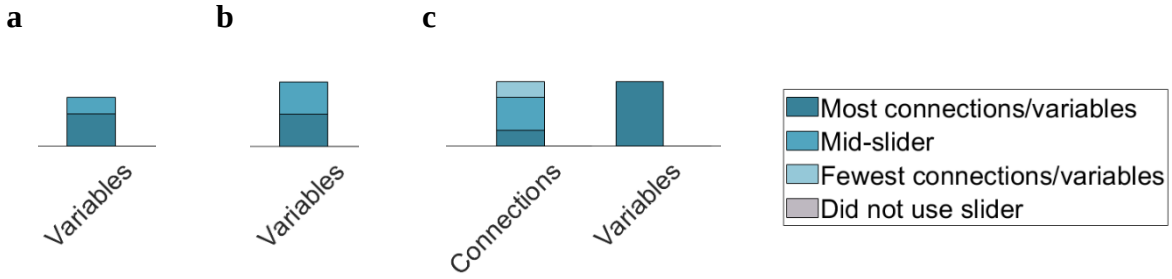


Figure 57: Comparison of the number of **AAP employees** (4 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default.* **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

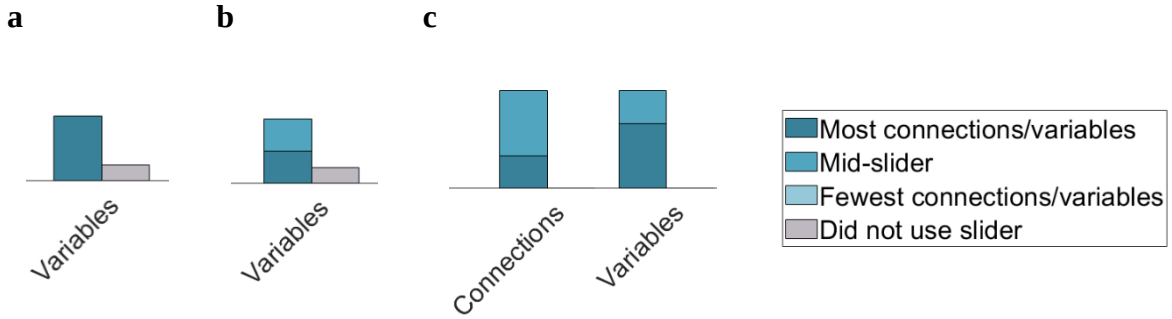


Figure 58: Comparison of the number of **employees of an IoT company in the mineral processing industry in the Western Cape** (6 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default.* **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

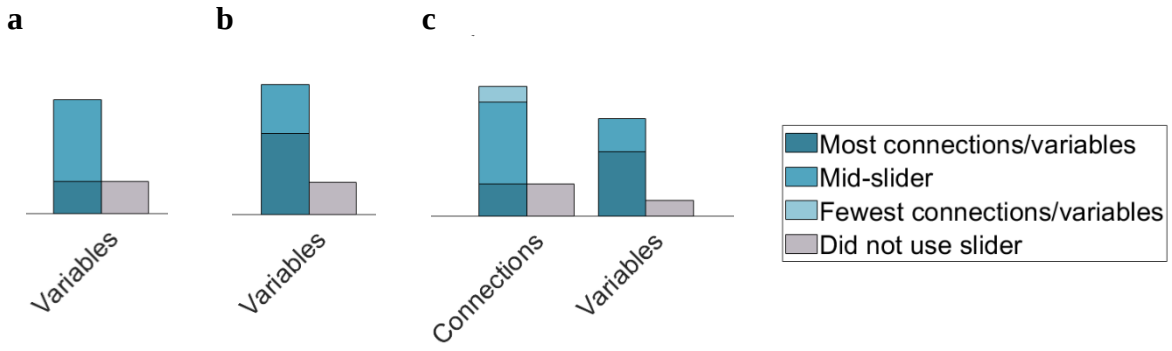


Figure 59: Comparison of the number of **final-year/postgraduate Chemical Engineering students at Stellenbosch University** (10 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default.* **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

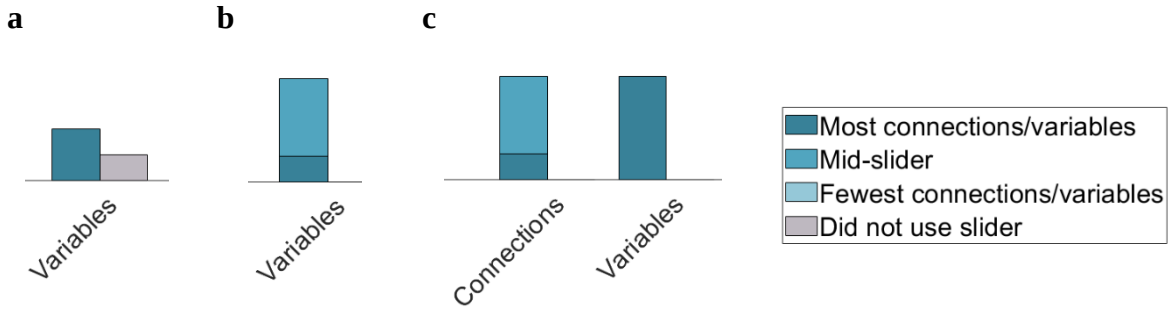


Figure 60: Comparison of the number of participants with **extensive prior experience with causality analysis** (4 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default.* **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

