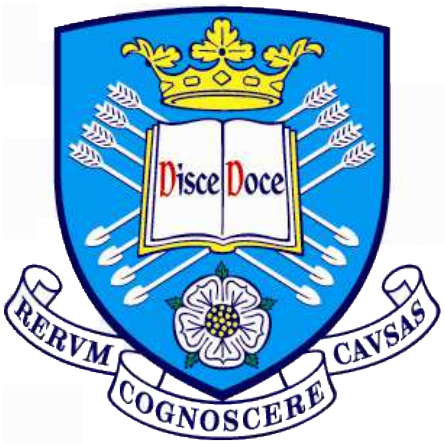
On Nonstationarity from Operational and Environmental Effects in Structural Health Monitoring Bridge Data



A Thesis submitted to the University of Sheffield

for the degree of Doctor of Philosophy in the Faculty of Engineering

by

I. Iakovidis

Department of Mechanical Engineering

University of Sheffield

August 2018

Abstract

Structural Health Monitoring (SHM) describes a set of activities that can be followed in order to collect data from an existent structure, generate data-based information about its current condition, identify the presence of any signs of abnormality and forecast its future response. These activities, among others, include instrumentation, data acquisition, processing, generation of diagnostic tools, as well as transmission of information to engineers, owners and authorities. SHM and, more specifically, continuous monitoring can provide numerous measures, which can be generally classified into three categories; vibrational-based, which includes natural frequencies, modeshapes, damping ratios, component-based, such as strains, tensions, deflections and *environmental and operational variations* (EOVs), associated with temperature, wind, traffic humidity and others.

One of the main technical problems that SHM has to tackle is that of data normalisation. In abstract terms, this describes the impact that EOVs can have on SHM measures. In many cases, with interest placed on bridges, it has been observed that EOVs introduce nonstationary to SHM signals that can mask the variability that can be associated with the presence of damage; making damage detection attempts difficult. Hence, it is desirable to quantify the impacts of EOVs on damage sensitive features, project them out, using methods such as the *cointegration*, *Principal Component Analysis* (PCA) or others, in order to achieve a stationary signal. This type of signal can be assessed over time using tools, such as *statistical process control* (SPC) charts, to identify the existence of novelty, which can be linked with damage.

As one can understand from the latter, it is important to detect the presence of nonstationary in SHM signals and identify its sources. However, this is not a straightforward procedure and one important question that need to be answered is; how

one can judge if a signal is stationary or not. Inside this work, this question is discussed, focusing on the definition of *weak stationarity* and under which assumption this judgement holds. In particular, the data coming from SHM are finite samples. Therefore, the mean and variance of a signal can be tracked, using a sequence of moving windows, something that needs a prior determination of the width of window. However, the major concern here is that the SHM signals can be characterised as *periodically-correlated* or *cyclostationary*. In such cases, it seems that it is better for one to use more advanced statistical tools to assess a signal's nonstationary. More specifically, nonstationary tests coming from the context of Econometrics and time-series analysis can be employed. In order to use such proxies more extensively, one should build trust on their indications by understanding the mechanism under which they perform.

This work concentrates on the *Augmented Dickey-Fuller* (ADF) nonstationary test and emphasis is placed on the hypothesis (*unit root*) under which performs its assessment. In brief, a series of simulations are generated, and based on *dimensional analysis*, it is shown that the ADF test is essentially counts the number of cycles/periods of the dominant periodic component. Its indications depend on the number of observations/cycles, the normalised frequency of the signal, the sampling rate and *signal-to-noise ratio* (SNR). The most important conclusion made is that knowing the sampling frequency of any given signal, a critical frequency in Hz can be found, which can be derived from the critical normalised one, as a function of the number of cycles, which can be directly used to judge if the signal is stationary or not. In other words, this investigation provides an answer to the question; after how many cycles of continuous monitoring (i.e. days), an SHM signal can be judged as stationary?

As well as considering nonstationary in a general way, this thesis returns to the main issue of data normalisation. To begin with, a laboratory test is performed, at the laboratory (Jonas lab) of Sheffield University, on an aluminium truss bridge model manufactured there. In particular, that involved vibration analysis of the truss bridge inside an environmental chamber, which simulated varying temperature conditions from -10 to 20 ^oC, while damage introduced on the structure by the removal of bolts and connecting brackets in two locations of the model. This experiment provided interesting results to discuss further the impact of EOVs on data coming from the monitoring of a small-scale structure. After that, the thesis discusses the use of *Johansen's approach to cointegration* in the context of SHM,

demonstrate its use on the laboratory truss bridge data and provides a review of the available methods that can be used to monitor the *cointegration residual*. The latter is the stationary signal provided by cointegration which is free from EOVs and capable for *novelty detection*. The methodologies reviewed are various SPC charts, while also the use of ADF is also explored, providing extensive discussion.

Furthermore, an important conclusion from the SHM literature is that the impact of EOVs on SHM signals can occur on widely disparate time scales. Therefore, the quantification and elimination of these impacts from signals is not an easy procedure and prior knowledge is needed. For such purposes, refined means originated from the field of signal processing can be used within SHM. Of particular interest here is the concept of *multiresolution analysis* (MRA), which has been used in SHM in order to decompose a given signal in its frequency components (different time-scales) and evaluate the damage sensitivity of each one, employing the *Johansen's approach to cointegration*, which is able to project out the impact of EOVs from multiple SHM series.

A more principled way to perform MRA is proposed here, in order to decompose SHM signals, by introducing two additional steps. The first step is the ADF test, which can be used to assess each one of the MRA levels in terms of nonstationary. In this way, a critical decomposition level (L*) can be found and used to decompose the original SHM signal into a non-stationary and stationary part. The second step introduced includes the use of *autocorrelation functions* (ACFs) in order to test the stationary MRA levels and identify those that can be considered as *delta-correlated*. These levels can be used to form a noisy component inside the stationary one. Assuming that all the aforementioned steps are confirmed, the original signal can now be decomposed into a stationary, a mean, a non-stationary and a noisy component. The proposed decomposition can be of great interest not only for SHM purposes, but also in the general context of time-series analysis, as it provides a principled way to perform MRA. The proposed analysis is demonstrated on natural frequency and temperature data of the $Z24$ Bridge.

All in all, the thesis tries to answer the following questions:

- How an SHM signal can be judged as non-stationary/stationary and under which assumptions?
- After how many cycles of continuous monitoring an SHM signal that is initially

non-stationary becomes stationary?

- Which are the main drivers of this nonstationary (i.e. EOVs, abnormality/damage or others)?
- How one can distinguish the effect of EOVs from this of abnormality/damage?
- How one can project out the *confounding* influence of EOVs from an SHM signal and provide a signal that is capable for novelty detection?
- Is it possible to decompose an SHM signal and study each one of these components separately?
- Which of these components are mostly affected by EOVs, which from damage and which do not include important information in terms of novelty detection?

Understanding and answering all the aforementioned questions can help on identifying signals that can be monitored over time or in data windows, ensuring that stationarity achieved, employing methodologies such as statistical process control (SPC) for *novelty detection*.

Acknowledgements

First and foremost sincere thanks must go to my supervisors Dr. Elizabeth Cross and Professor Keith Worden. Without their help, guidance and continuing support, this work would not have been possible. Dr. Cross gave me the opportunity to do my PhD, supported me very much on difficult situations faced during my study and made always her best to help my progress. In addition, I would like to thank personally Prof. Worden, whose experience and direction helped me to organise my work in better way and progress faster. On this front thanks must also go to Dr. Graeme Manson, who was initially my second supervisor, as well as, Dr. Evangelos Papatheou, who supported me over my PhD study.

Furthermore, I would like to thank Dr. Ki Young Koo, who helped me numerous times during my PhD work, in order to acquire SHM data and set the sensors for my laboratory experiment. I would like also to thank Mr. Jamie Booth, Technician at the University of Sheffield, who helped me setting up my experiment inside the Jonas laboratory, providing his expertise for any technical problem faced.

I would like to thank the Dynamics Research Group of Sheffield University, it has been a privilege to work with such a lovely group of people.

Finally I would like to thank my Maria for her support during my thesis, as well as my family (Giorgos, Maria and Anastasia Iakovidi).

Author Publications to Date

Journal

K. Worden, I. Iakovidis and E. J. Cross, 2018, New Results in Nonstationarity Signal Analysis with a View Towards Structural Health Monitoring. Submitted to Proceedings of the Royal Society A.

Conference Papers

K. Worden, I. Iakovidis and E.J. Cross, 2018, On Stationarity and the Interpretation of the ADF Statistic, Proceedings of the 36^{th} International Modal Analysis Conference, Orlando, FL (IEEE).

I. Iakovidis, E.J. Cross and K. Worden, 2018, A Principled Multiresolution Approach for Signal Decomposition, IOP Modern Practice in Stress and Vibration Analysis Conference, Cambridge, UK.

Table of Contents

List of Figures

List of Tables

Chapter 1

Structural Health Monitoring and Bridges

1.1 Introduction

Structural health monitoring (SHM) is a relatively new scientific field, which was initially established in aerospace/aeronautics and during the 80s found useful application in civil engineering and especially in large-scale structures, such as bridges, ports, offshore structures, dams, buildings, tunnels, pipelines, oil rigs and others [5].

The main objective of SHM is the efficient monitoring and evaluation of the response and behaviour of in-situ structures, using sensor technology, in order to assess their condition under varying operational and environmental loads, detect damage or faults, evaluate their integrity and identify their existent pathologies. In essence, an effective SHM system should assure the fast, reliable and efficient transmission of information to engineers, owners and authorities with respect to serviceability and integrity of the structure.

For civil infrastructure, this information can help in establishing informative background during construction phases, assessment of existent structures for strengthening and rehabilitation purposes and after construction, for checking the function of structure and if it performs as expected, while also in cases of testing and evaluating the life-cycle expectancy of a bridge. In this way the information provided by SHM can be integrated within inspection and service strategies planned, in order to help

in establishing a more educated decision making (i.e. service, rehabilitation, operation breaks, maintenance etc.), which aims to avoid cases of unrecoverable damage (faults), optimise the distribution of resources: materials and services, reduce downtime and service costs.

SHM has a broad topical coverage due to its multidisciplinary nature as it fosters the intersection of different technologies to address the varied needs and applications. In particular, some of the topics that SHM deals with are diagnostics, condition-based maintenance, vibration methods for damage assessment, signal processing techniques, sensor design (new materials), data mining and management approaches, monitoring conducted on infrastructure structures, wired and wireless communication, embedding technology, development of self-repairable structures and many others [6].

Closely related disciplines to SHM are those of: *i*) condition monitoring (CM), which main subject is damage detection in rotating and reciprocating machinery, *ii*) non-destructive testing (NDT), which is conducted on specified damage critical regions or in the members where damage identified, *iii*) health and usage monitoring machines (HUMS), with damage detection applications in rotorcraft drive trains, *iv*) statistical process controls (SPC), which use a variety of sensors to monitor manufacturing processes for measuring and controlling industrial productivity and *v*) damage prognosis (DP), which is associated with a system's life expectancy and productivity [7].

In SHM literature, there is a large amount of discussion about the objectives of SHM. A commonly accepted idea is given by Rytter [8]. In particular, four hierarchical levels are described with respect to the procedure of damage identification:

- Damage detection, which refers to the procedure of detecting the existence of novelty/damage in the response of a structural system.
- Localisation, which describes the procedure of detecting the damaged region or component of a structural system.
- Quantification, which is associated with the damage type and assessment of its severity.
- Prognosis, which refers to the estimation of structural life expectancy and determination of decision making (preventative measures).

1.2 What SHM Includes

A typical SHM system involves the coordination of different sub-systems (components) in order to provide valuable information that can be used for the purpose of damage identification. Based on [6], a typical SHM system consists of six components:

- The data acquisition system (including the sensing system).
- Communication system available for information transmission.
- Data storage system.
- Diagnostic Tools/Methodologies (i.e. SHM pattern recognition techniques, finite element analysis).
- System of Information Retrieval.

According to [9], sensors can be used to measure numerous different SHM magnitudes. The most commonly measured magnitudes during Bridge SHM are i) accelerations, ii) strains, iii) deflection/displacements, iv) tilts and v) tensions (cable-bridges). Accelerations can be used to extract the natural frequencies and mode-shapes of a bridge, which are commonly employed for damage detection and finite element model update (FEMU) [10]. In addition, environmental measures are also recorded, such as temperature, wind speed and direction, vehicular amounts (weight-in-motion systems), relative humidity levels and other depending on the environment where bridges are situated.

SHM in bridges [11] can be classified into four main categories in terms of the way in which data are collected, as well as the time-scales of data acquisition. More specifically, these categories are: i) static field testing, which includes static/mechanic behaviour (identify certain method of analysis), diagnostic (evaluate the effects of components' interaction) and proof (establish safe load-carrying capability) tests, ii) dynamic field testing, including stress history, dynamic load allowance (DLA), ambient vibration and pull back tests, iii) periodic monitoring, which includes field testing and structural capacity determination tests and finally iv) continuous monitoring, including active monitoring and passive monitoring. The main objective of this thesis concerns continuous monitoring.

Although there are many available data that can be measured, the cost of a complete SHM system (system purchase, installation, monitoring, service etc.) is significantly high, while its usefulness is sometimes under question. Most of the time, organised monitoring systems are available in major and newly constructed bridges [10]. In SHM literature, there is an attempt to decrease the number of transducers without affecting the efficiency of the SHM systems and rely on cost-effective solutions [2] (i.e. a great amount of literature work [12] deals with optimised sensor placement).

In addition, some further parameters, such as the bridge type, the construction material and the type of transducers that will be placed on the structure should be taken into consideration. This is mainly because different construction materials manifest different pathologies and deficiencies and are also affected in different manners by environmental and operational conditions, fatigue and ageing. An interesting investigation conducted by the Department of Transportation of Oregon, US [13] on 13 in total highways bridges included, explaining that is useful to know the damage or deterioration mechanism to be monitored before installing an SHM system.

The latter is important, because otherwise it is probable that monitoring the whole structure placing many different sensors will be costly without profound reason. For example for bridges observed to have foundation instability or an ongoing tension crack progressing, appropriate SHM sensors scheme should be adopted in order to monitor efficiently the aforementioned observations, such as sensors measuring tilts, pressure levels, displacements, strains and environmental features.

Furthermore, the limitations of the acquisition system should be considered in order for data to be collected, stored, processed and presented easily. For example, the data monitored are transferred to a central computer, which needs a system power supply. Although, bridges in urban areas typically have direct supply from the grid, as well as phone lines or fibre optic connections, rural bridges should use alternative supply methods, such as solar panels, wind turbines, radio modems, cellular phones to transfer data to the central server etc. From the aforementioned, it can be concluded that the SHM system design is a complex procedure, considering many factors, besides the cost of equipment and installation.

1.3 Conventional Practice for Bridges SHM

Bridges have high importance for civil infrastructure, as they link different cities and countries, connecting isolated villages and regions with the main highways and substantially play significant role for the economic development of a region or city, reshaping and adding value to society. Due to the increased awareness of the importance of bridges for economic growth and sustainability of civil infrastructure of modern societies, the bridge stock of each country should be efficiently managed. The evaluation of current bridge stock condition, the investment in planning, upgrade and improvement of the service provided to society and the evaluation of available budget are of critical importance for the ability of a country to participate in economic forums and receive funding for further development.

1.3.1 Visual Inspection and Condition Evaluation

The most common way for assessing the state of bridges and characterising their condition is by the use of conventional inspection methods complemented by non-destructive (NDT) and destructive (ND) testing [14]. The main aim of these activities is to collect information about the state of the structure, visually detect pathologies, inspect all components for defects and characterise, in general, the condition of the bridge. This information can be used to update the available bridge management databases, which can be analysed to establish future maintenance schemes, to assure the longevity and sustainability of the structures.

Visual inspection is a fast, inexpensive and effective method to identify the pathologies of a bridge and to observe internal, external defects, deficiencies and damaged regions/members [15, 16]. The deficiencies which are anticipated for each bridge, depends on the construction material used. For example, in reinforced/pre-stressed concrete, deficiencies such as scaling, cracking, delamination, spalling, rebar-steel corrosion can be found; in timber, surface depression, warping, loose connections, failure of protective coating etc.; in steel, corrosion, fatigue cracking, buckling coating failures etc.); in fibre reinforced polymers, blistering, wrinkling, fibre exposure, discolouration etc.; in masonry, weathering, abrasion, plant and marine growth, splitting and others [15, 16].

Although, visual inspection provides a general appreciation and basic evaluation of

bridges' condition, there are some disadvantages associated with its use. Among them, is that the effectiveness of inspection is solely dependent on the experience and knowledge of the inspection team and main investigator. In addition, visual inspection cannot provide specific evidence on defects and cracks occurring within the material, in microscopic scale [16], as well as scheduled maintenance and periodic inspections can offer limited information about structural response [17]. Furthermore, access to specific parts of the bridge can be restricted or inspection activities can be dangerous (high-rise layouts, ravine, mountains) [13] (i.e. inspection of viaducts). Some of the aforementioned drawbacks of visual inspection can be tackled by employing NDT [16], as well as with the use of new technologies, such as inspection via drones.

It is important to add here, that although visual inspection and condition evaluation are relative cost-effective activities, are mainly performed every two or three years or in exceptional cases annually (i.e. army bridges) [16]. One can argue that the latter periodic inspection activities may sometimes considered as insufficient for a thorough condition evaluation and especially difficulties can occur when past inspection material are not very informative for future inspections [17]. In addition, the travel costs should also be considered in the case of site visits [17].

Therefore, SHM can be significantly useful here, because the installation of sensors on the structures and the remote data acquisition can provide a more thorough investigation of structural response and limit also the associated travel costs. Besides that there are cases that inspection, condition evaluation and assessment based on load rating factor cannot be representative of structural evaluation. In particular, in [13] there are cases of reinforced concrete girder bridges, which were assessed based on load rating factors and periodically inspected and judged as incapable to carry trunk loads. However, after performing SHM activities, such as dynamic and static tests on them it was observed that the bridges had adequate strength to carry these loads.

Another fact is that condition evaluation and visual inspection is mainly focus on the examination of particular bridge parts, looking for in-prior determined types of deterioration expected, such as the examination of particular bridge members, referred to as fracture critical members (FCM). These are more critical to the overall safety of the bridge than others, because their failure can lead in irreversible damage [16], such as absence of redundancy (load path; number of available load paths, structural; continuity in load path: statically indeterminate structures, internal; failure

spreads from one member to the other), which describes the ability of neighbouring members to carry the re-distributed load resulted by the members' failure. As one can understand, thorough knowledge of structural behaviour, mechanics and level of experience is needed to identify such parts. Here SHM can play a critical role as sensors can be placed to such critical members and their response can be monitored and evaluated over time, something that can provide more knowledge to the engineer and also help evaluating if already conducted service activity was sufficient assuring the proper function of the structure.

In other words, here it can be said that SHM can provide a very useful alternative to the traditional condition evaluation and visual inspection activities. In particular, the existence of sensors on structures can provide remote monitoring and evaluation of structure, which can enhance the knowledge of engineers and also help them organise better the needs for maintenance. In addition, for cases of critical members monitored, SHM can provide data that can be processed in order to identify the presence of any on-going deterioration mechanisms, which could not been identified by inspection activities. All the aforementioned show the preventative character of SHM and its direction to provide safety and organise better the needs for maintenance activities, as well as the opportunity for a more effective bridge stock management.

1.3.2 Bridge Management

In order to manage efficiently their bridge stock, countries created inventories and unified databases, which include multiple information about the bridges, such as their type, material, components (superstructure, substructure), location, support condition, specifications, number of lanes, traffic accounts, operational conditions, geometric data, age, which can help to categorise them and keep track of inspection data, such as ratings, surveys, time of inspection, maintenance activities [18]. This analytical collection of data for numerous bridges is referred to as bridge management [19].

For structural evaluation of bridges, the most common guidelines followed are provided by the Federal Highway Administration Guides (FHWA), which are established by the U.S. Department of Transportation and introduced as code of practice *AASHTO* [18]. These are, in general sense (slight differences), adopted from Eu-

ropean Standards. The equivalent specifications for U.K. is the design manual of roads and bridges ($DMRB$) [15].

Following the $AASHTO$ determined evaluation here, the evaluation of structural integrity is performed on a scale from 0 to 9, where 9 describes that bridges meet the desirable criteria and perform even better, while 0 describes that the bridge's operation should stop as soon as possible. For ratings under 5, corrective and priority actions are required [14]. The bridges of high concern are also classified as structurally deficient or functionally obsolete, however, neither term necessarily implies lack of safety [14–20]. In particular, a bridge is referred to as structurally deficient when it includes a significant defect which often means that speed or vehicle weight limits must be enforced on the bridge to ensure safety. On the other hand, structurally obsolete is a bridge that can function properly, however, at levels which are below the contemporary standards. The latter means that traffic associated restrictions should be imposed, such as closing a traffic lane in order to reduce the traffic tonnes that the bridge sees [15].

1.3.3 Bridges' Condition Nowadays

Based on the previous way of assessment and classification, bridge stocks from different countries have been assessed and important conclusions have been drawn [14, 21, 22]. In particular, the most commonly observed defects and deterioration are associated with design and construction issues, such as insufficient concrete cover, dense placement of steel rebars, inappropriately placed joints, problems with drainage systems, alkali-silica reaction (ASR) susceptible aggregates, insufficient foundation capacity, poor concrete quality, bad compaction, inadequate curing, poorly fixed reinforcement, faulty ducting for post-tensioning systems, inadequate grouting, painting and coating, overloading, vehicle impact, chloride attack, carbonation, poor maintenance, freeze-thaw cycle, dynamic loading [16].

According to recent studies and projects [21, 22], counties with big population of bridges, such as USA, China, Australia, Canada and Japan have a high percentage of structurally deficient bridges, of which the majority were build within 60s and 70s, something that is also happening in European countries, such as United Kingdom, France, Austria, Switzerland and Germany [23]. From this, one can understand that societies face the necessity to conduct maintenance, that costs billions each year, to

upgrade, strengthen and/or reconstruct many bridges. Some reasons behind the observed problems were, a) the significant increase in bridge construction until late 60*s* following the rapid growth of economy, which led to cost cutting solutions, b) the limited design knowledge and the conservative speculation, which in many cases led to the use of low quality materials, insufficient detailing and overestimation of load-carrying capacity [21, 22] etc.

After major bridge collapses, such as the Silver Bridge (1967), I-35W in Minnesota, Kisogawa Ohashi and Honjo Ohashi Bridge in Japan, the engineering community [14] became alarmed and started developing ways to improve bridge condition, something that provides ground for further development. In this direction, SHM was introduced for bridge monitoring. In particular, in the last 30 years, there has been a lot of investment and research orientated into monitoring technologies capable of improving the current state of bridges and other infrastructure projects, aiming to allocate resources towards repair, replacement/rehabilitation, as well as future estimation of life expectancy.

1.4 Motivation behind SHM for Bridges

To begin with, SHM can provide understanding of the behaviour of full-scaled in-situ structures, assessing, both physically and numerically, their actual response and bearing capacity, under full cycles of environmental and operational loads. In addition, SHM can be utilised to understand the effects that environmental and operational variations (EOVs), time-dependent effects (material deterioration and fatigue) and live loads have on structural response.

Furthermore, SHM is very useful for the testing and evaluation of innovative structural designs and new materials (i.e. ensuring safety of as yet unproven materials and systems). For example it can provide ground for testing and evaluation of newly developed, functional and economical sensors, data acquisition systems (DASs), communication technologies, new data transmission, collection, archiving, retrieving systems and novel data processing methodologies [11].

Moreover, monitoring can improve the level of management available for existing structures. This is achieved as monitoring results provide background to better understand structural behaviour, evaluate bearing capacity and life expectancy [11].

For example, the evaluation of bridges' load carrying capacity was based for many decades on the concept of *allowable stress/strength design* (ASD), where each component was designed for a level of stress much lower than its maximum allowable strength. In other words, the *nominal* strength of a member provided by multiple laboratory tests is divided by a safety factor (i.e. 2), which was mainly an empirical assumption [10]. This working stress methodology introduced large strength reservoirs. The past 30 years ASD method was substituted by the load and resistance factor design (LRFD), which is a probabilistic/reliability-based design method, which aims to account for the uncertainties occurring both during load estimation and structural resistance [15]. In addition, the assessment of existent bridges is also based on LRFD concept.

According to the revised versions of DMRB [15], there are parts describing some typical actions for monitoring, however without providing extensive knowledge background, something that is left to engineer's judgement. SHM has a very important role in order to efficiently manage bridges. In particular, the response of existent bridges with over 50 years of life can be monitored in order to evaluate their strength reservoirs and consequently the remaining life of the structure. Here one can argue that a simple assessment process as required by standards cannot be always sufficient for a thorough assessment before conducting maintenance. More specifically, the requirements for assessment do not include any dynamic analysis of a structure, which is of great importance to identify the boundary conditions of the structure and consequently determine, more precisely, the static configuration of a bridge. For example an simple assessment can hide locations of structure were moment and shear forces have an increased effect, such as the connections between piers/abutments with the deck and therefore inaccurate structural analysis configuration can lead to questionable results.

Furthermore, for bridges constructed the last 30 years, SHM can provide a great background in order to verify and help in developing further the LRFD method, which is already mainly based on the statistical and probabilistic analysis performed on data collected from SHM. By aiming for further development of SHM one aims to save inspection time, reduce needed maintenance activities (downtime), re-evaluate inspection budgets, re-distribute financial resources and improve the state of practice and design. In addition, the degree of confidence with respect to visual inspection is improved, because SHM can help to assess a larger portion of structures and minimising human factors/errors.

1.5 SHM in Bridges

1.5.1 Damage Sensitive Features (DSFs)

SHM can provide a wide variety of measured data. Representations of these are also referred to as *features* and certain of them demonstrate great sensitivity to change associated with damage [7]. In a general sense, DSFs can be categorised into vibrational-based features, which can be natural frequencies, modeshapes, mode-shape curvatures, modal strain energy, operational deflection shapes (ODS), frequency response functions (FRFs), transmissibilities, structural flexibilities and damping ratios. On the other hand, there are the component-based features, which are displacements, tensions, strains, moments, deflections, tilts and others [22].

Some of these measures and more specifically the vibration-based data are obtained after a procedure of manipulating the measures of time domain (i.e. accelerations) and decomposing them into the frequency domain. These methods are based on operational (OMA) and experimental (EMA) modal analysis. Some of the most commonly used methods are: fast Fourier transform (FFT), singular value decomposition (SVD), peak picking method (PP), frequency domain decomposition (FDD; Enhanced FDD, Curve-fit FDD), stochastic sub-space identification (SSI) and others. A great review of these methodologies can be found in [24].

1.5.2 Approaches to SHM

A plethora of SHM approaches have been developed throughout the years for damage identification. These can be classified into two main approaches. The first is the physics-based or model-based approach, which involves the use of advanced numerical methods, such as finite element analysis (FEA), for the simulation of the governing dynamics (stiffness, mass, damping) of a structural system. On the other hand, there are the data-based approaches, which rely on statistical pattern recognition (SPR) to make inferences directly from data. An interesting study for their strengths and weaknesses can be found in [25].

The physic-based modelling, includes FE bridge modelling, updating and calibration, which is performed based on monitored data. The FE model updating methods

can be generally categorised into sensitivity and direct ones [26]. The former is based on the formulation of parametric models of the bridge, aiming to minimise appropriate penalty functions, which approximates the relative error between monitored data and FE extracted results, as well as the selection of sensitivity parameters, which are mainly based on the judgement of the engineer and their physical/mechanical importance (sensitivity analysis). On the other hand, the direct FE update methods perform a direct update of the dynamic properties (mass, stiffness, damping matrices) of the FE model, aiming to approximate the monitored data.

The FE model can be used to model the dynamic behaviour of the structure, which can give a relationship between the matrices affecting its dynamic response. These are the stiffness, mass, damping matrices and the vectors associated with the force or acceleration introduced to the system. Here it is important to mention that sensitivity analysis provides very useful information about the sensitivity of each localised member and can be used to identify the most sensitive parts of the structure in terms on how much they change the dynamic features, such natural frequencies, modeshapes etc. In addition, interesting information can be provided, such as the identification of boundary conditions of a structure, which can be obtained by the close examination of modeshapes and the associated natural frequencies.

The results provided by a FE model have been used inside SHM literature in many different ways. Firstly, there are available damage detection methods based on vibration measures. An conclusive overview is provided in [27]. The main four vibration methods for damage detection are based on i) the change on natural frequencies [28], ii) modal residuals [29], iii) modal curvatures (modeshapes) [30, 31] and iv) the dynamically measured flexibility [32]. Secondly, FE models can be used to simulate different possible damage scenarios [31, 33] to quantify their effects and identify the most sensitive components. Furthermore, FE bridge models can be used for probabilistic analysis aiming to predict the remaining life of bridges [34, 35]. Next, there are cases where the FEM used to model the bridge and from these to extract physics-based relations or create statistical models capable to model their behaviour in order to perform damage detection [36].

Although, physic-based modelling is used extensively in the industrial and research engineering worlds, there is great uncertainty associated with it [25]. This is mainly the outcome of the involvement of engineering judgement during the procedure of formulating the modelling assumptions. For example, the determination of boundary conditions is not a trivial procedure and there are a lot of assumptions involved,

which can significantly affect vibration and component-based FE results/projections. However, the existence of FE models simulating the structural behaviour are helpful to establish a general appreciation of the structural response (dynamic and static), perform parametric analysis (i.e. traffic load distribution) and back-up visual inspection results, relevant analysis performed and decision making.

On the other hand, data-based modelling (interest of current thesis) is mainly based on SPR [7]. This is a wide scientific field including many different methodologies, such as linear and non-linear neural networks, supervised and unsupervised learning methodologies, intelligent feature selection and extraction, novelty and outlier detection, decision trees etc. [5]. From the aforementioned methods, those of interest here are supervised and unsupervised pattern recognition, as well as the concept of machine learning. In abstract terms, the concept of machine learning can be described as the process of learning or training based on the relationship between measured data. In a general sense, the learning problem can fall into two classes, the supervised and the unsupervised [7]. The former describes the process of learning/training of data relations, when the data can be classified and labelled from the beginning of the process (i.e. undamaged and damaged), while the unsupervised describes the case where data classification and labelling is absent or a single class exists, something that requires for the learning to be based on intrinsic relations within the data [7]. A review of the most commonly employed methods used for bridge SHM will be presented in Chapter 2.

1.5.3 Basic Axioms of SHM

At this point, it is important to mention the basic axioms used for SHM. These can be used as general guidelines for damage identification [7]:

- Every structural material has inherent flaws or defects in microscopic level.
- Two system states (i.e. undamaged/training and potentially damaged/testing) should be compared to investigate for damage existence.
- Damage identification (existence and location) is possible with unsupervised SPR, but severity and type of damage can be assessed only in supervised mode.
- Measures provided by sensors, should be firstly modified (i.e. signal processing, intelligent feature extraction, statistical classification) before used for damage

identification attempts.

- The sensitivity of SHM features to damage is highly associated with the impact of changing operational and environmental conditions. Hence, intelligent feature selection is required.
- The requirements of an SHM system depends highly on the length, time-scale associated with damage initiation and propagation investigated.
- There is a trade-off between the sensitivity to damage of an algorithm and its noise rejection capability.
- The size of damage that can be detected from changes in system dynamics is inversely proportional to the frequency range of excitation.
- Structural complexity increases due to damage existence.

Here it is important to mention some of the main challenges that SHM face nowadays. To begin with, is the evaluation and quantification of the confounding impact of environmental and operational variations (EOVs) on damage senstive features. It is desirable to project out the impact of EOVs, because it has been observed in many cases inside SHM literature, that the EOVs impact can mask the variability of features that can be associated with damage, something that makes damage detection attempts difficult. This is one of the main aims that the current thesis has and wish to investigate further.

Another important challenge is the attempt to retrieve as many as possible information for as less as possible number of sensors. The latter means that methodologies involved and implemented inside SHM literature try to extract as many as possible valuable information from sensors. The number of sensors used to retrieve the aforementioned information is associated with the overall cost of the monitoring system installed in a bridge. The current thesis aims to deal with this issue, however indirectly by trying to develop further the already existent method of cointegration inside the context of SHM and provide further information for non-stationary/stationarity assessment and time series decomposition.

Finally, it should be mentioned that an additional challenge is the lack of data available from destructive testing and/or the damage state of bridges. In particular, there are not many cases and available data from damage state of bridges in order to develop and use *supervised learning* methods for damage detection and evaluation.

Due to this reason, the methods developed in this thesis are following the *unsupervised machine learning* approach, where data are available only from the undamaged state of the bridge. The methods developed are based on normal condition data, from which damage sensitive features are extracted and monitored over time on the newly recorded data provided by real-time bridge monitoring. These methods are tested on an existent showcase where damage is involved (Z24 Bridge) and the laboratory truss bridge experiment conducted, which also provided damaged data.

1.6 Conclusions

In this Chapter, the concept of SHM has been introduced. In particular, its origins, fields of application, multi-disciplinary nature and generalised objectives are discussed. In addition, the sub-systems comprising a typical SHM system are mentioned, focusing on the available sensor technologies and categories of SHM testing. Furthermore, the role and motivations behind SHM and especially bridge SHM are described, through a discussion on conventional practice (maintenance, inspection, NDT) employed for effective bridge stock management, underlining some basic considerations faced in some developed countries nowadays. Then, an introduction on bridge SHM for damage identification is made, concentrating on the measures provided by SHM, methods to extract them and modify them for damage-sensitive features and the approaches commonly used to manipulate the DSFs. The main focus of the thesis is placed on data-based approaches and unsupervised SPR.

Chapter 2

Damage Identification and the Data Normalisation Issue

2.1 Introduction

The unsupervised data-based modelling approach is commonly applied for bridge SHM, due observed lack of data associated with the damaged state of bridges [5]. The high structural/communal importance of bridges and their significantly high cost, means that destructive testing should in general be avoided, however it is important to achieve the targeted aims of SHM for bridges, to minimise the interruptions of bridge operation (downtime), ensuring the free traffic movement and goods transportation [5].

Numerous SHM techniques have been developed throughout the years for damage identification purposes in bridge SHM. In the current Chapter, a review of commonly employed methods for bridge SHM is made, whilst the main focus is placed in methods based on a specific approach referred to as data normalisation, which seeks to tackle the influence of confounding influences on damage sensitive features.

2.2 Damage Identification in Bridges SHM

Looking in SHM literature, the majority of work on bridges has been focused on vibration-based data. In such cases, damage identification is commonly performed by employing features extracted from acceleration measurements for bridges, such modeshapes, natural frequencies, modal curvatures and others. These are able to provide a generic idea of the dynamic response of the structure, whilst they exhibit significant damage sensitivity. Furthermore, there are many methodologies used for damage identification on bridges structures. Commonly methods include time-series models, outlier detection, neural networks, statistical process controls, support vector machine and many others [7].

On of the most commonly used method for damage identification is autoregressive (AR) modelling. In the case of AR models, the training set (assumed undamaged state) is used to obtain the parameters of the regression models, such as the coefficients, number of lags and the corresponding statistics (i.e. residual error, mean, standard deviation). These can be used either directly as DSFs or in part for creating intelligent DSFs. After the DSFs are established and tested for its appropriate use, the remaining data representing the potential damage state, are projected on the established AR model. Any deviation observed on the DSFs, between the two states, can be assumed as an indicator of novelty existing in the data (i.e. damage related). In many cases, statistical thresholds are employed to define the abnormality limit in order to cope with erroneous abnormality detection (false alarms) or statistical control charts used to capture novelty. An interesting review about the available AR models used for bridge SHM is provided in [7].

In particular, in [37] an AR model is employed to inspect the recorded acceleration time-histories from the undamaged and the damaged state of a reinforced concrete (RC) bridge column. The mean and the variance of residual errors of the AR model employed as DSFs and statistical control charts used (X-bar and S-charts) to observe their variation in the presence of damage in different locations.

Additionally, other AR model modifications have been explored, such as by introducing exogenous terms (ARX) and moving averages (ARMA). More specifically, in [38] an ARMA model is employed for damage identification in both time and frequency domains, from the recordings provided by the experimental testing of a IASC-ASCE four-storey benchmark frame, the Z24 Bridge and the second link of

Malaysia-Singapore bridge. A linear deterministic trend introduced in the ARMA model in order to account the long-term impact of creep and in addition a diurnal trend introduced in short-term without triggering any false alarms.

Moreover, the response of a laboratory constructed lightweight aluminium truss model was monitored in [39]. Damage indication and detection are performed employing ARX models using numerous DSFs (time-domain i.e. accelerations; frequency domain i.e. PSDs and FRFs employing DSFs such as residual, residual variance, maximum log likelihood function, residual correlation factors and fitting coefficients).

Moser and Moveni [40] monitored and tested the modal response of Dowling Hall Footbridge. The natural frequencies obtained are modelled using many mathematical techniques such as ARX, linear bi-linear and polynomial regression. Fourth-order polynomial regression provided the best accuracy. The authors noticed that temperature affects the standard deviation, which can be used to identify damage with higher accuracy.

Omerzetter and Brownjohn [41] formulated a vector seasonal autoregressive moving average (ARIMA) model to investigate the strain measures recorded on Singapore-Malaysian second link bridge (post-tensioned continuous box girder), allowing its coefficient to vary with time employing a Kalman filter. The in-service condition of the bridge was monitored under normal and operational conditions and the model was able to detect unusual events in the strain histories, which some of them investigated for structural change (damage onset).

Kullaa [42] proposed to use different univariate and multivariate control charts, such as Shewhart, x-bar, CUSUM and EMWA, in order to monitor online modal properties, such as natural frequencies, mode shapes and damping ratios. The results shown that the variability of DSFs is influenced by environmental and operational variations (EOVs) and suggested the use of data dimension reduction approaches (i.e. principal component analysis-PCA) for increasing data damage sensitivity.

Mosavi et al. [43] generated a DSF based on the coefficients and residuals of a vector autoregressive (VAR) model and use it for damage localisation in a steel beam tested under ambient vibration. This method provided good results showing high sensitivity to low and high levels of damage, as well as damage in different locations. The authors suggested further investigation in cases of data exhibiting

high level of nonstationarity, such as ARIMA or vector ARIMA (VARIMA), as well as further study on the effect of sensor locations and optimum spacing.

From the aforementioned one can understand that there are numerous approaches to damage identification inside the context of SHM. As mentioned in the introduction, the majority of this work is based on approaches that employ the DSFs directly without further modification. However, from a great amount of SHM literature, it has been observed that many DSFs show significant sensitivity to EOVs, such as temperature, wind, traffic, humidity and others. The issue dealing with the impact of EOVs on DSFs is referred to SHM literature as data normalisation.

2.3 The Data Normalisation Issue

Inside the context of SHM, data normalisation can be defined as the process of distinguishing the variability of DSFs associated with damage from that associated with EOVs [2]. More specifically, one of the main issues observed during monitoring and damage identification efforts is that existing EOVs tend to mask damage-induced variability present in response series. In particular, if effects, such as those introduced by EOVs, dominate the damage state of the structural response, the damage identification efforts become difficult, while their results can be uncertain [2]. Hence, the identification and isolation of the impact of EOVs on DSFs can be of great interest, because it can enable one to purge the effects of EOVs enabling outlier and novelty detection.