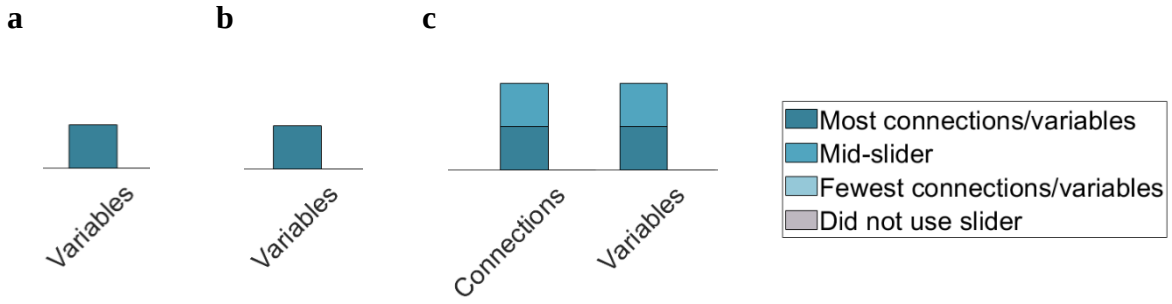


Figure 61: Comparison of the number of participants with **limited prior experience with causality analysis** (2 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default*. **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

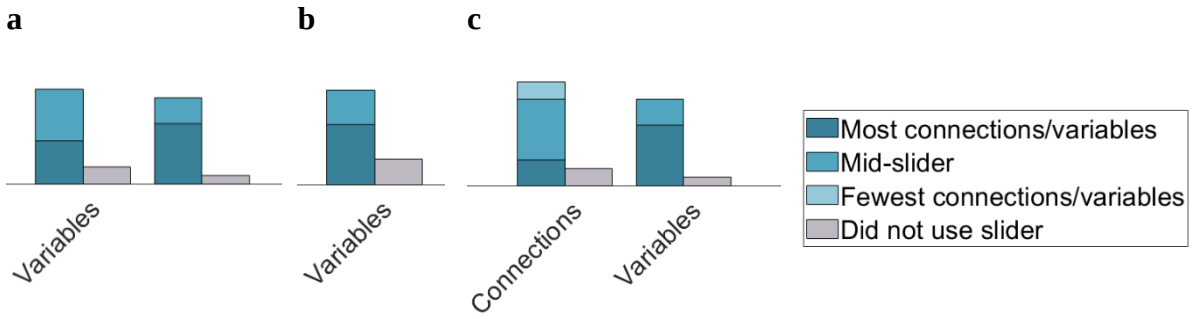


Figure 62: Comparison of the number of participants with **no prior experience with causality analysis** (14 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default*. **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

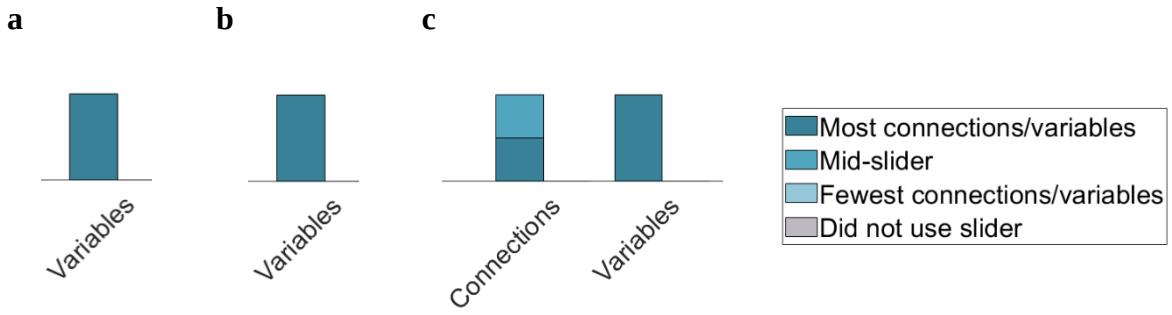


Figure 63: Comparison of the number of participants with a **National Senior Certificate (Matric) or equivalent** (2 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default.* **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

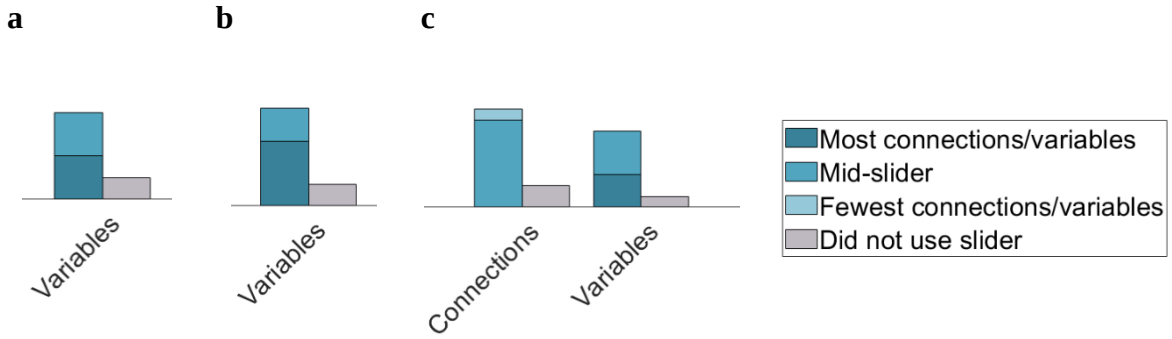


Figure 64: Comparison of the number of participants with a **Bachelor's degree** (11 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default*. **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