The approaches associated with data normalisation can be generally distinguished into subtraction and projection methods. Their difference is subjected to the way that the impacts of EOVs are compensated from the DSFs. For example, a common subtraction method is the estimation of the impact of EOVs through regression and its consequent elimination from the DSFs. In contrast, a projection method (which in many cases does not require recorded EOVs) can be thought of as a means of mathematically reconfiguring measured feature data into a lower dimension to remove any sensitivity to benign changes. However, a more systematic categorisation of data normalisation approaches is provided in [7]. In particular, these can be categorised into five categories: (i) experimental or empirical approaches, (ii) regression modelling, (iii) look-up tables, (iv) machine learning approaches and (v) intelligent feature selection. Some basic concepts behind these approaches, as well as related

literature will be discussed in the following.

i) Experimental/Empirical Approaches
The experimental/empirical approaches are those whose main aim is to focus and record the effects associated with data normalisation issue, attempting to identify their potential sources (i.e. EOVs). These studies have an empirical character, due to the fact that the majority of cases are experimental or are the findings of planned real-time destructive testing on systems, in order to gain valuable inference from them.

The most commonly described source of EOVs is temperature. The impact of temperature can affect vibration-based data in both the short and long term. More specifically, the long term impact of temperature is discussed in [4], Peeters et al. [44], Bricker et al. [45], Peeters and De Roeck [46], Rohrmann et al. [47], describing the seasonal effects that temperature can have on natural frequencies (highly increased value during winter) on many bridges (i.e. $Z24$, Pedro de Ines, $I-39$ Kishwaukee River Bridge). Temperature impacts have been associated with boundary conditions' change (bearings and expansion joints), soil-structure interactions and construction material properties variation (i.e. wearing surface/asphalt stiffness increase). In addition, the impact of diurnal temperature variation can be associated with daily frequencies variability. For example, Doebling and Farrar [48] mentioned that the observed daily natural frequency variation on Alamosa Canyon Bridge was the impact of the air temperature variation on a 24 hour basis, and more specifically it has been associated with the daily temperature difference variation between the top and bottom bridge's flange.

Besides temperature, the impact of traffic mass introduces significant variability on the measured natural frequency of bridges. In particular, during the monitoring of three bridges (Nogro, Sangjin and Namhae) in Korea [49], it was mentioned that both light and heavy traffic has impact on the variation of first five fundamental deck natural frequencies, which significance was a function of bridge span and dead weight. Farrar et al. [48] performing vibration tests and destructive testing on the $I-40$ Rio Grande Bridge and on the Alamosa Canyon Bridge in New Mexico, US, mentioned that the temperature and the traffic can have a significant impact on the variation of natural frequency and tend to mask the variation associated with damaged introduced. In the case of Tamar Bridge [2], the traffic has been observed to be the main EOV impact associated with the fluctuation of the fundamental deck natural frequency, something that can be explained by the significant increase

of daily vehicle numbers, which led also to the re-strengthening of bridge's deck in order to reinforce its bearing capacity.

Moreover, according to [50–52], wind induced vibration can have an impact on the structural response of bridges, especially in cases where extremely high wind speeds are observed or when the deck-cross section has a compact closed form (concrete decks). Cable stays, in suspension and cable-stayed bridges, are prone to exhibit large amplitude and continual oscillations, due to their large flexibility, small mass, and small damping, which can cause fatigue and corrosion of strands in surprisingly short periods. The cable tensions are highly sensitive to air temperature and the latter sensitivity depends on many factors, such as their configuration, angle of inclination, initial curvature, level of pre-tension, cable sag effects and others, which have a significant impact on the cables stiffness [53–56]. In many cases, the combined action of wind and rain, introduces significant vibration on cables leading to instability. The latter is associated with droplets formulating a liquid film and rivulets affected by the influence of drag, gravitational effects, friction forces and more severe air induced phenomena, such as galloping [57, 58].

From all of the above, it can be concluded that EOVs, such as temperature, wind, traffic mass, humidity have a significant impact on the variability of the most commonly employed monitored features for damage identification efforts. These can be referred to as sources of operational and environmental variability (EOVs) and their monitoring is necessary for any SHM study.

ii) Modelling via Regression
In the case that EOVs measures are available, data-based models can be developed, trained, tested and updated, in order to predict DSFs (i.e. natural frequencies) based on measured EOVs data. The main concept of this approach is to capture, using a variety of regression techniques, the impact of EOVs on DSFs. These models can use the available EOVs measures to express the DSFs as functions of EOVs (temporal terms) and can also involve dynamics by adding history dependence between DSFs and EOVs (e.g. autoregressive models) [7].

iii) Look-up Tables
The main concept of this approach is to establish damage sensitive features in the absence of EOVs measures that can be used to formulate look-up tables to compare different states of structural response. In the majority of cases, a feature vector is established from the available undamaged state of the data, which has been trained

and tested inside the undamaged state.

For example, an AR model can be used for this purpose. The normal condition data can be divided into multiple samples, which will be used for model training and testing, and the AR can be used to obtain the coefficients for these samples. Then a discordancy measure, such as the Euclidean distance (univariate data) or Mahalanobis Squared Distance (MSD) (multivariate data), can be used to find the training sample which coefficient vector is closest to the testing sample coefficients, in terms of distance. In this way, a look up table can be established including the available training states that are compared to the tested ones. Then DSFs can be generated based on the statistical inference provided by the AR model, such as residual errors, that can be used for novelty detection.

iv) Machine Learning Approach
Some of the most common machine learning algorithms that are used for the data normalisation issue are: 1) the auto-associative neural networks (AANN), 2) factor analysis, 3) discordancy measures and outlier analysis, and 4) singular value decomposition (SVD). More information can be found in [7, 59]. A brief review is provided below.

AANS
AANNs are neural networks that map network inputs onto themselves in order to learn internal data structure. In many cases these do not include EOVs measures and treat them as hidden variables in their network structure [7]. The network is trained and tested based on its functional dependencies using the undamaged condition. In case of damage, an increase of AANNs predictions error is expected. A good example on the use of AANNs on bridge SHM data is given in the work of Sohn et al. [60], which presented a combination of an *AR-ARX* model, *AANNs* and sequential probability ratio to examine the response of Alamosa Canyon Bridge and discriminate the change caused by ambient conditions from changes caused by damage.

Discordancy Measures
Discordancy measures for outlier analysis are commonly employed for novelty detection. These measures can be distinguished into univariate and multivariate. The most common univariate measure is the Euclidean distance, while the most common multivariate is Mahalanobis-squared distance (MSD). These have been used either directly or as a part of more complex algorithms. For example, Dervilis et

al. [61], introduced a robust multivariate method using the minimum covariance determinant estimator (MCD) and the minimum volume enclosing ellipsoid (MVEE) as novelty indexes in multiple data, such as the FRFs of a piper tomahawk aircraft wing, a wind turbine gearbox vibration data and Z24 bridge. These studies pointed out the different way in which outliers associated with EOVs manifest themselves with respect to those associated with damage and how classification can be made. Some more examples of using discordancy measures in SHM literature are provided in [62, 63].

Factor Analysis and Singular Value Decomposition (SVD)
Factor analysis is based on the correlation of a small number of variables of dimension (f) and other dependent variables of dimension (m; $f < m$), which are assumed to correspond to EOVs (e.g. temperature, vehicular load). Some further information can be found in [42, 64]. In addition, SVD is a technique, which is commonly employed for operational modal analysis (OMA). However, it can be used for data normalisation and novelty detection. In particular, it is based on the determination of the rank of the feature vector matrix (including EOVs), which is re-generated introducing augmentation of a feature vector that is assumed for possible damage. It's novelty detection concept is that in case of damage, the rank of feature matrix will increase [2]. Some further information for the application of SVD as novelty detection measure is provided in [65, 66].

v) Intelligent Feature Selection
In general sense, the selected SHM feature should be damage sensitive and insensitive to EOVs impact. One approach that is used extensively inside SHM is the Principal Component Analysis (*PCA*) [67, 68]. In particular, *PCA* is a data dimension reduction method, which is used to project the original features into a space spanned by the eigenvectors associated with the largest eigenvalues of the feature covariance matrix. This is the reason why *PCA* is considered a projection method. *PCA* using minor components is explored by Manson [69] and Cross [2], based on the concept that the main variability of DSFs caused by EOVs is trapped inside the main principal components of *PCA* and consequently the minor components include minimised variance, which can be used for damage detection (benchmark testing data provided by EU project Damascos).

Furthermore, non-linear extensions of *PCA* have also been employed for novelty detection, such as non-linear *PCA* (NLPCA), kernel regression *PCA*, local *PCA*, greedy Kernel *PCA*. In [62], Sohn et al., proposed and compared *PCA* and *NLPCA*

in combination with *AANNs* for damage detection. In addition, *PCA* and kernel regression has been used for damage diagnosis in a laboratory bridge [70]. Some further information about the application of *PCA* and *NLPCA* inside SHM context can be found in [71–73].

vi) Cointegration
Another projection method that has been used in the past for dealing with the issue of data normalisation in the context of SHM is cointegration achieved via *Johansen's* approach. In particular, this concept originates from the field of econometrics [74, 75] and was employed in the past for novelty detection on the Tamar Bridge [2, 76], laboratory experimental applications on aluminium plates and composites [77, 78], monitoring of dam structures [79] and offshore pipelines [80]. A guide for using cointegration for SHM purposes can be found in [76].

Cointegration method can be used when multiple nonstationary variables/series are available, such as in the case of SHM. The latter nonstationarity is manifested through the existence of common trends shared by the SHM series, which can be associated with the impact that EOVs impose on them over time. Cointegration can project out the aforementioned common trends by obtaining a linear combinations between the series. This combination is referred to as cointegration residual and represents a stationary signal. In the absence of common trends or in other words in the absence of EOVs effect on SHM series, the cointegration residual is damage sensitive and is capable for novelty detection.

In addition, there are some work related to non-linear cointegration, where time series are nonlinearly combined. In particular, Cross and Worden [81], proposed two possible approaches to non-linear cointegration, while Shi et al. [82] proposed an non-linear cointegration approach based on Augmented Dickey-Fuller (ADF) test and Gaussian process regression, which was validated using modal data of Z24 bridge.

Cointegration was introduced to the SHM community in 2011 [76]. Although it has shown some promising results, this has mostly been limited to experimental data or limited monitoring campaigns. It is important that the proposed methods are tested in cases where damage is small and where multiple EOVs are considered. Further to this, little work has been done to investigate how best to use the cointegrated residual, how different tests for nonstationarity can be used effectively, or how decomposition methods may prove useful.

Although the main concern of this thesis is bridge SHM, it is useful to say that, in general sense, the aforementioned SHM methods can apply to data coming also from other SHM applications, such as from aerospace, offshore structures, mechanical systems and others. The main difference has to do with the time-scales in which the data are collected and the level of nonstationarity the data possess. The latter can be associated with the impact of environmental and operational variations on the SHM data collected from different SHM applications, since each structure has its unique behaviour and is tested under different environmental and operational situations.

2.4 Scope of Thesis

According to the previous literature review, it can be said that the main interest of this thesis is placed on bridge SHM and damage identification based on novelty detection and machine learning approaches. The main interest lies on projection methods and especially cointegration. In particular, the linear cointegration approach will be explored further to identify possible drawbacks, optimise its use, validated through different experimental and real-time data investigations, as well as explore the possibilities of extending the methodology or combine it with other SHM based methods.

The analysis in this thesis will be performed under the theoretical framework of data normalisation and using multiple data coming from both experimental laboratory bridge structures testing, as well as by real-time bridge monitoring data, which are provided by organised campaigns. The thesis layout is presented below.

2.4.1 Thesis Layout

This thesis includes seven Chapters, excluding the first two (9 in total). The third Chapter describes a laboratory experiment, conducted in Jonas Lab at the University of Sheffield, where the aim is damage detection from data acquired by the vibration analysis of an aluminium truss bridge. This Chapter discusses the experimental set-up, the vibration analysis, the process of data acquisition and damage introduction and finally a summary of the experimental results.

Chapter 4 will provide an overview of the available real-time Bridge SHM data. Firstly, the data sets of Tamar Bridge will be presented and special attention will be placed on the cable tension data. After that a preliminary analysis will be performed on the cable tensions data which will include simple plotting and response surface modelling in order to evaluate the impact of the available environmental and operational variation measurements (wind speed, traffic mass, temperature and humidity) on them. The obtained results from simple plotting and surface modelling will be discussed. This preliminary analysis aims to understand when SHM signals can be considered as appropriate for use inside cointegration (the latter will be demonstrated in greater extend in Chapter 5) and how simple plotting can help identifying common trends in the signals. Finally, a short description of Z24 Bridge data sets will be provided, which will also include some preliminary analysis in order to draw important conclusion about the relationships between series.

In Chapter 5, the Johansen approach to cointegration will be introduced for damage detection purposes and the best practice, as recommended by the author, for its application on the available data will be discussed, focusing on important procedures during its application (i.e. training set selection, lag-length specification and common trends manipulation). In addition, SHM data from three bridges (Tamar Bridge, $Z24$ and laboratory truss bridge) will be used to demonstrate the use of Johansen's approach to cointegration, as well as discuss how the interrelations between SHM series are affected by EOVs and how are these form the cointegration vector.

Chapter 6 will discuss the use of *Augmented Dickey-Fuller* (ADF) test for testing the stationarity of SHM series inside cointegration. In particular, an alternative interpretation of the ADF t-statistic will be introduced, which is based on the fundamentals of signal processing and on dimensional analysis. The latter interpretation will be introduced employing the continuous SHM signal of the second natural frequency of Z24 bridge, whilst a brief example on Tamar Bridge's 1^{st} natural frequency signal will be used to show how one can define the critical frequency of a signal and estimate after how many cycles the signal will become nonstationary.

The seventh Chapter will focus on the attempt to understand the novelty detection capabilities of Johansen's approach to cointegration. In particular, an attempt will be made to answer the question of what information one can actually gain from the cointegrated residuals and how (with which techniques) this information can be tracked. For this purpose the available methods in literature for monitoring the

cointegration residual will be demonstrated and discussed on the data generated using a simple mass-spring system and on the data sets of the experimental aluminium truss bridge.

Furthermore, Chapter 8 will present the use of *multi-resolution analysis* (MRA), which is based on *orthogonal wavelet analysis*, in order to decompose SHM signals. This decomposition scheme is based on the Debauches wavelet family and the *ADF* test is used to test each level obtained by the decomposition. In this way, the signal can be decomposed in principal manner into a nonstationary, stationary, noisy and a mean component. The importance of this procedure will be discussed within the bridge SHM context as the analysis is performed on the frequency and temperature signals of the $Z24$ bridge.

Finally, Chapter 9 includes the conclusions of thesis and the recommendations for future work.

Chapter 3

Experimental Data for Analysis

3.1 Introduction

The main aim of the experimental work that will be presented in this Chapter is to investigate how damage affects features that are commonly used for bridge SHM and their consequence on cointegration. Another aspect of the experimental plan was to study the structural response when the system was 'repaired' or returned to its initial state.

More specifically, this Chapter provides information about the experimental setup, measurement system and data collected from an experimental vibration analysis performed on a laboratory aluminium truss bridge. The experiment was conducted at the George Porter's laboratory (Jonas Lab), at the Mechanical Engineering Department of the University of Sheffield. In particular, a vibration analysis is performed on a bridge truss model aiming to monitor its vibration-based response under a varying temperature field, which was simulated inside an environmental chamber in the Jonas Lab. The environmental field covered temperatures between 25 to -10 ^{o}C. In addition, four damaged states were simulated by the removal of bolts at selected connections of the truss.

The results of this study will be used in the next Chapters to demonstrate how cointegration can work in order to purge temperature related variability from the recorded data aiming to capture the presence of damage in the signals. However, in this Chapter, the experimental procedure and the main results are presented.

The Chapter is organised as follows: in the first section a description of the structure is given, while in the second the measurement system details are covered. Then, the procedure of selecting the most appropriate sensor locations is discussed and in the fourth section, the experimental procedure and an overview of the obtained results is provided.

3.2 Description of the Structure

The laboratory truss bridge (Figure 3.1) was manufactured by the Dynamics Group of Sheffield University using $T6082$ aluminium alloy. The truss consists of 17 independent members (Figure 3.1). The horizontal top ties are referred to as top chords, while the horizontal bottom ones as bottom/lower chords. The inclined vertical members are referred to as end posts (side diagonals) and diagonals (internal ones). Finally, the transversal members linking the chords (top and bottom) between them are commonly referred to as struts.

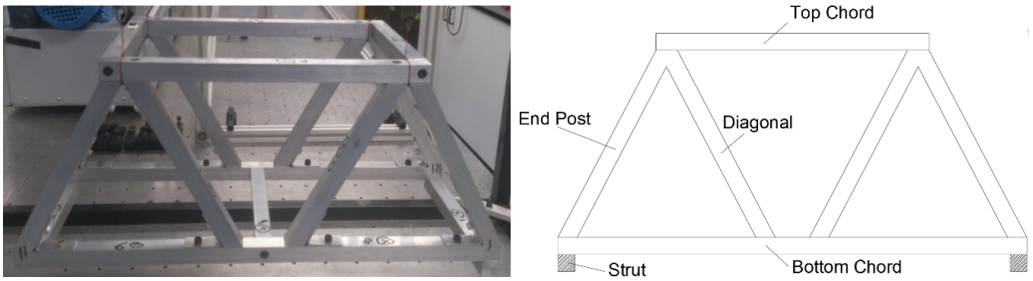


Figure 3.1: (a) Laboratory Testing Structure (Truss Bridge) hanged by four springs (free-free support) and (b) designation of truss members.

Moreover, for vibration testing, the structure is suspended by four threads (high resistance fishing-net threads), which are attached to 4 springs, in the upper face of the truss.

3.2.1 Material, Mass and Dimensions

All members share the same cross-sectional characteristics $20 \times 20mm$, however they have different lengths. The exact dimensions of the truss structure are displayed in Figures 3.2, 3.3, which show the front and bottom plan view of the structure. As for the material properties, all components were constructed using $T6082$ aluminium

alloy, which has an elasticity modulus between $68 - 70GPa$, a *Poisson ratio* (ν) near 0.32 and weight per unit volume equal to $2.7kg/cm^3$. The net weight of the structure (excluding springs, threads and sensors) was measured to be $4.41kg$, each bolt weighted $4gr$, whilst the angle bracket weight is $12.5gr$.

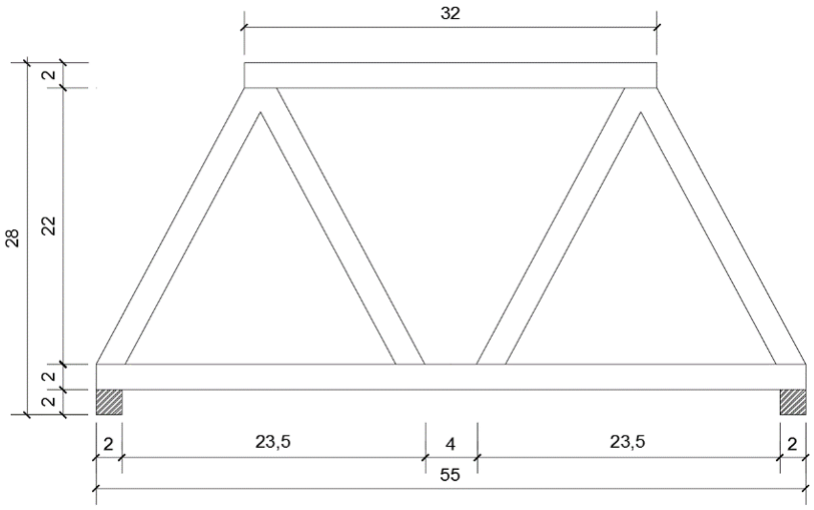


Figure 3.2: Front View of the laboratory truss bridge (in cm).