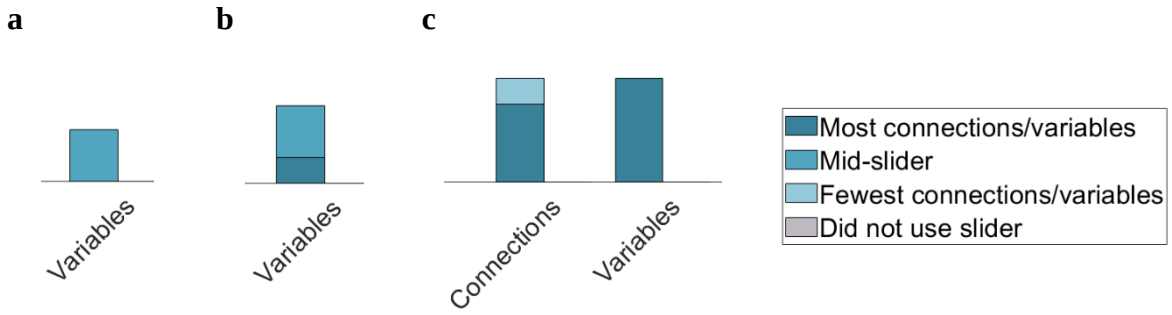


Figure 65: Comparison of the number of participants with a **Master's degree** (4 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default.* **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

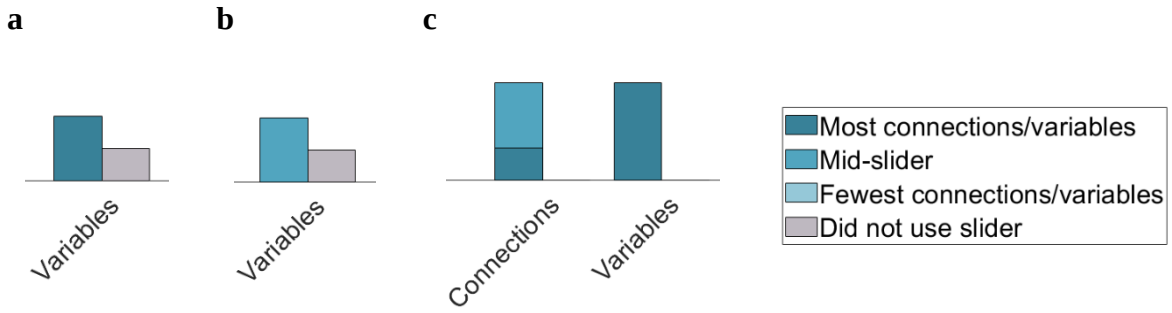


Figure 66: Comparison of the number of participants with a **PhD** (3 participants) who did and did not use the sliders, and which map they used to make their final decision w.r.t. the root cause. *For connections, 'mid-slider' is the default. For variables, 'most variables' is the default.* **a**. PS-PC1. **b**. Mod-PC1. **c**. TR.

H.3.2. Characteristics / tools ranking results

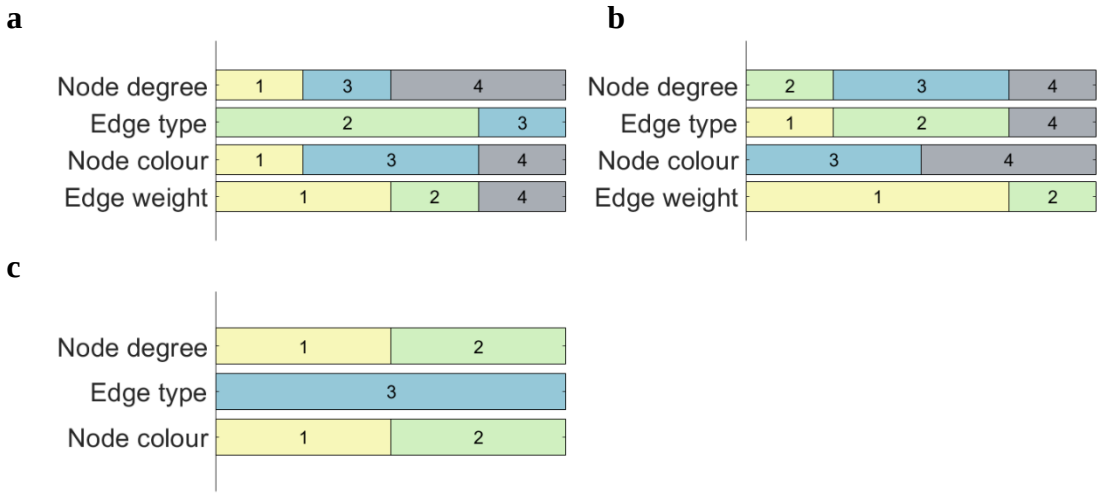


Figure 67: Rankings of the extent to which each characteristic/tools played a role in decision-making of **AAP employees** (4 participants) while identifying a root cause from a causality map. *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

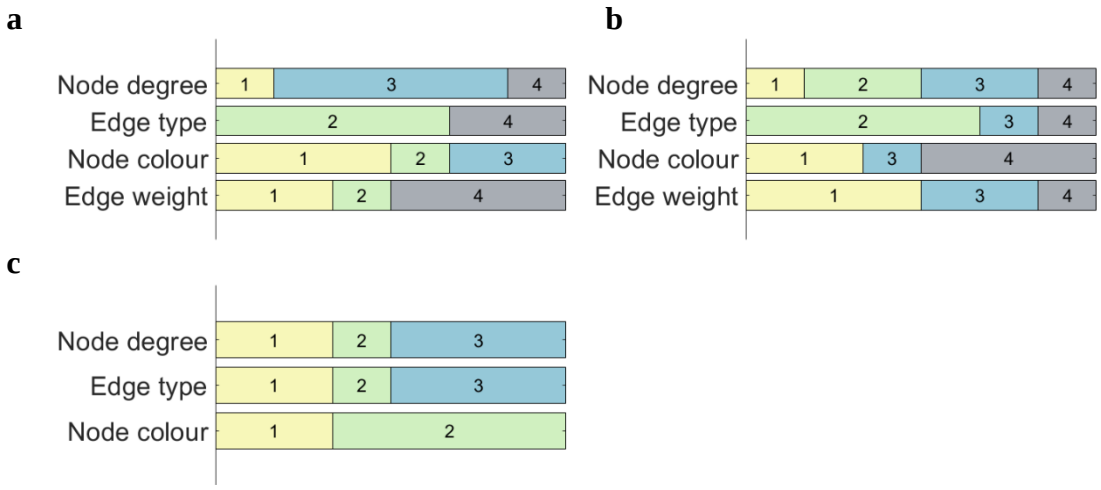


Figure 68: Rankings of the extent to which each characteristic/tools played a role in decision-making of **employees of an IoT company in the mineral processing industry in the Western Cape** (6 participants). *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

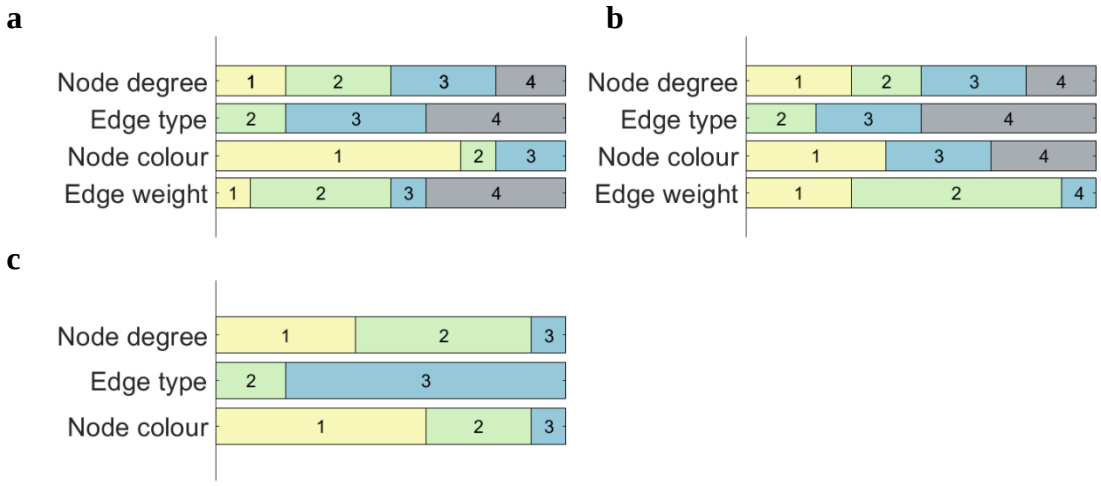


Figure 69: Rankings of the extent to which each characteristic/tools played a role in decision-making of **final-year/postgraduate Chemical Engineering students at Stellenbosch University** (10 participants). *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

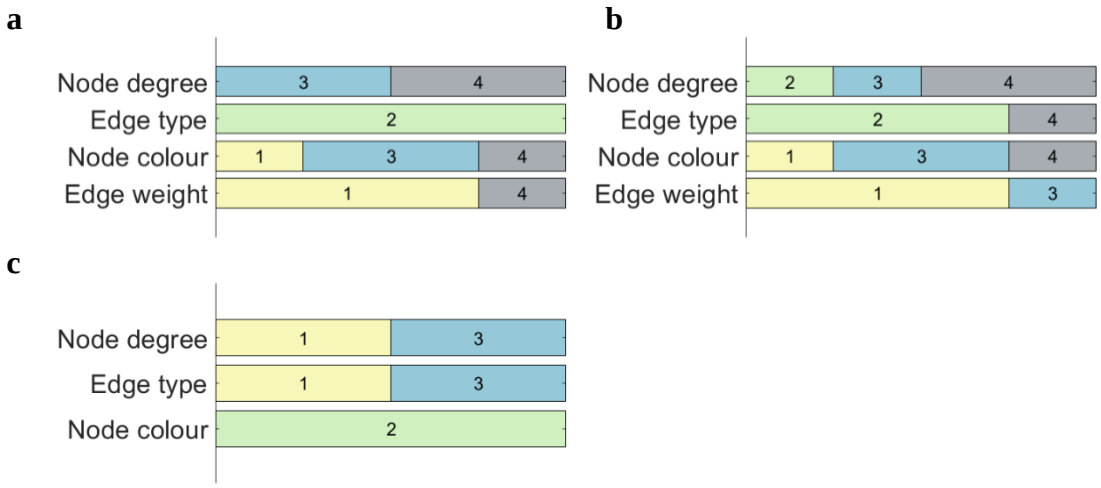


Figure 70: Rankings of the extent to which each characteristic/tools played a role in decision-making of participants with **extensive prior experience with causality analysis** (4 participants)**.** *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

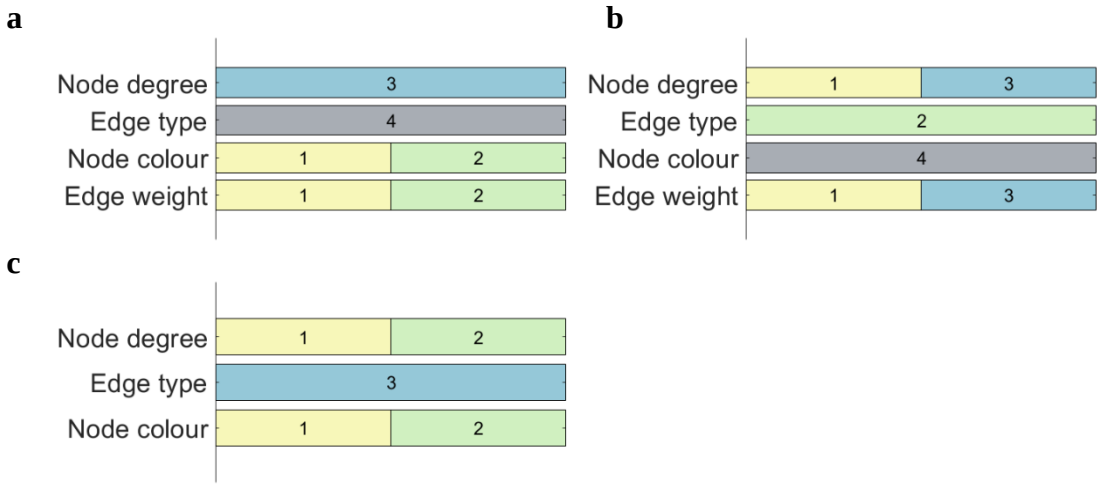


Figure 71: Rankings of the extent to which each characteristic/tools played a role in decision-making of participants with **limited prior experience with causality analysis** (2 participants). *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