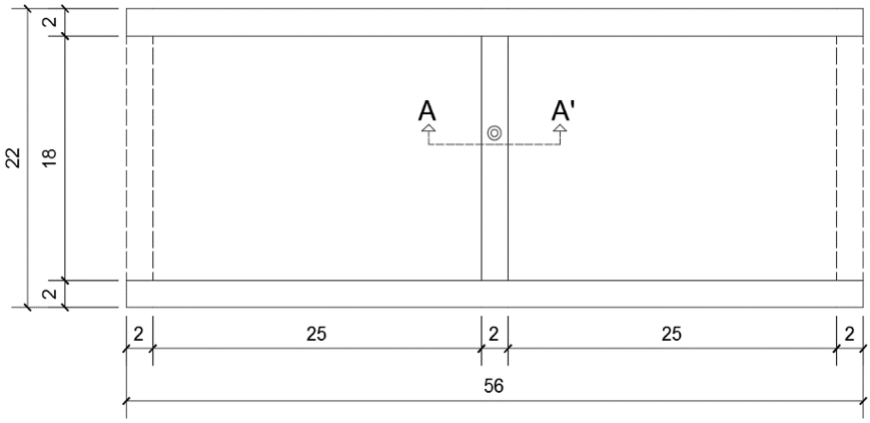


Figure 3.3: Ground level plan View (in cm) and the location of pin (AA'), where shaker is connected.

3.2.2 Connections' Details and Bolts

The members of the tested truss are connected between them using bolts, while angle plates, where bolts are mounted, are used in order to connect attached members, such as the diagonals between them or with chords (Figure 3.4). In particular, 34 bolts are used in the structure, from which 30 are of type $M4$ *Grade* $A2$ (Stainless

Steel, cap head hexagon socket), while the remaining four are of type $A2\ M6-40mm$ (Figure 3.4 detailing).

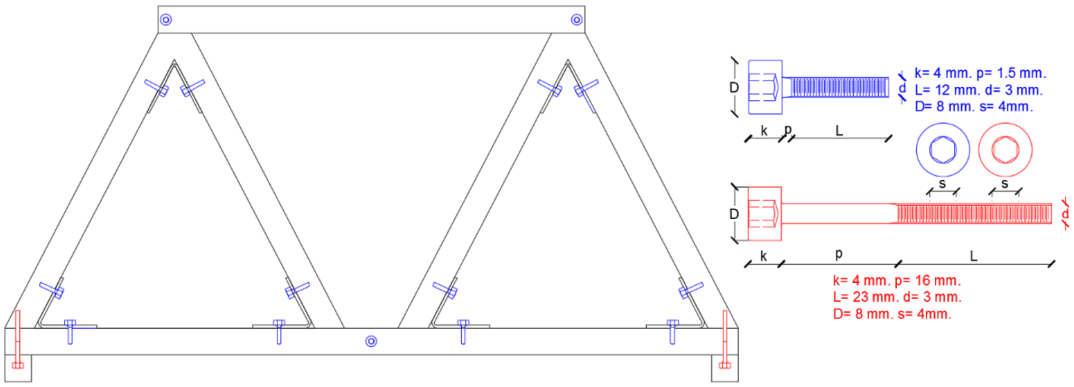


Figure 3.4: Vertical Cut View (Front View) of Truss, locations and types of bolts.

A torque wrench was used for fastening all the bolts at the same level of torque (5.5 Nm), which represents the 75% of maximum torque resistance of a standard $M4$ bolt (according to $DIN7984 - Metric$) [83]. It should be mentioned that additionally, aluminium $M4$ bolts were tested for the experiment, however, their use led to an increase of noise during vibration analysis as a consequence of inadequate fastening (unexpected loosening under temperature variation inside chamber) and sensitivity of cap hexagon heads to increased torque. The selection of stainless steel bolts proved to be the most appropriate for this experiment, both in terms of data extraction and resistance to temperature variation. Furthermore, at the lower face of the plan view (Figure 3.5) and more specifically in the mid-strut, there is a fixed pin located, which is the point where the structure is connected with the shaker. Figure 3.5 provides the detailing of AA' section-cut shown in Figure 3.3, as well as the dimensions of the fixed pin.

Before moving further to system's description, it is important to talk about the connections between independent truss members and especially about the locations where was damage simulated (bolts removal), because it will be useful for the explanation of vibration analysis results (section 3.4). In particular, there are three damage locations, $A1$, $A2$ and B (Figure 3.6), where bolts were removed. Location $A1$, connects three truss members together, the horizontal bottom chord (designated as 2 in Figure 3.6), the vertically inclined end post (external diagonal; designated as 1 in Figure 3.6) and the transverse strut (designated as 3 in Figure 3.6).

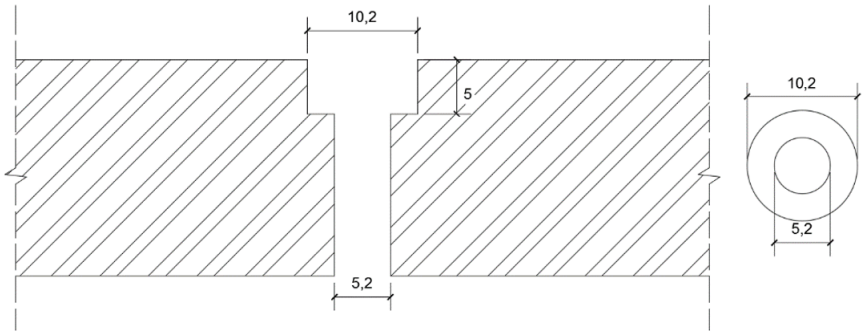


Figure 3.5: Cross-Cut AA' (in mm) of middle strut in the lower face of plan view (see Figure 3.3.)

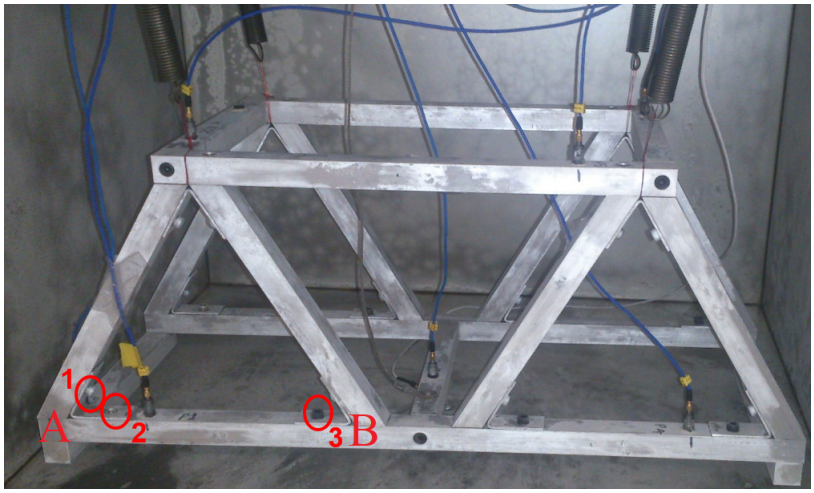


Figure 3.6: Bridge truss model inside chamber and damage locations.

Although there is an angled bracket in damage location A, with two mounted bolts, connecting the end post and bottom chord, members 1 and 2 are connected through a different bolt (with greater pitch; Figure 3.4) which is located at the lower face of the strut. Therefore, removal of bolts mounted in the angled bracket at location A, are not expected to significantly affect the stiffness of the truss, while a low percentage of mass reduction, due to bolts weight, is anticipated. Considering the masses of bolts and bracket, a mass reduction of 0.1% is anticipated for the first damage case and 0.465% for the second.

In contrast to location A, in damage location B, the diagonal and bottom chord of the truss are connected only through the angled bracket. Removal of a bolt at location B is expected to significantly affect the stiffness of the structure (mass slightly) and consequently the modal properties of the truss, since the diagonal tends to slide. Finally, it should be mentioned that the bolts attached to the end

posts and diagonals, mounted into brackets, have an inclined orientation, something that proved to be problematic during fastening them and during the restoration of normal state.

3.3 Measurement System Description

The measurement system used for the experiment consists of two main sub-systems: (*i*) the environmental chamber, in which the structure was placed and was able to provide different temperature conditions and (*ii*) the data acquisition system (DAQs), which includes the sensors, transducers, data analysing system, data storage and the software for data retrieval and processing. Table 3.1 provides a summary of the sub-systems comprising the measurement system used during the laboratory experiment, whilst Figure 3.12 provides the arrangement of the experiment.

Table 3.1: Description of Measurement System

Name	Description	Properties
Piezoeletric Accelerometers	PCB ICP P1-P7 (7)	Sensitivity: ($\pm$30%) $1mv/(m/s^2)$ Measurement Range: ($\pm$5000 gpk) Frequency Range: (0.40 to 10000) Elect. Filter Corn. Freq: 21 kHz Mec. Filter Reson. Freq.: 35 kHz Temp. Range: +66 to -18 ^oC
El/Dyn. Shaker	V406 LDS: MA-CE	2.503 mV/N
Raspberry Pi	DS18B20 Sensor	4.7 $k\omega(Ohm)$Resistor +100 to -40 ^oC
Impact Hammer	PCB 086C03 ICP	Sensitivity: ($\pm$15%) $2.25mv/(N$ Measurement Range: $\pm$2224 Npk Hammer Mass: 0.16 kg
Thermocouple	RS Pro 206-3738 Dual Channel K Type	Best Accuracy: $\pm$0.2% Resolution: 0.1 ^oC Dimensions: $160x64x26$ mm Weight: 430 $gr.$
Amplifier	PA 100E Power Amplifier	Freq. Range: ($1o$ to 200 MHz) High Output Power: +34 dBm Low VSWR:¡1.5 High Gain: +22 dB

3.3.1 Environmental Chamber

The environmental test chamber allows dynamic testing of small-scale structures across a broad temperature range (-45 to 120^oC) using either manually or programmable temperature control. Its internal dimensions are $600 \times 600 \times 700mm$. A port in the base of the cabinet enables electrodynamic shaker excitation in the vertical direction at loads up to $100N$, while there are two available side ports for internal-external cable connections (i.e. DAQs). The chamber of Jonas Lab is depicted in Figure 3.7.

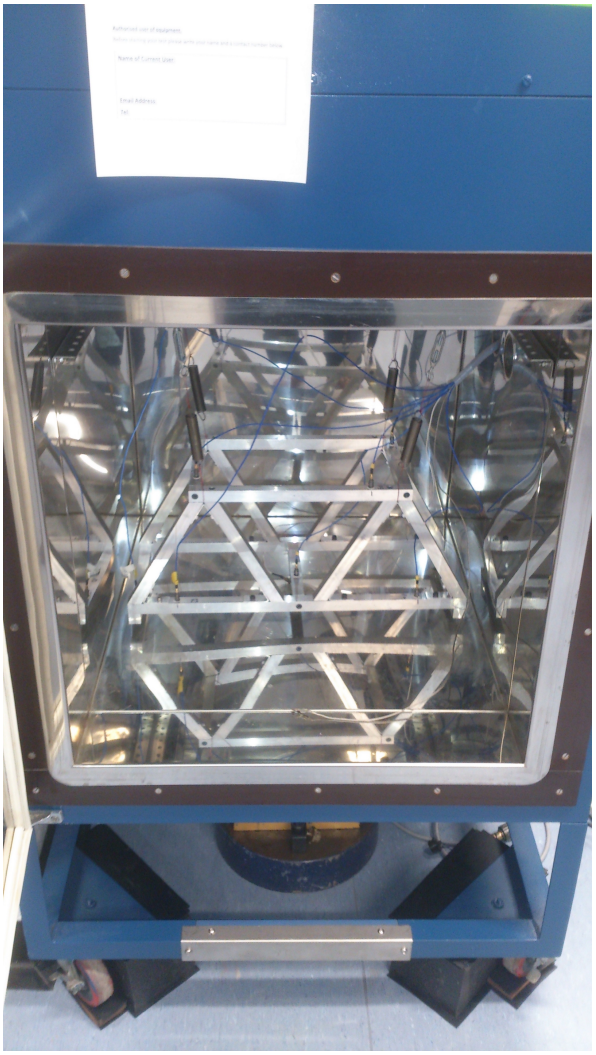


Figure 3.7: Environment chamber in Jonas Lab.

For this analysis, the chamber temperature field was manually controlled for temperatures between -15 to $+25^oC$. The temperature of the chamber and the structure was measured using a thermocouple and a temperature acquisition Raspberry Pi

sensor (provided by the Department of Civil Engineering of Exeter University), to validate that the temperature of the structure is in accordance with the indications of chamber's control panel. The Raspberry Pi has two built-in cable wired sensors, providing measures with accuracy up to 0.001^oC, which are placed, during the experiment, on chamber's wall (inside) and on truss model respectively.

3.3.2 Data Analyser and Processor

The data acquisition and frequency analysis system used is an 8-channel modal testing system (LMS SCADAS). The impact hammer, shaker (through amplifier) and accelerometers (inputs) are connected directly to the LMS system, while the latter is connected via Ethernet port to the central computer, where the data are stored and analysed (see Figure 3.12). The LMS system includes software for data management and processing. In brief words, the acceleration data are recorded and stored in the system and then the modal analysis software, polyreference least-square complex frequency algorithm (PolyMax) [1, 84], is used to extract the modal properties of the structure under different temperature and damage conditions (more information in section 3.3.4).

3.3.3 Sensors and Instrumentation

According to [85], the selection of the most suitable transducers should be based on the consideration of some parameters, such as: (i) their sensitivity, (ii) resonant frequency, which depends on the mass and the size of the transducer and (iii) the effect that the motion of a transducer can have to a structural system, especially in cases of small-scale systems, which can lead to additional inertia forces and moments. For this experiment, piezoelectric accelerometers PCB used, of sensitivity around $1mV/(m/s^2)$ and temperature resistance between 66 to -18^oC. The sensors are small and lightweight to minimise their affect on structural response.

3.3.4 Extraction of Modal Properties

The procedure for acquiring the modal properties of the structure includes two main steps. Firstly, the frequency response functions (FRFs) during the excitation

are recorded and then used employing the available software to extract the modal properties.

The FRFs obtained were estimated using a random signal of 2V generated in LMS (amplified by a 1.5 level gain using the amplifier) and transmitted to the shaker which is connected through a port in the base of the cabinet and fixed in its upper end in middle bottom strut (Figure 3.12). The sampling frequency was $2048Hz$, while a *hamming window* used to account signal *leakage* present due to non-periodic random signal. The bandwidth selected was $1024Hz$, with 4096 spectral lines, providing a resolution of $0.25Hz$. The approach followed here for the estimation of FRFs is the H_v. For this, random noise and distortion are assumed to be present both in the input and output of the system [85]. The FRF is calculated through the relation between the excitation and response of the system, as described in Equations (3.1), (3.2) and (3.3), where $S_{xx}(\omega)$ is the response auto spectra, $S_{ff}(\omega)$ is the auto spectra of the excitation signal and $S_{xf}(\omega)$ is the cross spectrum between the two signals. The H_v is provided according to Equation (3.4) [84].

$$S_{xx}(\omega) = \left|H(\omega)^2\right| \times S_{ff}(\omega) \tag{3.1}$$

$$S_{fx}(\omega) = H(\omega) \times S_{ff}(\omega) \tag{3.2}$$

$$S_{xx}(\omega) = H(\omega) \times S_{xf}(\omega) \tag{3.3}$$

$$H_v = \frac{S_{xf}(\omega)}{|S_{xf}(\omega)|} \times \sqrt{\frac{S_{xx}(\omega)}{S_{ff}(\omega)}} \tag{3.4}$$

The second step includes the PolyMax algorithm [1]. The PolyMax was selected because it is the industrial standard and was incorporated inside the LMS Scadas environment. This employs as primary data the FRFs determined above, which are processed and then using stabilization diagrams, the natural frequencies, damping ratios and modeshapes are obtained. A brief description is provided below.

The FRFs representation is assumed according to right-matrix fraction model, Equation (3.5), where H_ω is the FRFs matrix of size mxl (m inputs and l outputs), $[\beta_\gamma]$

describes the numerator matrix polynomial coefficients, $[\alpha_\gamma]$ is the denominator matrix polynomial coefficients, p describes the order of the model, while z is the z-domain model (frequency domain model) based on time discretised model $z = e^{j\omega\Delta t}$, where Δt is the sampling time and ω the frequency axis of FRFs.

The unknown parameters are the model polynomial coefficients $[\alpha_\gamma]$ and $[\beta_\gamma]$, which are obtained using the ordinary least-squares (OLS) method and used to obtain the poles and modal participation factions, needed to construct stabilization diagrams [1]. Employing appropriate stabilization criteria the natural frequencies, damping ratios and modeshapes can be obtained. More information about the derivation and application of Polymax algorithm can be found in [84].

$$[H_\omega] = \sum_{\gamma=0}^{P} z^\gamma [\beta_\gamma] \left(\sum_{\gamma=0}^{P} z^\gamma [\alpha_\gamma] \right)^{-1} \tag{3.5}$$

$$[H_\omega] = \sum_{i=1}^{n} \frac{\{\nu_i\} \times \langle l_i^T \rangle}{j\omega - \lambda_i} + \frac{\{\nu_i\} \times \langle l_i^H \rangle}{j\omega - \lambda_i^*} - \frac{[LR]}{\omega^2} + [UR] \tag{3.6}$$

$$\lambda_i, \lambda_i^* = -\xi_i \times \omega_i \pm j\sqrt{1 - \xi_i^2 \omega_i} \tag{3.7}$$

The modal parameters are estimated following the pole residue model, described in Equation (3.6), where n describes the number of modes, $\{\nu_i\}$ the vector of mode shapes, l_i^H, l_i^T the complex conjugate transpose and transpose of modal participation factors. The poles (λ_i,λ_i^*) occur in complex conjugate pairs associated with natural frequencies (ω_i) and damping ratios (ξ_i) according to Equation (3.7), while $[LR]$ and $[UR]$ describe the lower and upper residuals, which models the influence of the out-of-band modes in the frequency band considered. The LLS is employed to identify the unknown parameters ν_i , $[LR]$ and $[UR]$.

3.4 Description of Experimental Procedure

The experimental procedure included two main steps. In the first step, the roving impact hammer test was employed aiming to identify the most appropriate locations for placing the accelerometers. After selecting these locations, the second step

includes the main experimental procedure, which involves the introduction of the truss model inside the environmental chamber and the vibration analysis.

3.4.1 Roving Impact Hammer Test

To understand optimal locations for accelerometers a roving hammer test was conducted. The locations are chosen on this basis as those that provide the majority of the FRF peaks and consequently natural frequencies. During this procedure, many different locations are excited using the hammer (this excites the structure by introducing an impulse force). In Figure 3.8, an impact hammer is depicted, consisting of a head, an force cell and a tip.

Figure 3.8: Impact Hammer, including different tips, cable and additional mass [1].

The impact magnitude (energy amplitude level) is a function of the hammer head mass and the velocity, in which the hammer hits the structure (linear momentum). Due to the fact that is difficult to control the velocity of the hit, the hammer head can be adjusted appropriately by adding or removing mass. On the other hand, the frequency content (magnitude and phase response) of the energy applied to the structure is a function of the contacting surface stiffness and the hammer mass. The shape of force impulse can be affected by the contact surface stiffness. Due to the fact that the stiffness of the test structure cannot be changed, the hammer tip can be adjusted (i.e. made stiffer). In particular, the harder the hammer's tip, the shorter the impulse duration and the higher the frequency content, which means that in order to excite higher frequencies harder tips are needed. In general, the tip should be selected in such a way that the force spectrum amplitude is no more than 10-20 dB down at the maximum frequency at interest [85].

An important decision, at this point, is the selection of the most suitable accelerometer in terms of sensitivity. In a general sense, a good choice can be an accelerometer

with high sensitivity, having a resonant frequency, which depends on the mass and the size of the transducer, far from the structure frequencies and the size of structure tested. The latter is associated with the effect that the motion of a transducer can have to a structural system, especially in cases of small systems, which lead to additional inertia forces and moments. For this experiment, a piezoelectric accelerometer PCB is used with a sensitivity of 1.063 mV/(m/s^2).