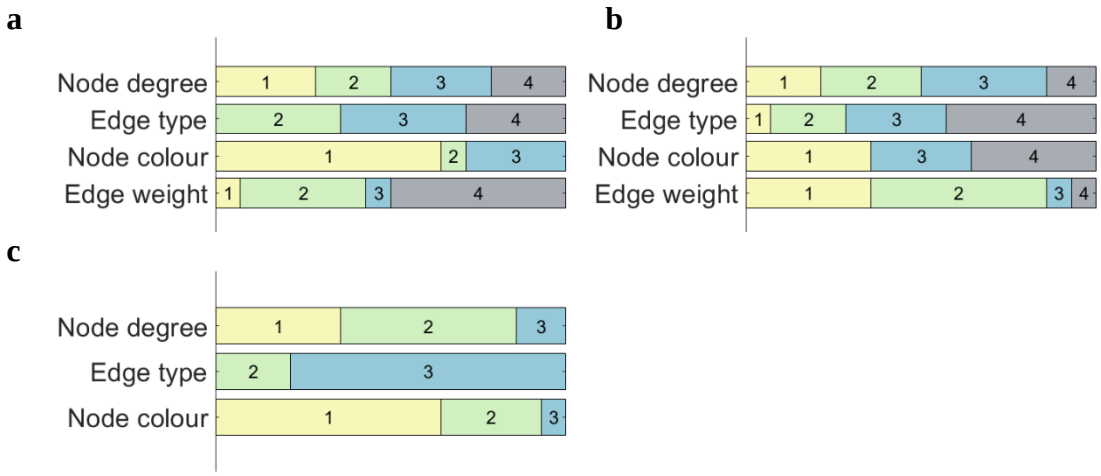


Figure 72: Rankings of the extent to which each characteristic/tools played a role in decision-making of participants with **no prior experience with causality analysis** (14 participants). *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR*

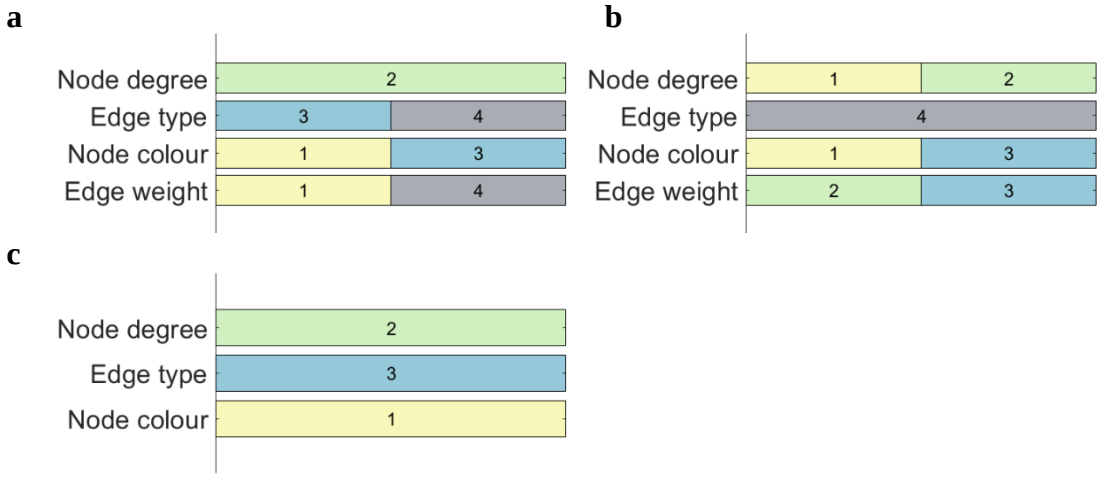


Figure 73: Rankings of the extent to which each characteristic/tools played a role in decision-making of participants with a **National Senior Certificate (Matric) or equivalent** (2 participants)**.** *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

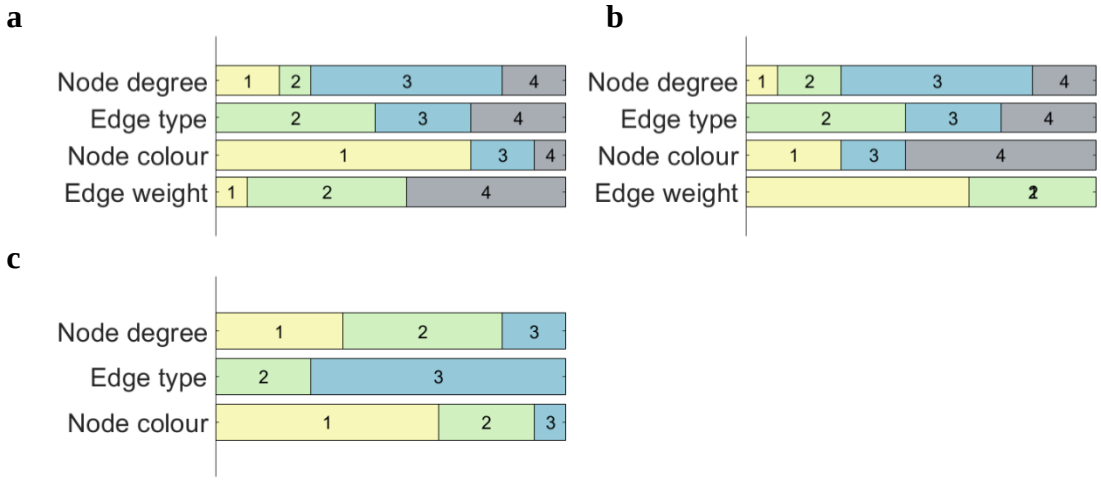


Figure 74: Rankings of the extent to which each characteristic/tools played a role in decision-making of participants with a **Bachelor's degree** (11 participants). *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