Furthermore, another two important issues are the attachment and the location of the accelerometer. The most appropriate location of the accelerometer should be identified prior conducting the roving hammer test, the accelerometer was attached in different locations using a thin layer of wax. This was selected instead of superglue in order to test many different locations on the structure. However, it should be mentioned that the wax cannot provide such a good contact (stiffness) between the accelerometer and the structure. In particular, the wax tends to deform causing change in the orientation of the accelerometer, which is not positioned normal to the structural face, providing poor contact. This can lead to a noise increase observed during the measurements.

The location of the accelerometer is also an important issue. In essence, the accelerometer should be located in a position, which is not close to a node (location of zero motion) of a structural modeshape, in order to provide an effectively measured point. An approach that can be used in order to obtain the most appropriate location of accelerometer is to use the point mobility test and examine the measured FRFs (selecting the accelerometer providing the most resonant frequencies).

The impact hammer test requires little hardware and is relatively inexpensive, there is a difficulty to obtain consistent results. As a matter of fact, there are two particular problems, which affect the results quality. The first is noise, which can be present in either the force or response signal as a result of a long time record. Moreover, leakage can be present in the response signal as a result of a short time record, due to the fact that the force impulse is very short relative to the time record length. For both aforementioned problems, windowing techniques can provide compensation.

The first step of the analysis includes the determination of system inputs, which are the hammer, the accelerometer and the units. Based on the manufacturers details, for a hammer PCB 086C03 ICP with actual sensitivity 2.25 mV/N. and an accelerometer PCB ICP with sensitivity 1.063 mV/(m/s^2). Next, the bandwidth and the spectral lines employed should be defined. Considering the latter, a plastic

tip was selected in contrast to the metal one (excites higher frequencies/ more energy to the system), since the former can provide enough energy to excite the frequencies of interest. The bandwidth selected was 1024 Hz, with 2048 spectral lines, providing a resolution of 0.5 Hz.

Furthermore, the trigger settings were determined, verifying the bandwidth selection and apply the most appropriate windowing function to avoid problems with regards to noise and leakage. In this procedure, the structure is excited, in order to determine the percentage of input's cut-off frequency and the percentage of exponential decay for the response (acceleration). In this case, an exponential window function is used, due to the fact that the impulse test is a transient vibration technique, where much of the important information of the time record is concentrated and suppressed in its initial part.

Finally, before conducting the roving impact hammer test, it should be added that because the structure is suspended using threads (to provide a free-free test set up), after every hit sufficient time should be left for the structure to become stable (not oscillate) before the next hit.

3.4.2 Proposition for the most Appropriate Accelerometer Locations

The most appropriate locations of accelerometers were found by examining the measured FRFs (most resonant peaks), as well as the stability diagrams used during the modal identification [1, 84]. This analysis is performed inside LMS Scadas, employing the PolyMax algorithm (section 3.3.4)

The roving hammer test is used here to obtain the FRFs of 24 different structural locations. For reasons of space only, the 12 most appropriate locations are presented here. The 12 locations tested are depicted in Figure 3.9, whilst the obtained FRFs are depicted in Figure 3.10 and 3.11. From analysing each one of the FRFs, those that provide the majority of the resonant peaks are $P1$, $P2$, $P3$, $P4$, $P5$, $P7$, $P9$ (Figure 3.9, circled in red colour).

Moreover, the LMS Scadas used in order to test if the selected locations of accelerometers can provide good results in terms of stabilisation diagrams, which means that the poles informing the presence of resonant frequencies were appropriately dis-

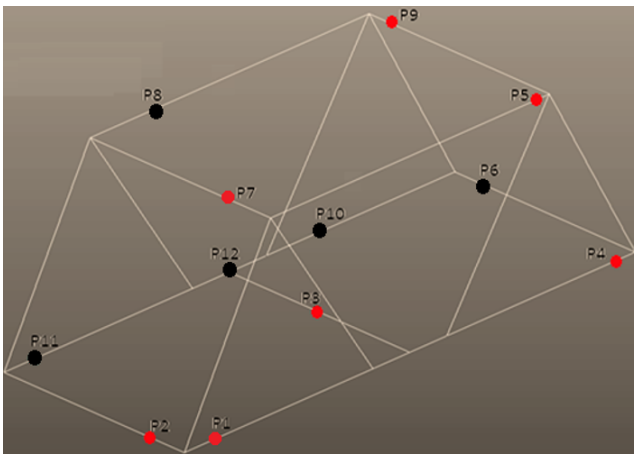


Figure 3.9: Accelerometer locations tested using the roving hammer method. The red nodes selected for further analysis.

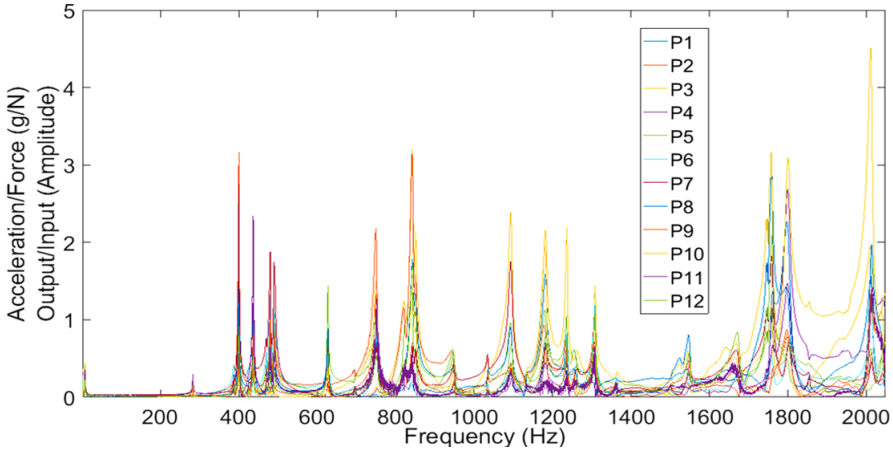


Figure 3.10: The measured FRFs of the 12 tested points using the roving impact hammer.

played.

3.4.3 Vibration Analysis inside Chamber

As the locations of the accelerometers have been determined, then the truss bridge model was introduced inside the environmental chamber. In particular, Figure 3.12 provides the experimental schematic, which shows the arrangement of the experimental setup. In brief the structure was hung from a hanger grid inside the chamber using fish-net threads and was mounted on an electrodynamic shaker using a driving rod, through a base port (CC'). The accelerometers and temperature cables were connected on an external central computer through a side port available on the right

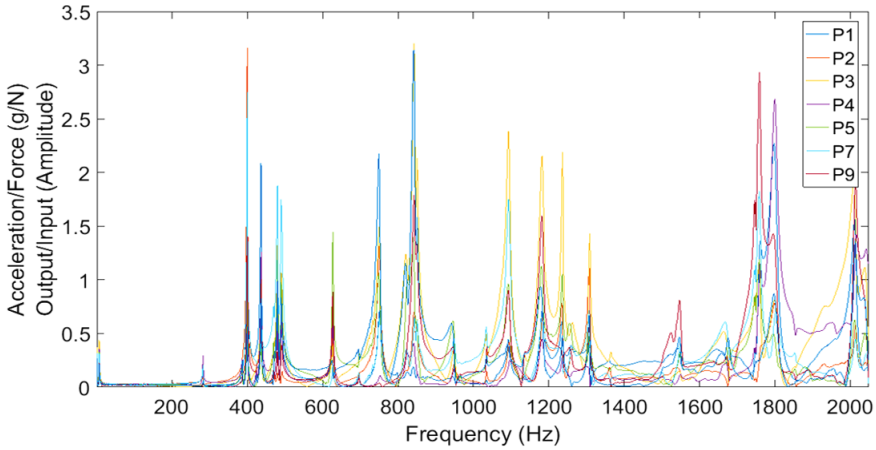


Figure 3.11: The measured FRFs of the most appropriate 7 locations using the roving impact hammer.

side of the environmental chamber.

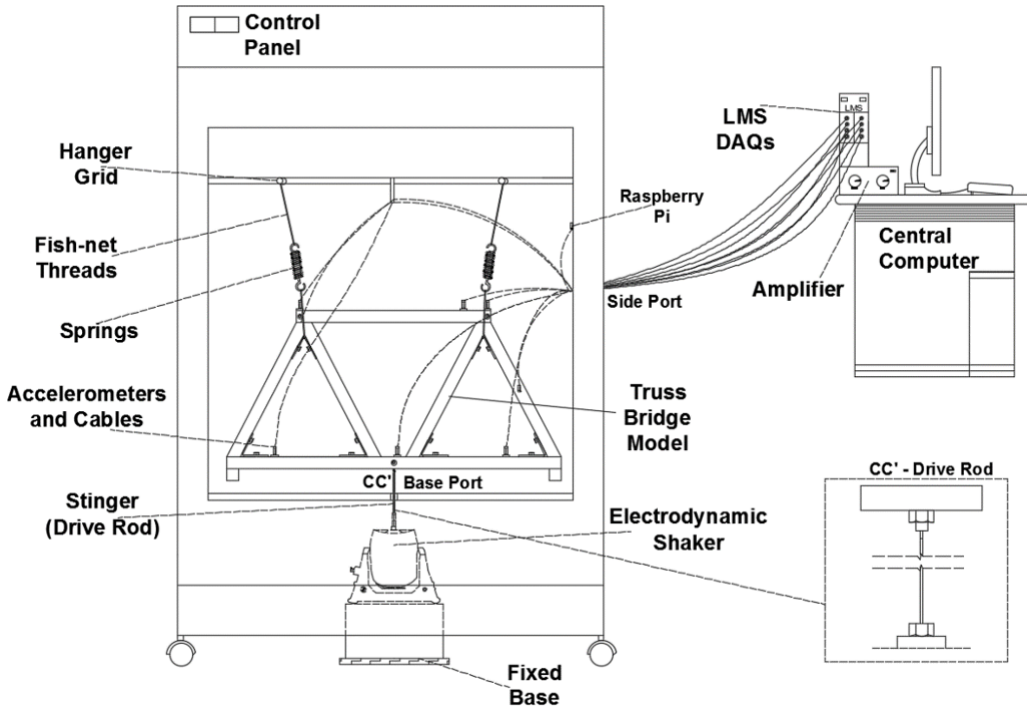


Figure 3.12: Experimental schematic/arrangement inside environmental Chamber.

To begin with, five states were determined and measured as depicted in Table 3.2. In particular, the first state is considered as the 'normal state' of the structure. The second state is the first damage state, in which one bolt was removed at damage location A (Figure 3.13). The third state describes the second damage, where both

bolts removed at damage location A (Figure 3.13) and the angled bracket attached to bolt holes. Next, in the fourth state three bolts removed, two in damage location A and one in damage location B (Figure 3.13). The last state describes the attempt to restore the normal condition of the structure.

Table 3.2: Structural states during laboratory experiment.

State	Location	Description
Normal	-	Normal State
Damage 1	Location A	Bolt Removal
Damage 2	Location A	Both Bolts removed + angled bracket
Damage 3	Location A&B	Three Bolts Removed + Angled Bracket
Restored	Location A&B	All Bolts & Bracket re-placed

Each one of the structural states includes 160 measurements and corresponds to one day of recording. In particular, twenty accelerations records were measured for each temperature level (8 levels) starting from 25 to -10^oC with a step of -5^oC. Each acceleration record duration was $60s$, averaging every $4s$ (15 averages per measure), to achieve higher precision. Temperature was also recorded every $4s$ and each time its mean value is used as correspondence to each acceleration record.

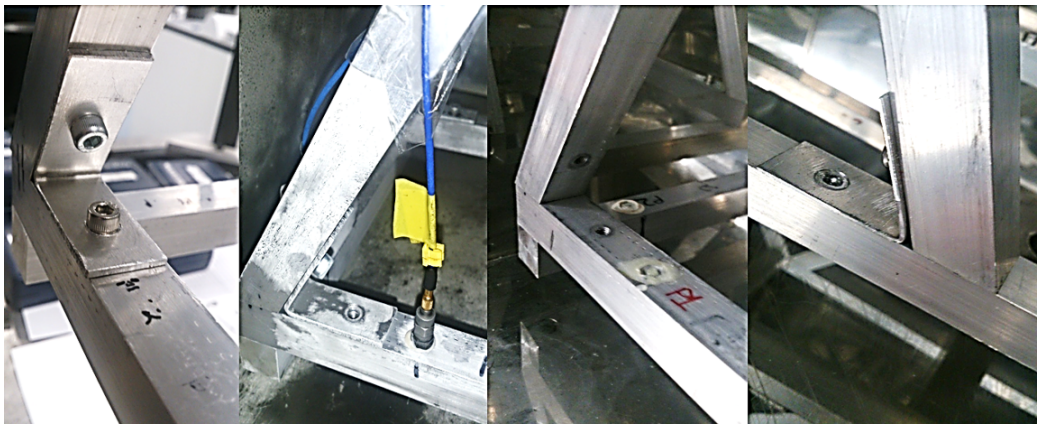


Figure 3.13: Undamaged state, first bolt removal, two bolts and bracket removed and finally, third bolt removed.

During the set up for each state, the bolts are removed and tightened up at temperatures around 25 ^oC. All bolts have the same pretension level, torque $5.5Nm$, verified before and after conducting the experiment. In addition, it was important to ensure that when a temperature cycle is performed (25 to -10 and back to 25^oC) the structure was left enough time for the structural temperature to stabilise (thermal inertia). The accelerometers are attached to the structure using $X60$ fast curing adhesive, which provide good contact and stability to the accelerometers.

3.5 Summary of Experimental Results

In Figure 3.14, the measured FRFs at a temperature of 19.985 ^oC are illustrated. These are obtained from a trial recording. Figure 3.15, shows the stabilisation diagrams used to obtain the natural frequencies inside the environment of LMS SCADAS software.

From the second plot, it is clear that poles can be identified, however, it can be observed that noise is also present during the measurements. Apart from the expected noise from the sensors cables, the factors that were identified to affect the level of noise, were: (i) the connection between truss members and the connection between bolts, holes and brackets. Although these cannot be perfect, the structure was reassembled twice to improve the noise levels. Then, (ii) the threads used for hanging the structure affected the noise level, and due to this many different fish-net threads with different resistance were trialled. Next, (iii) the connection between the model and the shaker is examined. This is done using a driving rod (stringer), whose cross-sectional area is quite small in order to fit in both impedance heads. The latter can be seen in Figure 3.12. However in each impedance head, the stinger should be partly-fixed, something that is very difficult to achieve. This is needed because otherwise fixing the rod in impedance head can deform the rod and change its orientation; and consequently can change the orientation in which the signal is transmitted from shaker to model.

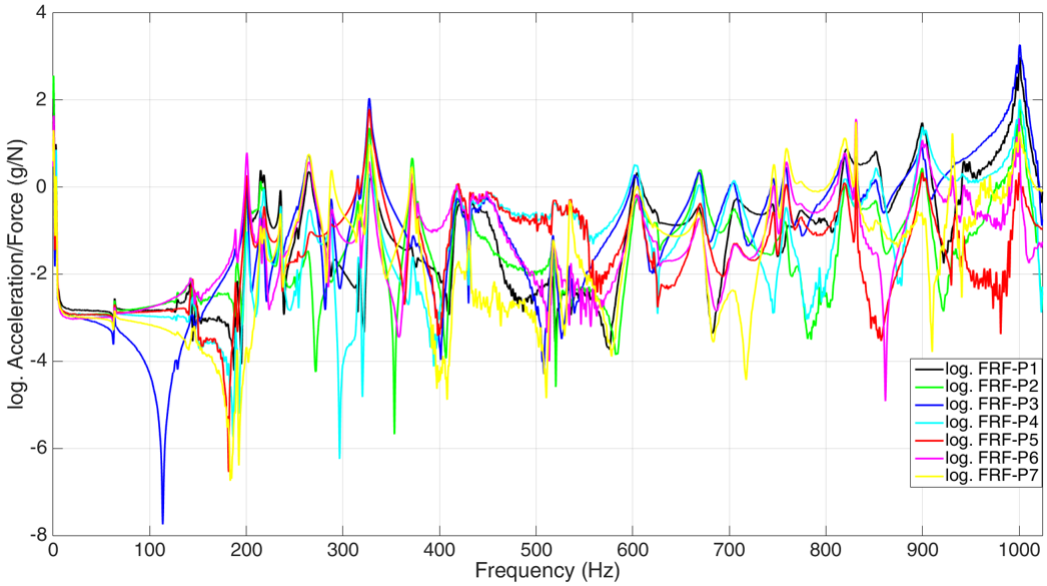


Figure 3.14: Measured FRFs from a trial recording.

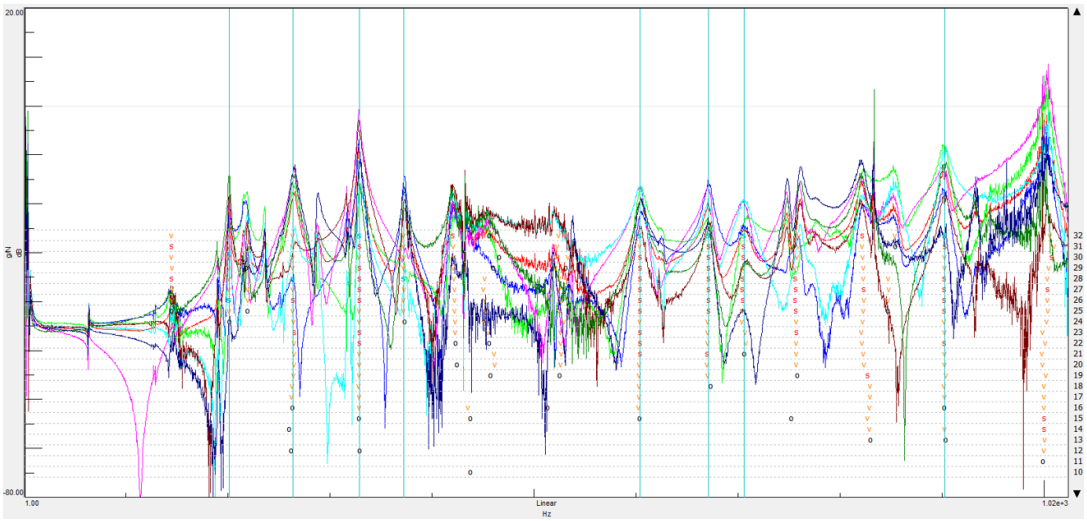


Figure 3.15: Polymax algorithm inside LMS SCADAS environment for stable poles detection.

Furthermore, (iv) the attachment of cables and accelerometers with the model's surface should be sound to provide good quality measures. In terms of acceleration recordings and natural frequency values extraction, it should be mentioned that the chamber during its function, sometimes makes slight vibrations when moving from one temperature to the other, even for small temperature difference. In order to mitigate this unavoidable vibration additional recordings were performed.

Finally, for every temperature step, the temperature was kept inside a $\pm 0.5^oC$ range. This means that all recordings, i.e. for 20^oC, are measured between 19.5 and 20.5^oC. In order to perform such a temperature stabilisation manually, re-heating and re-cooling was needed.

3.5.1 Results

The results that will be used for further analysis in this thesis are the first five natural frequencies obtained by the vibration analysis, which are illustrated in Figures 3.16 to Figure 3.20. The plots show the five different structural states, as well as the impact that temperature has both in long-term (25 to -10^oC) and short-term (i.e. 25 $\pm 0.5^oC$).