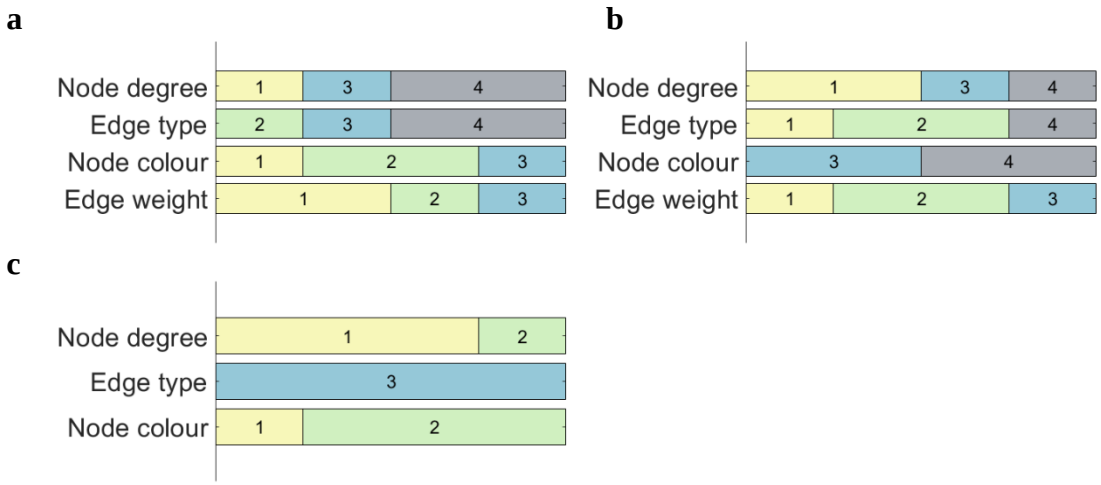


Figure 75: Rankings of the extent to which each characteristic/tools played a role in decision-making of participants with a **Master's degree** (4 participants). *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

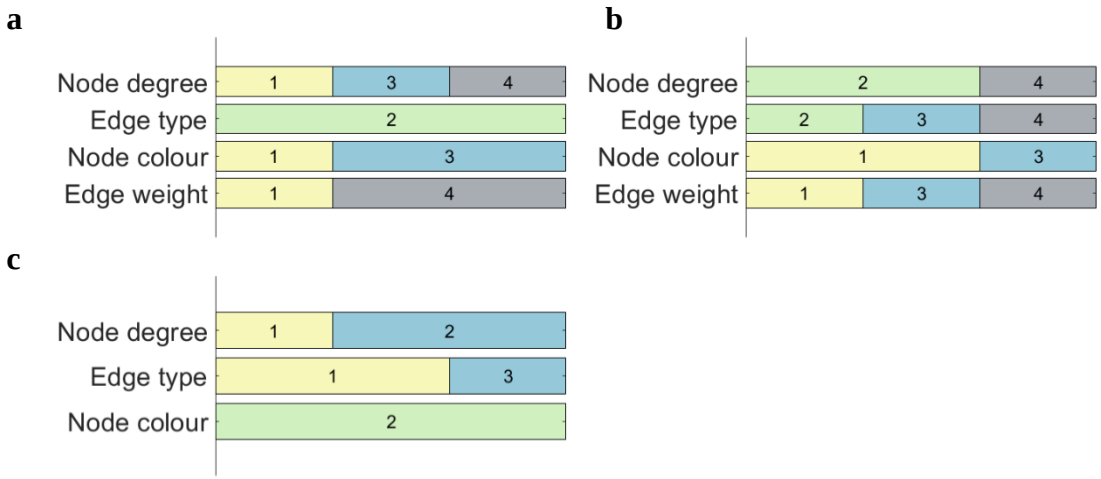


Figure 76: Rankings of the extent to which each characteristic/tools played a role in decision-making of participants with a **PhD** (3 participants). *1 represents the largest role in decision-making, and 4 represents the smallest role in decision-making.* ***a****. PS-PC1.* ***b****. Mod-PC1.* ***c****. TR.*

H.3.3.Summary of thematic analysis results

Table 16: Summary of the lines from participant responses grouped according to codes and subsequently the four themes. For each code, the participant lines are also separated according to whether they were in response to the question regarding TR, PS-PC1, or Mod-PC1.

THEME 1: Process knowledge & CA experience	
Code	**Lines from participant responses**
Need process knowledge	***TR:***
	It would help if one was looking at the values for a process that you know well.
	The user would ideally need process knowledge.
	Having a background of the mineral processing plant further makes it easier to understand why a certain stage would be influenced by the previous stage as to compared to just looking at the causality map alone.
	If one knows the plant well it could be beneficial.
	The grouping by plant section is most understandable if you are an engineer on site and understand the processes.
	PS-PC1:
	If one knows the plant well it could be beneficial
	The grouping by plant section is most understandable if you are an engineer on site and understand the processes.
	Mod-PC1:
	If one knows the plant well it could be beneficial.
	If applied with plant and experience knowledge. Actual plant process knowledge to bring in a reality check.
	I will be reserved to do this as I believe a pure data driven causality map without any intrinsic process knowledge will lead to causal mistakes.
Interpreting causality maps effectively requires prior experience/training	***TR:***
	I don't think it would be practical unless the person has some prior experience with causality graphs.
	The user would ideally need process knowledge AND an understanding of causality maps in order to effectively use causality maps.
	With some extra experience or training the base case approach can be applied easily for a small plant set up
	PS-PC1:
	With some more experience, fault finding can go fairly quickly and provide good insights.
	Mod-PC1:
	I feel like some form of training / experience is needed on my part as it doesn't come as intuitively as the prior examples.
Purely data-based approach avoids process knowledge subjectivity	***Mod-PC1:***
	As the more data-based approach, it is less susceptible to subjective interpretation.
	I feel this method would be most applicable due to the objectivity that data modules would provide into the root cause interpretation as knowledge about the process will not bias the investigator's view.

THEME 2: Grouping variables	
Code	**Lines from participant responses**
Fewer nodes improve interpretability	***TR:***
	Although some form of condensing the variable pool may be required for excessively large plants where more complex system interactions occur. For smaller plants the base case approach can be adequate, but once the variables start to become too high, it may become difficult to interpret.
	PS-PC1:
	Fewer variables allows for easier interpretation.
	The reduced causal map for a few variables is more manageable and understandable.
	Mod-PC1:
	The lesser number of dots also makes the interpretation simpler.
Grouping according to plant sections makes sense	***PS-PC1:***
	Grouping according to plant section made sense and remained stable.
	Grouping by plant layout just seemed more intuitive.
	When group according to plant sections, a certain degree of plant knowledge is introduced.
Grouping variables simplified root cause analysis	***PS-PC1:***
	The idea of grouping variables will likely improve/simplify the root cause analysis workflow.
Data-based modules can contain many variables	***Mod-PC1:***
	The sheer number of variables in certain modules can make identifying problem sections challenging.
Data-based grouping adds confusion to user	***Mod-PC1:***
	Unable to comment. Not familiar with modularisation
	No, the groupings did not quite make sense
Transitive reduction could be applied to plant-wide maps in the hierarchical approaches	***PS-PC1:***
	This seems the most promising, but hopefully transitive reduction could make it even easier to use?

THEME 3: Accuracy during simplification	
Code	**Lines from participant responses**
Accuracy must be maintained during map simplification	***PS-PC1:***
	Anything that simplifies the causality graphs is good (assuming it stays correct).
PCA may lose necessary info	***PS-PC1:***
	The PCA reduces the overall information content of a system, potentially suppressing propogation of certain variables' effects.
	The first PC captures bulk of variation - some important information contributing to a variation may be lost in subsequent PCs.
Uncertainty must be quantified	***TR, PS-PC1:, Mod-PC1:***
	you need to quantify your uncertainty with regards to the edge weights, connection strengths, etc. This is going to allow you to make more informed decisions about where the root cause is comping from. You are using noise-corrupted data to infer connections and this can dilute the 'information' you are trying to deduce. Noise is a real problem and you need to find a principled way or addressing it. Your results might be more conclusive if you have use noise-free data, but this isn't a realistic assumption in practice.
THEME 4: Sliders	
Code	**Lines from participant responses**
Maps changing with sliders causes confusion	***TR:***
	Interpretation of the maps changes with your sliders - seems clear on one and less on another.
	The effect of changing the connection and variables could just add more confusion unless one has had a more experiance with the causality map.
	The fact that there were so many contradicting results dependent on the connections and variable sliders does raise cause for concern.
	There was already some confusion on my part on the connection sliders vs variable sliders. Depending on the value of the slider the apparent fault changes.
	There are too many things to consider as the two slider options create multiple permutations of possible settings. This makes it difficult to know which setting is the most representative of what is to be determined.
	PS-PC1:
	The single slider makes it easier to identify and interpret the consequences of changing slider values. This makes the interpretation of certain settings easier and more logical as there is not interaction between the sliders.
	Mod-PC1:
	The changing of groups of variables, depending on the value selected in the slider, made this even more complex.
	Having the groupings change with the slider just added more permutaions and the answer again felt like guess work.
	The connections and variables slider are counter-intuitive.
	Rather than showing single plot with interactive selections; it is easier to show a matrix of plots - the user would not need to remember what happened in the previous selection in order to draw a conclusion.
	This approach was more confusing than the rest as variables changed according to how they are grouped.

H.4. Notice of ethics approval

UNIVERSITEIT
STELLENBOSCH
UNIVERSITY

NOTICE OF APPROVAL

REC: Social, Behavioural and Education Research (SBER) - Initial Application Form

28 May 2020

Project number: 14385

Project Title: Improving the interpretability of causality maps for fault identification

Dear Miss Natali Van Zijl

Your response to stipulations submitted on 27 May 2020 was reviewed and approved by the REC: Social, Behavioural and Education Research (REC: SBE).

Please note below expiration date of this approved submission:

Ethics approval period:

Protocol approval date (Humanities)	Protocol expiration date (Humanities)
25 February 2020	24 February 2023

SUSPENSION OF PHYSICAL CONTACT RESEARCH DURING THE COVID-19 PANDEMIC

Due to the Covid-19 pandemic and resulting lockdown measures, all research activities requiring physical contact or being in undue physical proximity to human participants has been suspended by Stellenbosch University. Please refer to a formal statement issued by the REC: SBE on 20 March for more information on this.

This suspension will remain in force until such time as the social distancing requirements are relaxed by the national authorities to such an extent that in-person data collection from participants will be allowed. This will be confirmed by a new statement from the REC: SBE on the university's dedicated Covid-19 webpage.