As for long term temperature impact, a trend over time can be observed due to increased stiffness, while for each temperature step (i.e. 25 to 20^oC), a mean shift

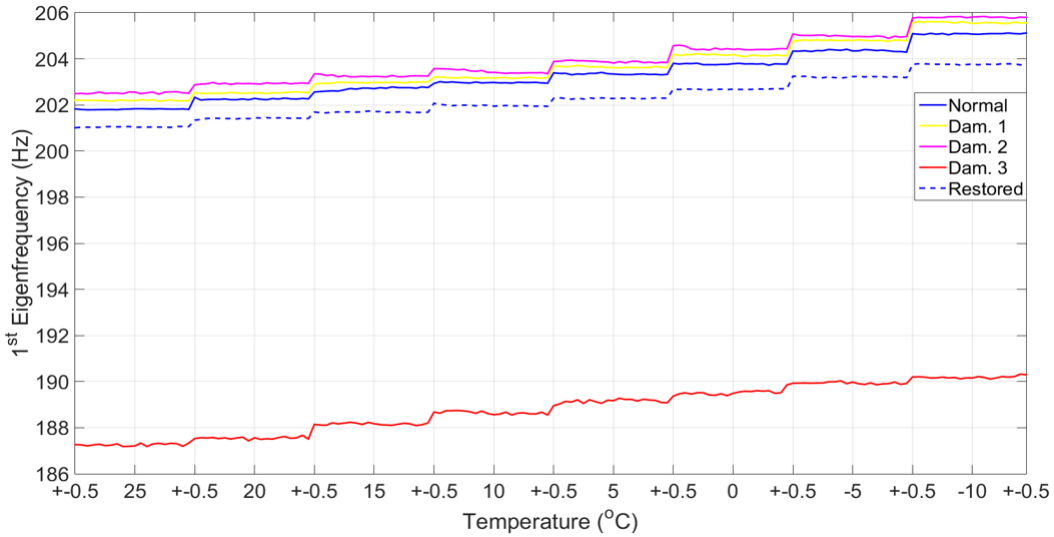


Figure 3.16: First natural frequency series for all states of experiment.

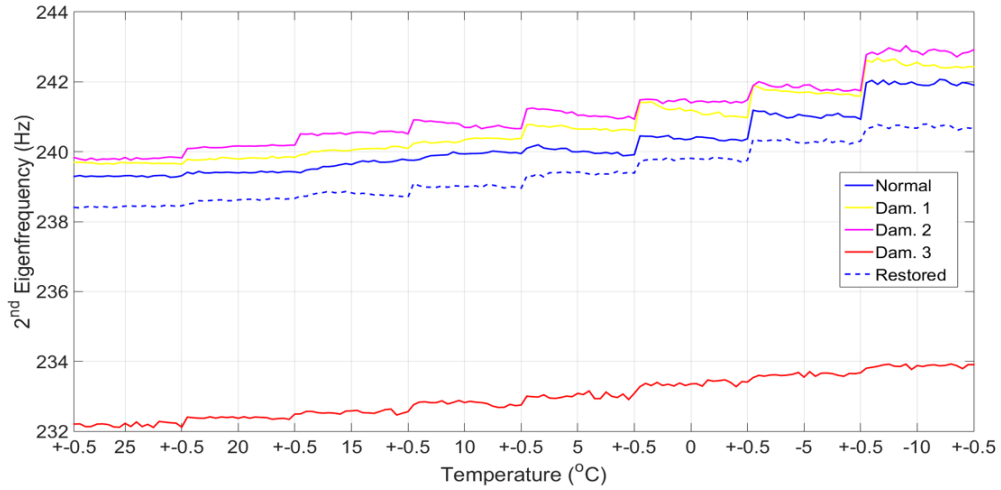


Figure 3.17: Second natural frequency series for all states of experiment.

is introduced to the data. On the other hand, the short-term effect of temperature is associated with the way that natural frequencies change inside each temperature step and is observed in all plots. The latter can be also associated with the variation of stiffness due to temperature change.

Furthermore, the different damage states can be observed clearly. In particular, the first two damage states are close to the normal state condition, showing that the frequencies increased. This observation can be explained by the fact that the bolts removed at damage location A, did not affect the stiffness of the structure and the change in frequencies can be associated with bolt mass and bracket removal (Figure

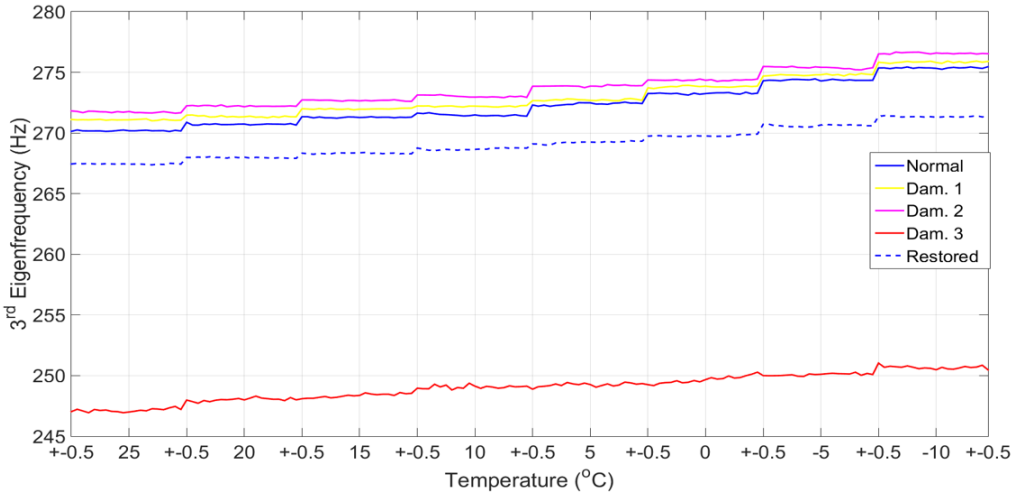


Figure 3.18: Third natural frequency series for all states of experiment.

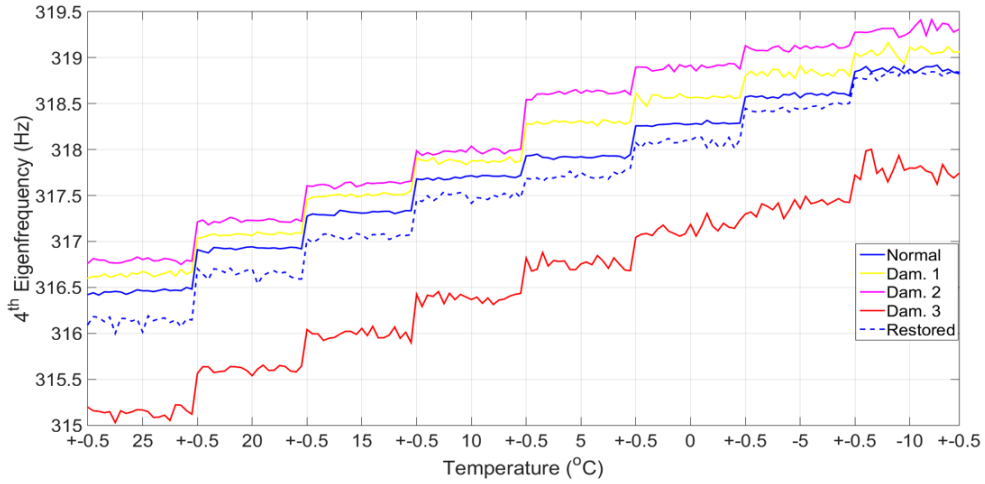


Figure 3.19: Fourth natural frequency series for all states of experiment.

3.12). In particular, a mass reduction of approximately 0.1% is expected from the first damage case, as the mass of a bolt is 4gr. and that of the whole structure is 4410gr. In addition, for the second damage, a mass reduction of approximately 0.45% was expected, since two bolts removed (8gr) and an angled bracket (approx. 12.5gr.). Assuming an average natural frequency value of 201.816Hz, a 0.1% mass reduction gives 202Hz, whilst a 0.45% gives 202.758Hz. These two values are quite close to those observed in Figure 3.17.

However, for the third damage a completely different result is found. In particular, damage 3 is the most severe and led to decrease of natural frequency, because the removal of the bolt led to truss diagonal sliding and change in connection at damage

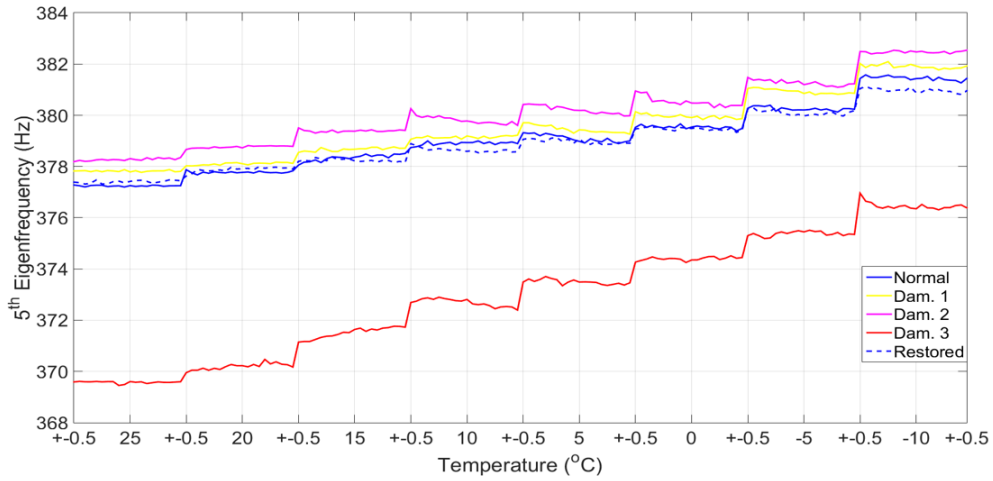


Figure 3.20: Fifth natural frequency series for all states of experiment.

location B. Finally, in all cases and tests, the results show that the system was restored successfully, however exhibiting lower natural frequency values. This shows that although the system is practically restored, it does not retrieve its exact same response, something that can be caused by the change in bolts tension and the impact of temperature cycles on the structure.

Finally, the repeatability of the results was tested. In particular, the system was brought to the initial normal condition, then all three bolts and angled bracket removed and put back on. This procedure repeated 8 times. For repeatability purposes, the average values of natural frequency at 25^oC were recorded, as shown in Table 3.3.

Table 3.3: Repeatability of Experiment for natural frequency values at 25^oC.

Reference	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}
Test	201.816	239.300	270.183	316.457	377.245
Repeat 1	202.756	239.811	270.167	316.486	375.047
Repeat 2	202.469	239.442	271.672	315.833	376.180
Repeat 3	202.208	239.178	271.419	315.618	375.248
Repeat 4	201.491	238.483	271.490	315.713	374.904
Repeat 5	200.930	238.359	267.658	315.797	377.957
Repeat 6	200.863	238.008	268.076	315.672	376.373
Repeat 7	200.568	238.291	266.373	315.008	374.814
Repeat 8	200.926	239.308	266.766	315.002	376.494

3.6 Conclusions

The aim of this Chapter was to describe the experimental setup and testing methodology followed in order to test the vibration-based response of a laboratory truss bridge model under varying temperature and simulated damage conditions. In particular, the procedure followed and presented included: *a*) the model description: material, mass, dimensions, connections details, bolts and damage locations, *b*) the measurement system description, which involved the environmental chamber, the data analyser and processor, the shaker and impact hammer (for vibration analysis), the transducers (temperature, acceleration) and their appropriate locations. Furthermore, *c*) the experimental procedure followed described including information about the FRFs estimation and the considerations faced during the extraction of modal properties.

The novelty in this chapter is that a new data set was generated, laboratory truss bridge vibration analysis data, that can be used inside this thesis for methodology development and further future research.

In the next Chapter, a description and a preliminary analysis, based on surface modelling, will be performed on the data sets of Tamar Bridge and Z24 Bridge.

Chapter 4

Real-Time SHM Data for Analysis

4.1 Introduction

The current Chapter aims to present the available structural health monitoring (SHM) data coming from two real-time campaigns (Tamar Bridge and Z24 overbridge), which will be used inside this thesis for further analysis and methodologies development on the following chapters. Besides, the presentation of the available data, a preliminary analysis will be performed on them, aiming to investigate the relations between the SHM series and any available environmental and operational (EO) measures, through simple plotting and first order linear regression (surface modelling). The latter investigation is very useful in order to establish an informative background about the SHM series available and the sources affecting their variability over time, before conducting a more advanced analysis on the upcoming chapters.

The Chapter includes: a description on Tamar Bridge, the available SHM data from it, especially on cable tension system and a preliminary analysis on them, a description on Z24 overbridge available data and a preliminary analysis on the natural frequency signals with respect to EO measures.

4.2 Tamar Bridge Description and SHM Data

4.2.1 Tamar Bridge Description

The Tamar Bridge is a three-span cable-stayed suspension bridge, extending over a total length of 643m, as a continuous stiffening girder of externally anchored type with vertical hangers arranged at 9.1m intervals. The bridge carries the $A38$ truck motorway passing over the River Tamar and links Saltash in Cornwall and Plymouth in Devon. The bridge was constructed by Cleveland Bridge & Engineering Company from 1959-1961 [86]. Due to the continuing increase of traffic requirements and in order to improve aero-dynamic behaviour, the bridge was upgraded (Hyder Ltd. Consulting) and re-opened on April 26^{th} of 2002.

In its current form [87], the superstructure of the bridge consists of a light-weight orthotropic open steel box truss, carrying three main traffic lanes, while extending in transverse direction by two side cantilevers (3m width each), which carry a supplementary traffic lane, parapets and pedestrian-cycle footpaths.

Cable saddles support six main suspension cables (3 North, 3 South) at the top of the two towers, which are anchored externally through mounted splay saddles into the anchorage. A total of 120 vertical hangers are fastened to the main 6 cables over the deck steel truss. Seventy two hangers are located on the main span (36 South and 36 North) and 24 at each side span. The main load carrying system consists of suspension cables with a diameter of 350mm, each includes 31 locked coil wire ropes, while a secondary stay-cable system reliefs the main suspension cables (to avoid over-stressing) during heavy traffic loads. Finally, the substructure of Tamar Bridge consists of two non-prismatic (varying cubic area per height) hollow cross-section reinforced concrete towers of portal shape sitting on caisson foundation founded on rock subgrade.

The secondary stay-cable system consists of 18 stay cables (4 around Saltash Towers $S1$, $S2$, $S3$, $S4$, another 4 around Plymouth Tower $P1$, $P2$, $P3$, $P4$, another 8 on the south side Towers and two additional longitudinal under the bridge deck) (Figure 4.1). Some of them are extended from the Tower saddles to internal anchors on the deck truss ($S2$, $S3$, $P2$, $P4$), while others are anchored externally on the abutments ($S1$, $S3$, $P1$, $P3$).

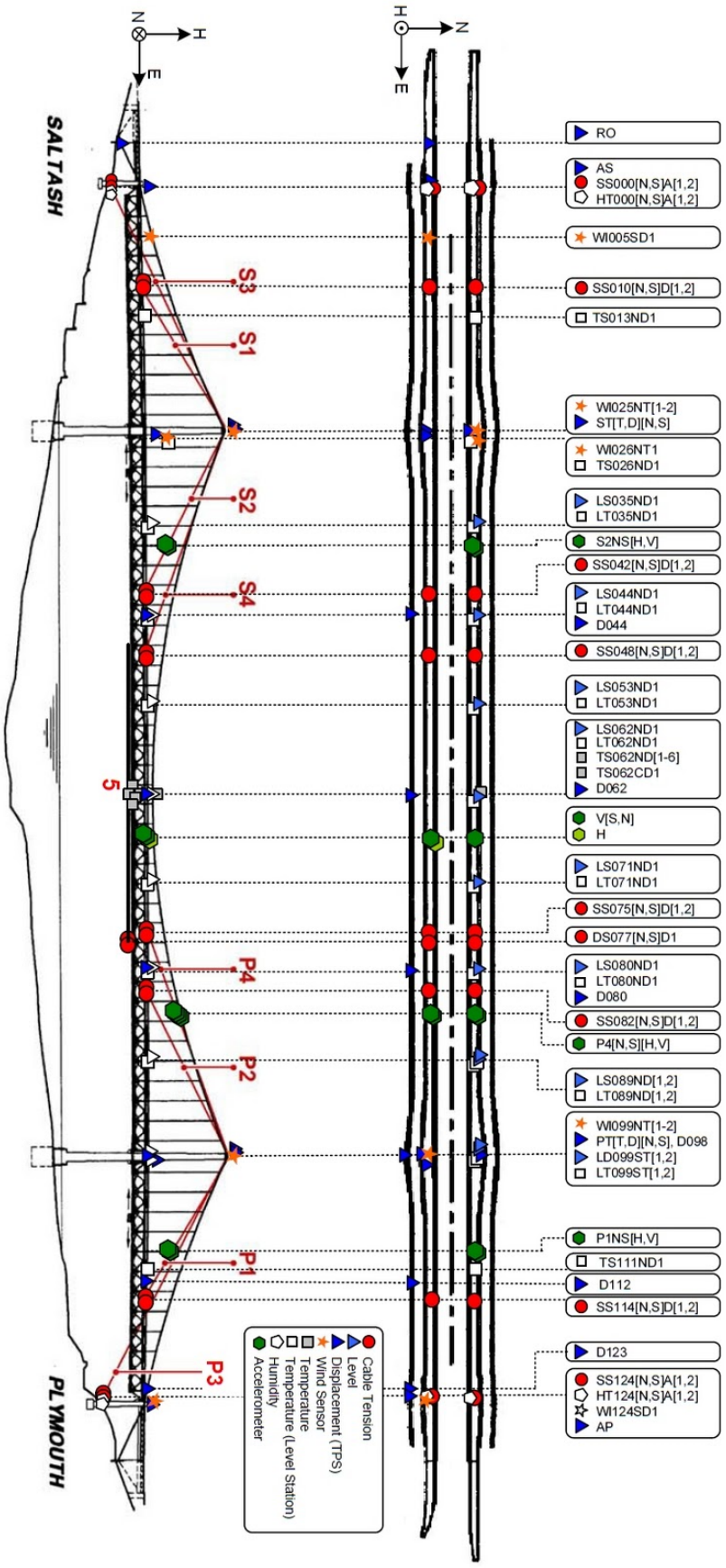


Figure 4.1: Locations of sensors placed on the deck of Tamar bridge [2].

4.2.2 Available SHM Data

The monitored data are provided by the SHM campaign organised by the Vibration Engineering Section (VES) (Department of Civil and Structural Engineering at the University of Exeter [2, 3]). This data includes recordings of deck accelerations, deck vertical displacements (fluid pressure-based level sensing system), natural frequencies (Honeywell QA750 accelerometers; at 8Hz sampling frequency; Figure 4.2), cable tensions (extensiometers in Figure 4.3; resistive strain gauges), damping values, deck and towers displacements (Leica robot total positioning system; Figure 4.4).

Figure 4.2: Accelerometers placed on deck and cable $P4$. From left to right i) horizontal in south deck, ii) vertical in south deck, vertical in north deck and iii) horizontal placed on cable $P4$ [3].

Figure 4.3: Extensiometers located at the north, center and south expansion joints in the bridge deck at the Saltash Tower [3].

Furthermore, environmental and operational features are monitored, such as air temperature, traffic mass, wind speed and percentage of relative humidity. In particular, the monitored air temperature was extracted from a sensor consisting of a temperature probe with radiation shield, sampled per 1Hz in degrees of Celsius (^oC). The wind speed sensor was located at the top South of Plymouth Tower (Figure 4.1), consisted of an anemometer and a vane, recording in miles per hour (mph) at a sample frequency of 1 Hz. The percentage of atmospheric humidity was measured by a hygrometer (Fugro) at sample frequency of 1 Hz. The traffic count was approximated from a toll, which monitored the traffic flow in one direction (east bound)

Figure 4.4: Leica TCA 1201M (Robotic total station) located on the South West corner of the roof of the bridge offices and reflectors placed on Plymouth and Saltash side Towers respectively [3].

across the bridge, providing the vehicle numbers per category (from motorcycles up to 4-axle heavy trucks) based on gross-weight estimations [2]. Further information about the monitoring campaign on Tamar Bridge can be found in [3].

4.2.3 Cable Tensions

The main measurements that will be used in the upcoming Chapters in order to demonstrate and discuss the methodologies presented in this thesis are the stay cable tensions of Tamar Bridge. In particular, there are three types of stay-cables. Firstly are those anchored on either sides of the deck, where tension is measured by resistive strain gauges attached onto the main tensioning bolts at the deck anchor points. The second type represents stay-cables anchored externally on abutments or inside the ground, whilst the third type are longitudinal stay-cables located under the deck, which are installed in order to restore creep deformations during the construction stages of Tamar Bridge [88]. In the latter, tensions are measured using load displacement strain gauges (LDSG) mounted onto the cable. The monitored data are measured in KN and extracted at a sampling frequency of 1Hz [3].

Finally, it should be mentioned that some cable tensions are expected to have a proportional relationship to temperature variability and some others inverse proportional. This depends on the orientation in which the strain gauges are installed. For example Figure 4.5 depicts two cases, one for stain gauges mounted on tower anchorage and one for those mounted on deck level anchorage.

After presenting the available SHM data of Tamar Bridge and a brief description of the cable system is provided, it seems beneficial to study and display the effect that

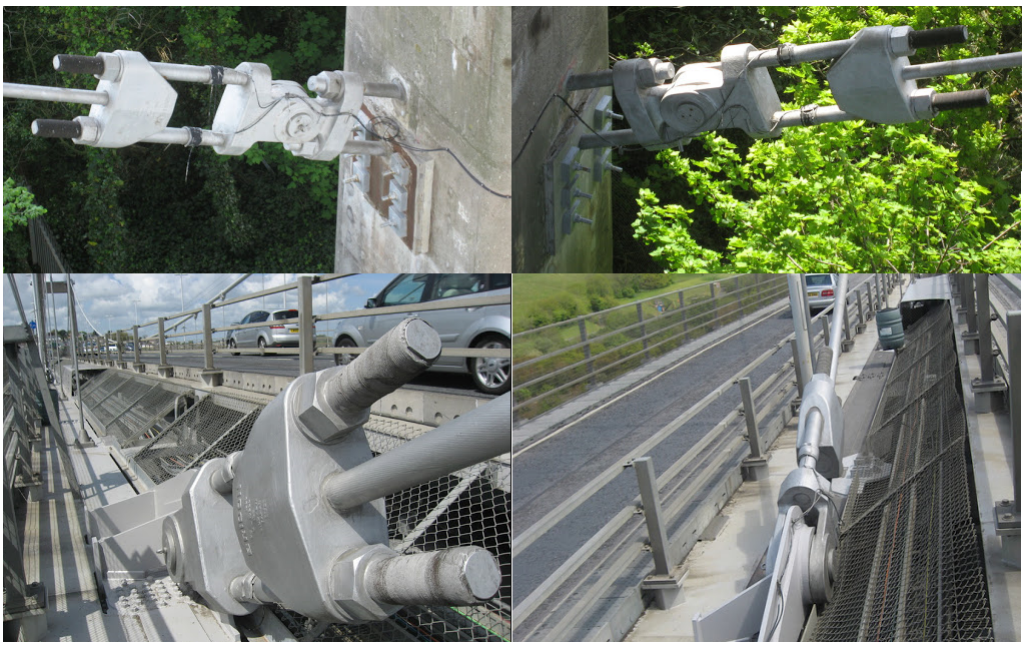


Figure 4.5: Tension sensors placed at the anchorage of the South side and the deck level of the north side of the bridge [3].

measured EOVs available impose on the tension signals of the Tamar Bridge. The latter information can be very useful in order to understand the characteristics of the series obtained by real-time bridge monitoring and detect the existence of signals' nonstationarity, as well as establish an informative background before attempting to use cointegration analysis for anomaly detection. Thirdly, the response surface methodology will be employed to study the effect that the measured EOVs impose on the cable tension series and extract linear relations between them, in terms of regression. A brief description on the response surface method is provided below. Finally, the observations will be discussed and an initial hypothesis about their use inside cointegration will be made.

4.3 Preliminary Analysis of Tamar Bridge Data

The proposed preliminary analysis includes three main steps. In the first step, the tension series are categorised according to cable type and their response into three groups, in order to be studied more efficiently. Secondly, a representative cable tension series will be plotted with respect to the available EOVs (traffic, air temperature, wind speed and relative humidity) and multivariate linear regression (response surface modelling) will be employed to establish informal relations between

tensions and EOVs. Finally, a comparison between series of different tension groups will be performed to explain any difference on the variability of the monitored signals.

To begin with, the aforementioned tension signals represent half-hourly measurements recorded over a motoring period of two years, as depicted in Figures 4.6-4.8. From these, it can be observed that the majority of tension signals exhibit the same pattern over time. However, in order to more efficiently study them, the series will be divided into three groups.

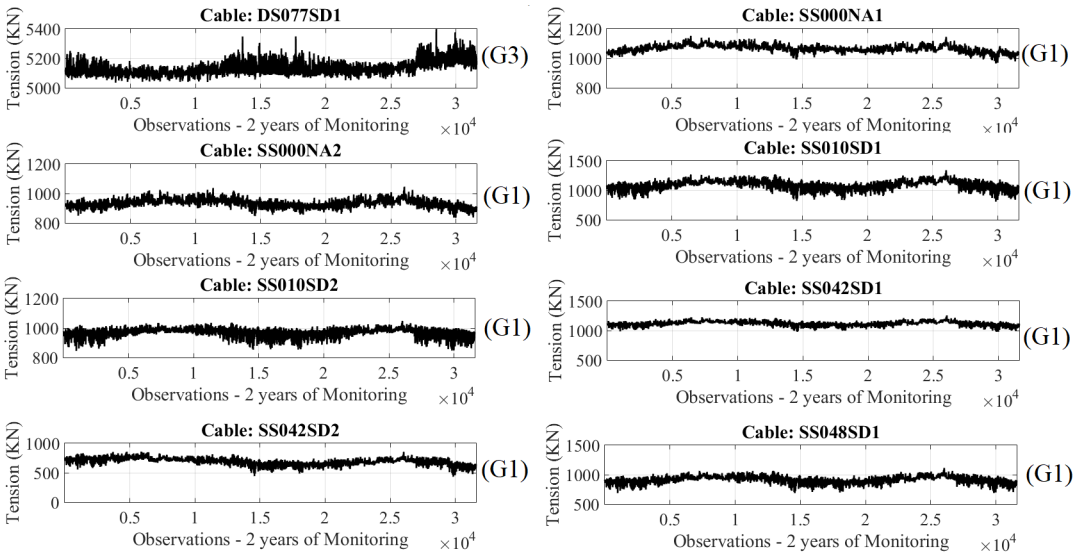


Figure 4.6: Cable Tensions recorded over the examined monitoring period.

The first group includes those series following an increasing trend between 5000-10000 observations, which then falls and again increases around 20000-25000 observations. These are denoted as $G1$ (group 1) in Figures 4.6-4.8. The second group includes those series that exhibit the inverse pattern with respect to the series of group 1 and are denoted as $G2$ (group 2). The third group ($G3$) includes two cable tensions measured using displacement sensors, $DS077SD1$ and $DS077ND1$, and are located under the deck on both sides of the bridge. The functional purpose of these two cables is to restore horizontal deformations and creep during construction stages. It should be mentioned that all the cable tensions shown in Figures 4.6-4.8, follow a generic notation; SS stands for tension sensor, N stands for north, S for south, A for anchorage level and D for deck level. The reference numbers on the notation describe the distance (in metres) from a reference point along the length of the bridge.

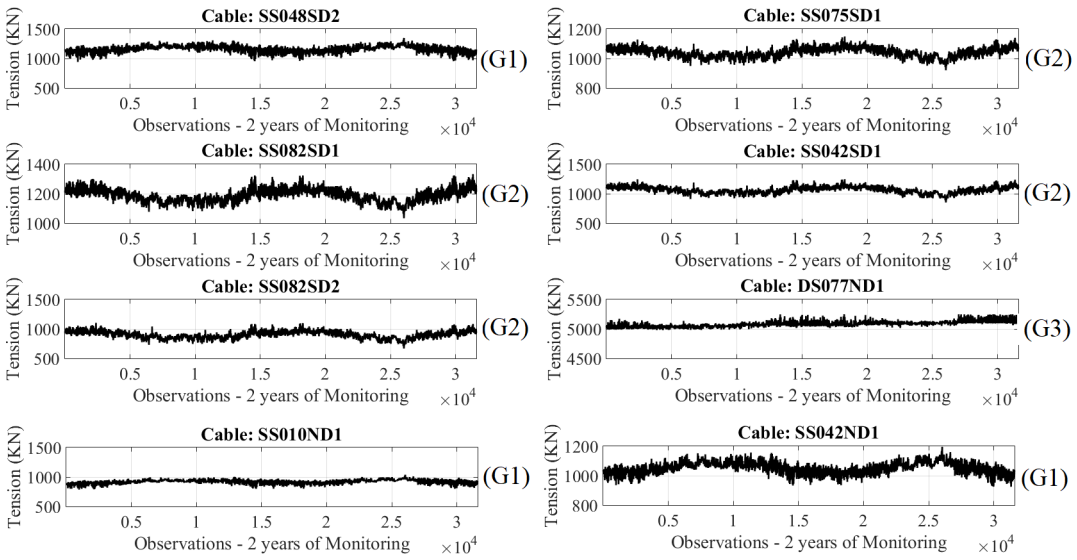


Figure 4.7: Cable Tensions recorded over the examined monitoring period.

4.3.1 Response Surface Models

Response surface models will be used here to study the correlation between series. The methodology follows the work of [89] and describes the use of a multi-variable regression model for the estimation of the regression coefficients. In particular, the method includes two steps. Firstly, the available variables are normalised, which means that the their mean (μ) value is subtracted from their initial value and then the latter is divided by its standard deviation (σ). This aims for all the variables to be of the same scale in order to establish a regression model. Finally, the second step includes the linear regression between the variables based on ordinary least squares (OLS). Assuming a linear regression, such as: $y_i = ax_{1i} + bx_{2i} + cx_{3i}$, the coefficients (a, b, c) of the regression can provide statistical evidence of which variables are the more informative in order to explain the variable in question (y_i). The greater the coefficient, in terms of absolute value, the stronger the influence of this variable on the y_i.

A normalised mean squared error (nMSE) is introduced here as a measure of model fitness and is defined as:

$$nMSE = 100\frac{\sum(model \quad errors)^2)}{n\left(\sigma\left[predictions\right]\right)^2} \tag{4.1}$$

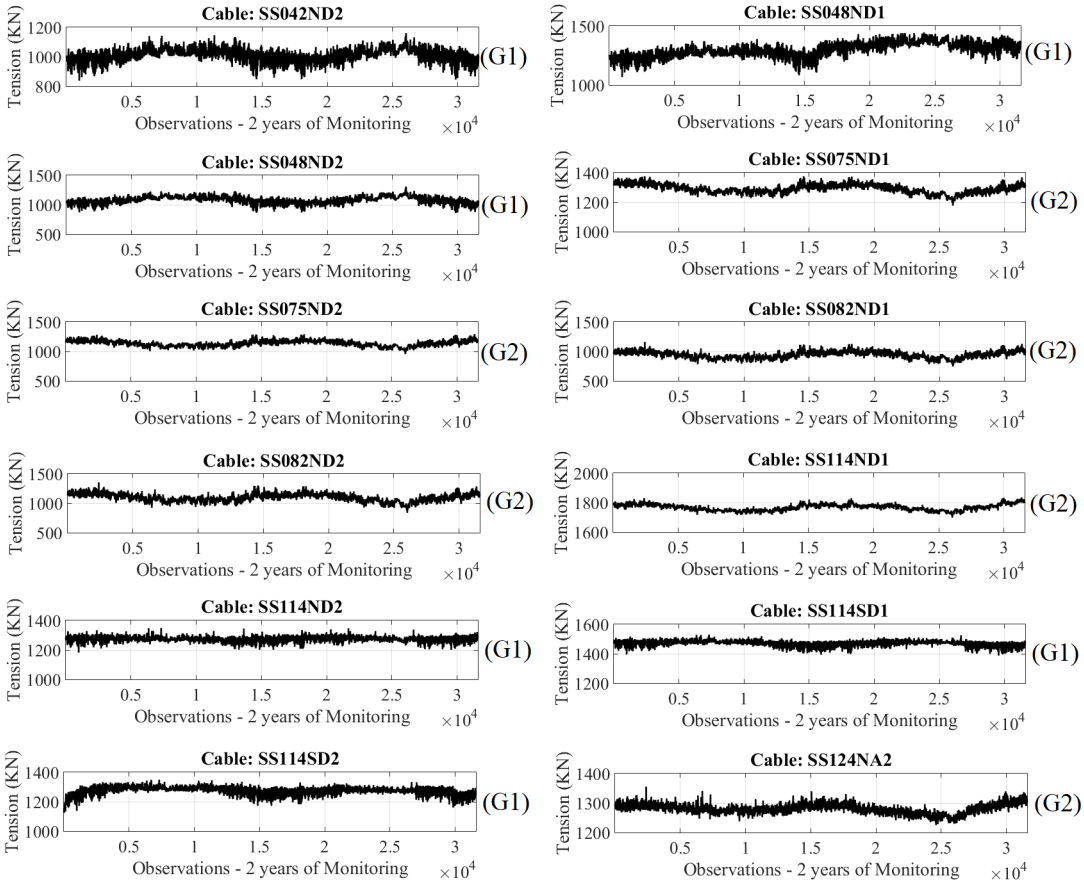


Figure 4.8: Cable Tensions recorded over the examined monitoring period.

where n is the number of data observations of the series and σ denotes the corresponding standard deviation. This nMSE has the property that, if the mean of the data is used as the model, the nMSE will be 100%. With this normalisation, values of nMSE below 100% are indicative of captured correlation. Inside the context of statistics, it is common to use the R^2, correlation coefficient, as a means of studying model fitness. In particular, the nMSE is related to the R^2 coefficient, as shown in Equation (4.2), where R^2 increases as nMSE decreases.

$$R^2 = 1 - \frac{nMSE}{100} \tag{4.2}$$

4.3.2 Simple Plotting

As the steps of the preliminary analysis are described, now the $S3$ back-stay cable tension series is examined, which corresponds to cable SS010SD1 in Figure 4.6. The first step of the analysis includes simple plotting aiming to extract functional relations between the $S3$ back-stay cable tension and the explanatory variables of air temperature, traffic mass variation, wind speed and percentage (%) of atmospheric humidity for the monitoring period starting at 12^{th} July 2007 and finishing at the 7^{th} July 2009.

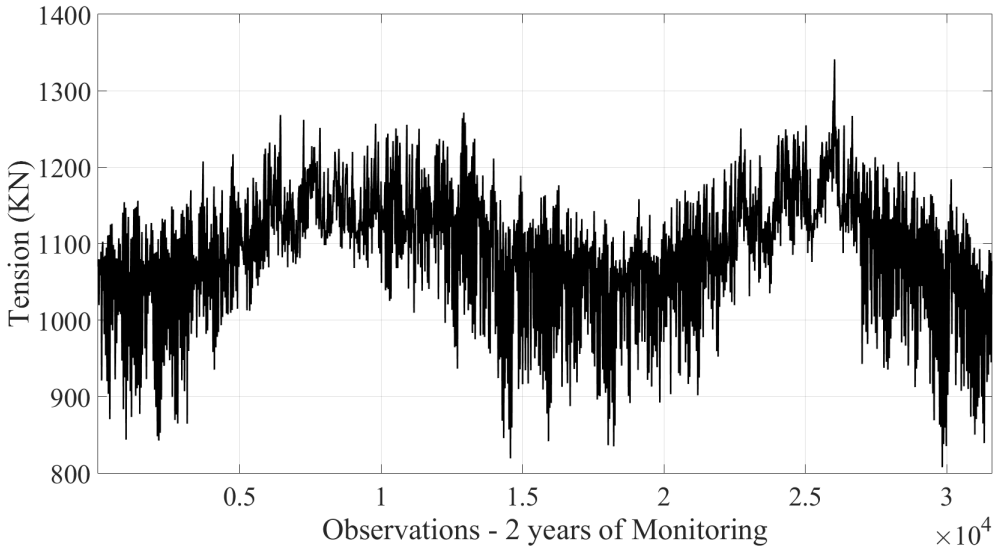


Figure 4.9: Cable Tension S3 recording over the examined monitoring period.

Figure 4.9 shows the S3 back-stay cable tension for the examined monitoring period. From the latter it can be observed that $S3$ exhibits a fluctuation over an mean tension value of 1098.1 KN and can be considered as nonstationary. As it is nonstationary, the next step is to identify the probable sources behind this nonstationarity. In order to do this, the cable tensions is plotted versus the available EOVs. More specifically, Figure 4.10 depicts the $S3$ back-stay cable tension with respect to air temperature. From this, a negative correlation between S3 tension and air temperature (T) can be assumed, as by linear regression it can be found that $S3 = -13.103 \cdot T + 1251.8$. The latter relation has been commonly observed in bridge SHM literature, considering that stay-cables tend to tighten when temperature falls [90], as a consequence of effective stiffness and cable length variation during cold temperatures.

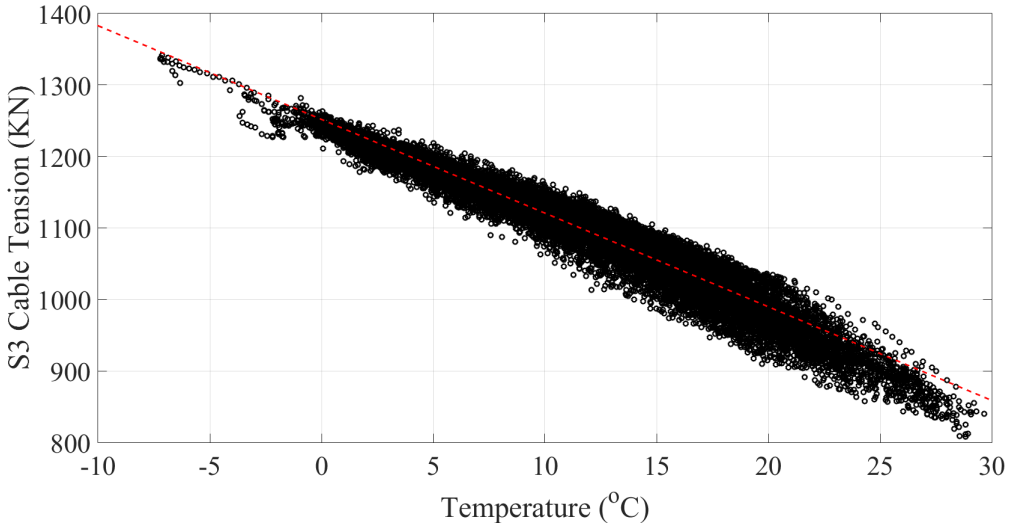


Figure 4.10: Cable Tension $S3$ with respect to air temperature over the examined monitoring period. The red dashed line provides the best linear fit based on OLS.

Furthermore, the $S3$ tension is plotted with respect to wind speeds recorded over the examined monitoring period. From Figure 4.11, it can be observed that it is difficult to assume a deterministic relationship between $S3$ tension and wind speeds (W). The latter is also observed when the $S3$ cable tension is plotted with respect to wind speeds over 25 mph, as shown in Figure 4.12. This behaviour has been observed in the majority of stay-cables of the Tamar Bridge. According to [91] the change in wind speed is expected to introduce some amount of vibration to suspension bridges' cables, however from different showcases it was observed that the wind speed did not have a particular effect on the variation of cable tensions.