Until such time online or virtual data collection activities, individual or group interviews conducted via online meeting or web conferencing tools, such as Skype or Microsoft Teams are strongly encouraged in all SU research environments.

If you are required to amend your research methods due to this suspension, please submit an amendment to the REC: SBE as soon as possible. The instructions on how to submit an amendment to the REC can be found on this webpage: [instructions], or you can contact the REC Helpdesk for instructions on how to submit an amendment: applyethics@sun.ac.za.

GENERAL REC COMMENTS PERTAINING TO THIS PROJECT:

INVESTIGATOR RESPONSIBILITIES

Please take note of the General Investigator Responsibilities attached to this letter. You may commence with your research after complying fully with these guidelines.

If the researcher deviates in any way from the proposal approved by the REC: SBE, the researcher must notify the REC of these changes.

Please use your SU project number (14385) on any documents or correspondence with the REC concerning your project.

Please note that the REC has the prerogative and authority to ask further questions, seek additional information, require further modifications, or monitor the conduct of your research and the consent process.

CONTINUATION OF PROJECTS AFTER REC APPROVAL PERIOD

You are required to submit a progress report to the REC: SBE before the approval period has expired if a continuation of ethics approval is required. The Committee will then consider the continuation of the project for a further year (if necessary).

Once you have completed your research, you are required to submit a final report to the REC: SBE for review.

Included Documents:

Document Type	File Name	Date	Version
Research Protocol/Proposal	Research Proposal_Improving the interpretability of causality analysis for fault identification	10/02/2020	1
Budget	Budget - Improving the intepretability of causality analysis for fault identification.pdf	10/02/2020	1
Informed Consent Form	Informed Consent	17/02/2020	2
Data collection tool	SURVEY_Improving the interpretability of causality maps for fault identification_v2	17/02/2020	2
Proof of permission	Stone Three Institutional Permission Letter 20200217	27/05/2020	1
Proof of permission	AAP-Permission-Letter	27/05/2020	1
Proof of permission	Institutional Permission_Standard Agreement 1680	27/05/2020	1
Default	Response letter - N van Zijl	27/05/2020	1

If you have any questions or need further help, please contact the REC office at cgraham@sun.ac.za.

Sincerely,

Clarissa Graham

REC Coordinator: Research Ethics Committee: Social, Behavioral and Education Research

National Health Research Ethics Committee (NHREC) registration number: REC-050411-032.
The Research Ethics Committee: Social, Behavioural and Education Research complies with the SA National Health Act No.61 2003 as it pertains to health research. In addition, this committee abides by the ethical norms and principles for research established by the Declaration of Helsinki (2013) and the Department of Health Guidelines for Ethical Research: Principles Structures and Processes (2nd Ed.) 2015. Annually a number of projects may be selected randomly for an external audit.

Principal Investigator Responsibilities

Protection of Human Research Participants

As soon as Research Ethics Committee approval is confirmed by the REC, the principal investigator (PI) is responsible for the following:

Conducting the Research: The PI is responsible for making sure that the research is conducted according to the REC-approved research protocol. The PI is jointly responsible for the conduct of co-investigators and any research staff involved with this research. The PI must ensure that the research is conducted according to the recognised standards of their research field/discipline and according to the principles and standards of ethical research and responsible research conduct.

Participant Enrolment: The PI may not recruit or enrol participants unless the protocol for recruitment is approved by the REC. Recruitment and data collection activities must cease after the expiration date of REC approval. All recruitment materials must be approved by the REC prior to their use.

Informed Consent: The PI is responsible for obtaining and documenting affirmative informed consent using **only** the REC-approved consent documents/process, and for ensuring that no participants are involved in research prior to obtaining their affirmative informed consent. The PI must give all participants copies of the signed informed consent documents, where required. The PI must keep the originals in a secured, REC-approved location for at least five (5) years after the research is complete.

Continuing Review: The REC must review and approve all REC-approved research proposals at intervals appropriate to the degree of risk but not less than once per year. There is **no grace period.** Prior to the date on which the REC approval of the research expires, **it is the PI's responsibility to submit the progress report in a timely fashion to ensure a lapse in REC approval does not occur.** Once REC approval of your research lapses, all research activities must cease, and contact must be made with the REC immediately.

Amendments and Changes: Any planned changes to any aspect of the research (such as research design, procedures, participant population, informed consent document, instruments, surveys or recruiting material, etc.), must be submitted to the REC for review and approval before implementation. Amendments may not be initiated without first obtaining written REC approval. The **only exception** is when it is necessary to eliminate apparent immediate hazards to participants and the REC should be immediately informed of this necessity.

Adverse or Unanticipated Events: Any serious adverse events, participant complaints, and all unanticipated problems that involve risks to participants or others, as well as any research-related injuries, occurring at this institution or at other performance sites must be reported to the REC within **five (5) days** of discovery of the incident. The PI must also report any instances of serious or continuing problems, or non-compliance with the RECs requirements for protecting human research participants.

Research Record Keeping: The PI must keep the following research-related records, at a minimum, in a secure location for a minimum of five years: the REC approved research proposal and all amendments; all informed consent documents; recruiting materials; continuing review reports; adverse or unanticipated events; and all correspondence and approvals from the REC.

Provision of Counselling or emergency support: When a dedicated counsellor or a psychologist provides support to a participant without prior REC review and approval, to the extent permitted by law, such activities will not be recognised as research nor the data used in support of research. Such cases should be indicated in the progress report or final report.

Final reports: When the research is completed (no further participant enrolment, interactions or interventions), the PI must submit a Final Report to the REC to close the study.

On-Site Evaluations, Inspections, or Audits: If the researcher is notified that the research will be reviewed or audited by the sponsor or any other external agency or any internal group, the PI must inform the REC immediately of the impending audit/evaluation.