Moreover, the relationship between traffic mass variation and $S3$ back-stay cable tension is studied. Figure 4.13 shows the plot of $S3$ cable tension with respect to traffic mass of Tamar Bridge during two and half years of monitoring. From this it can be said that there is no specific relation between traffic mass and $S3$ cable tension variation. In addition, Figure 4.14 shows the relation between traffic masses over 90 tonnes with respect to $S3$ stay cable tension. From the latter, once again it is difficult to assume a deterministic relationship. An probable explanation is that the stay-cable system is a secondary load carrying system and is mostly affected by traffic mass variation when a significant amount of vehicles is present on the bridge (traffic congestion). This can lead to increase of cable tensions as the cables are becoming tight when load increases [88]. However, the latter cannot be assumed

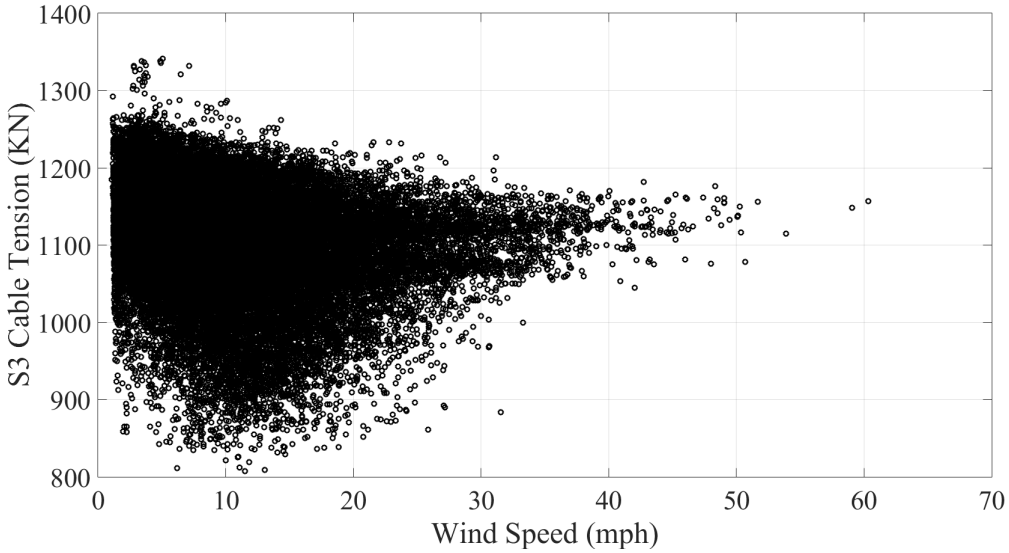


Figure 4.11: Cable Tension $S3$ with respect to wind speeds over the examined monitoring period.

based on the plot, something that holds for the majority of stay cables.

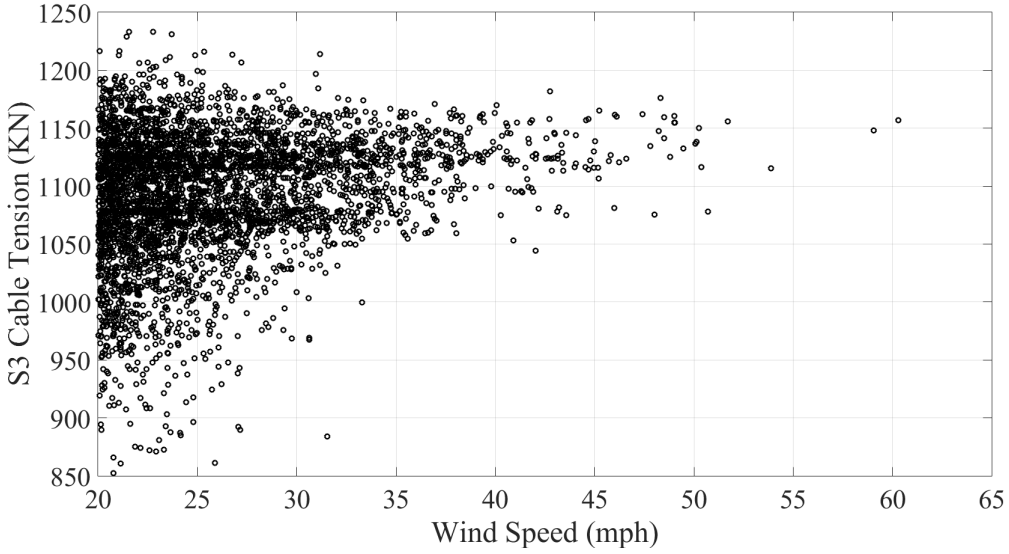


Figure 4.12: Cable Tension $S3$ with respect to wind speeds over 20 mph for the examined monitoring period.

Figure 4.15 depicts the $S3$ cable tension versus the percentage (%) of atmospheric relative humidity. From this plot it is difficult to assume that there is correlation between relative humidity and cable tension. As described in [92, 93], secondary EOVs impacts, such as those introduced by humidity and rainwater can be manifested as a combined action associated with wind induced vibration, however it is

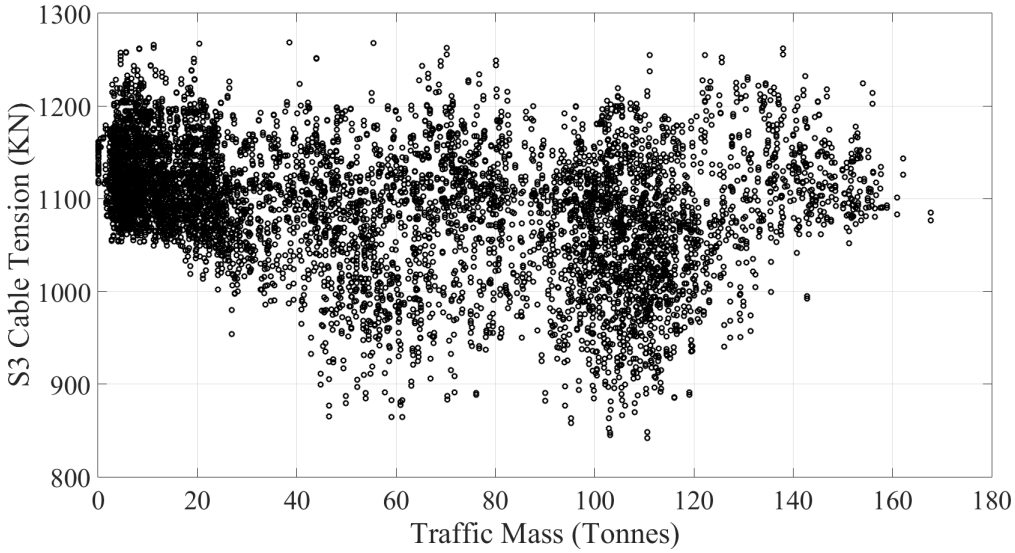


Figure 4.13: Cable Tension $S3$ with respect to traffic mass over the examined monitoring period.

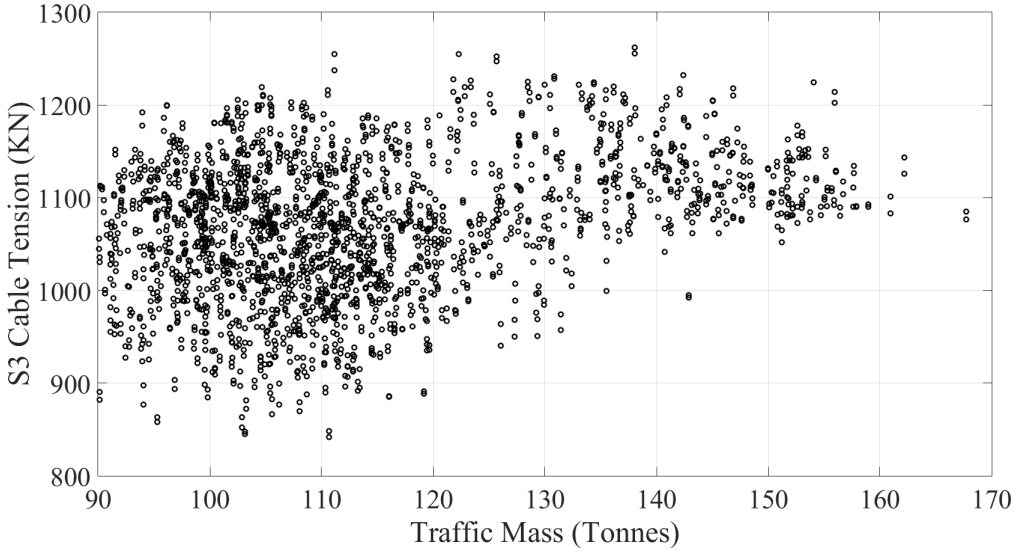


Figure 4.14: Cable Tension $S3$ traffic mass greater than 90 tonnes, over the examined monitoring period.

difficult here to draw a similar conclusion.

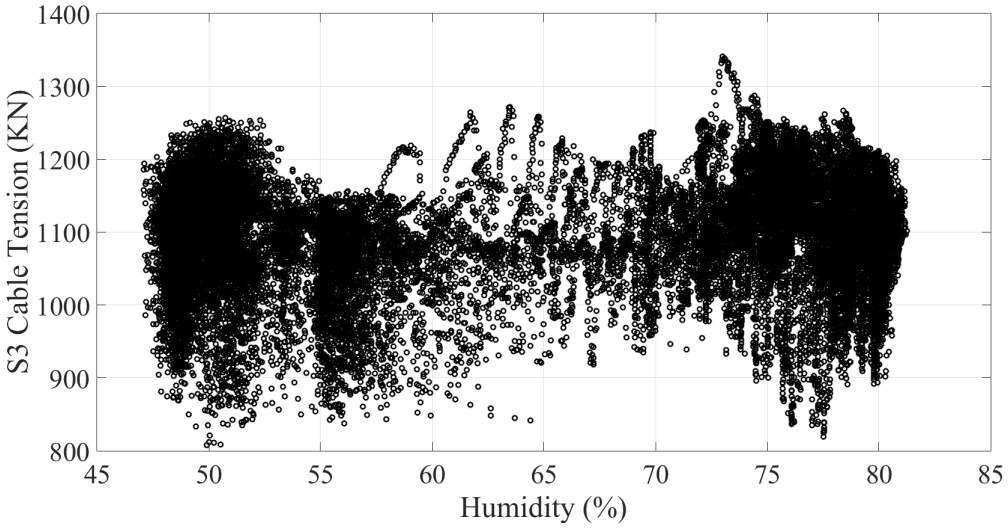


Figure 4.15: Cable Tension $S3$ with respect to the percentage of relative atmospheric humidity over the monitoring period of two years. The red dashed line provides the best linear fit based on OLS.

4.3.3 Results of Response Surface Model Analysis

After studying the plots and relationships between the explanatory variables and $S3$ cable tension, the next step is to investigate the data sets available by using the response surface model regression. In particular, Figure 4.16 illustrates the surface model forecast with respect to the normalised signal of $S3$ cable tension for the examined period of monitoring. Table 4.1 summarises the coefficients of the explanatory variables included in the response surface model regression. From Figure 4.16 it can be observed that the surface model provides good fitting with respect to a training period taken from the normalised $S3$ signal, providing a normalised mean squared error (nMSE) of 11.51%. Furthermore, from Table 4.1 it can be seen that the main driver of cable tension's variability is air temperature, with humidity to be second, wind speed third, while traffic has a slight effect on $S3$ tension variability. In other words, temperature is the main EOV that can be associated with $S3$ tension's variability.

At this point, the response surface model regression is employed again in order to analyse the cable signal coming from group 2. This is the signal of cable SS082SD1, which is depicted in Figure 4.17. As mentioned previously and comparing it with Figure 4.9, it can be observed that the two signals exhibit inverse proportional trends. The surface model results are shown in Figure 4.18, while the corresponding

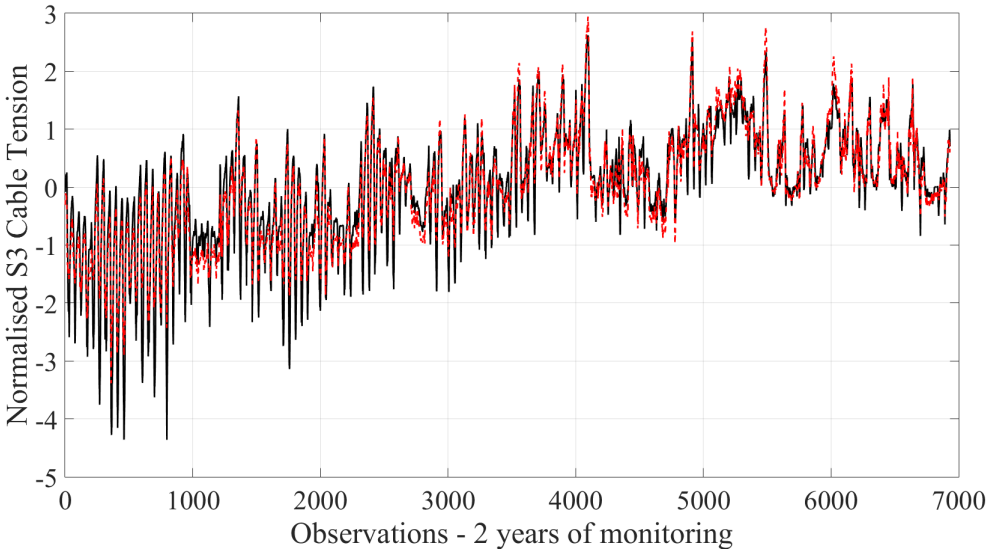


Figure 4.16: Normalised cable Tension $S3$ (black) and surface model forecast for $S3$ (red).

Table 4.1: Surface Model Regression Coefficients for normalised $S3$ cable tension.

Explanatory Variable	Coefficient
Traffic	-0.0032
Temperature	-0.9815
Wind Speed	-0.0589
Humidity	-0.0972

coefficients are summarised in Table 4.2. The response surface model provides good fitting with an nMSE of 11.06%. In addition, as expected looking at its coefficients, temperature is again the main driver of tension's variability, however providing a coefficient with opposite sign with respect to the tension series of group 1. In other words, cable tensions of group 2 follow similar behaviour to the series of group 1.

Table 4.2: Surface Model Regression Coefficients for normalised SS082SD1 cable tension.

Explanatory Variable	Coefficient
Traffic	0.0175
Temperature	0.8987
Wind Speed	-0.0562
Humidity	-0.0812

Although, the latter observations occurred in the majority of cable tensions, it should be mentioned that there was one case where the observations are different. More

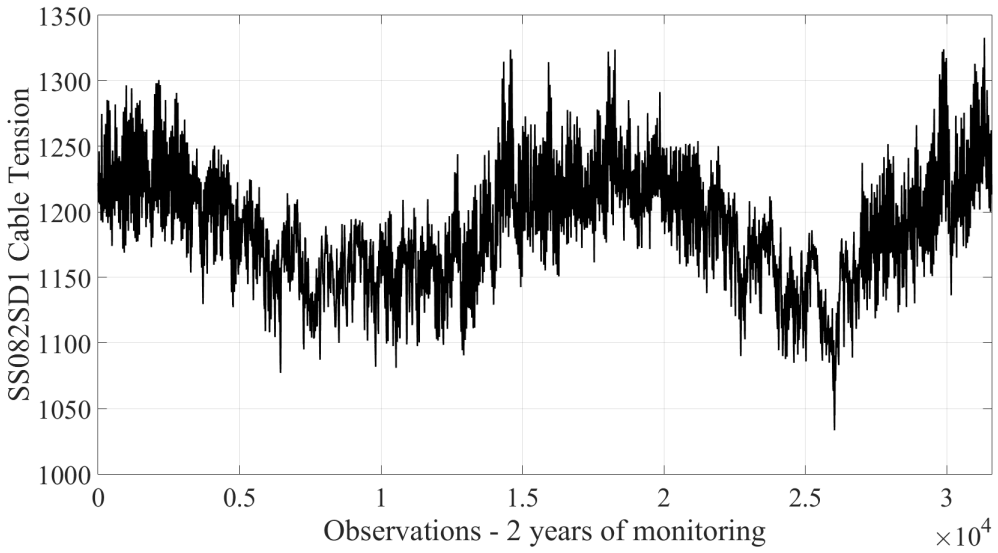


Figure 4.17: Signal of cable SS082SD1 tension over two years of monitoring.

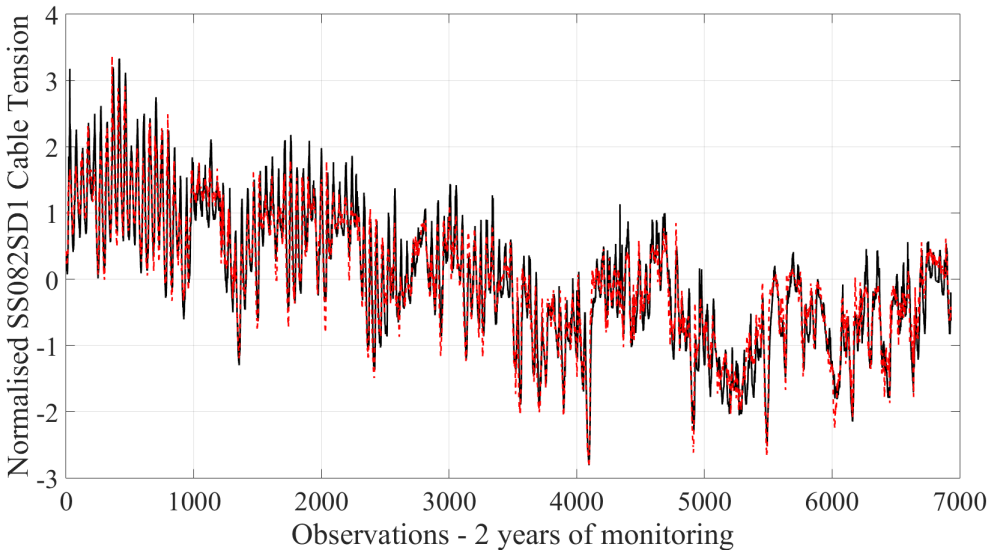


Figure 4.18: Normalised signal of cable SS082SD1 tension (black), group 2, and the corresponding surface model forecast (red).

specifically, in the case of cables, which design purpose is to restore deformations, such as creep, which occurs during the construction stages of the deck as well as during its in-service operation. This case is shown in Appendix A.1. as it is slightly outside the scope of the current analysis.

4.4 Z24 Data Description

The Z24 overpass was constructed in Switzerland in the beginning of the 1960s in order to link Koppigen and Utzenstorf by passing over the main national Highway $A1$ Lausanne-Zurich. The overpass was constructed by the Road department of Canton of Bern and its demolition was carried out in 1998, in order to accommodate newly designed train line. Before its demolition taking place, continuous monitoring and destructive testing activities were performed on it [4].

4.4.1 Z24 Structural Description

The deck of Z24 has 8.6m width including two vehicular lanes and the total span of the bridge is 58m (14+30+14m). The cross-section of deck is an post-tensioned reinforced concrete open rectangular box with average depth of 1m. The top side of the deck carries a 6cm surfacing (asphalt), inclination of 2%, while its plan-view is skewed (13^o). The bridge has full-height integral abutments, which each one consists of three rectangular columns (0.4×0.4m; 1.9m distance each other; 6.15m height) and were located inside ground. The foundations have a conical cross-section (height 1.65 and 1.80m respectively), while the foundation beam had a rectangular cross-section (0.80×1.00m; 7m length). Finally, Z24 has concrete leaf (wall) piers with a rectangular cross-section (0.4×3.2m) and an average height of 6m. The deck is continuous over the piers forming a longitudinal diaphragm. The aforementioned can be observed in Figure 4.19.

4.4.2 Available SHM Data

The sensors installed on the bridge included multiple temperature sensors placed in different structural locations, such as on bridge girder sections, on the pavement, approaches, soil, while also the air temperature is measured. Furthermore, in total 16 accelerometers were placed on the deck of Z24 bridge in order to measure accelerations. These are measured on hourly basis for a total monitoring period of 10 months. It should be mentioned that the effect of traffic on the natural frequencies of Z24 is assumed to be insignificant, because either one or no cars crossed the bridge during a day.

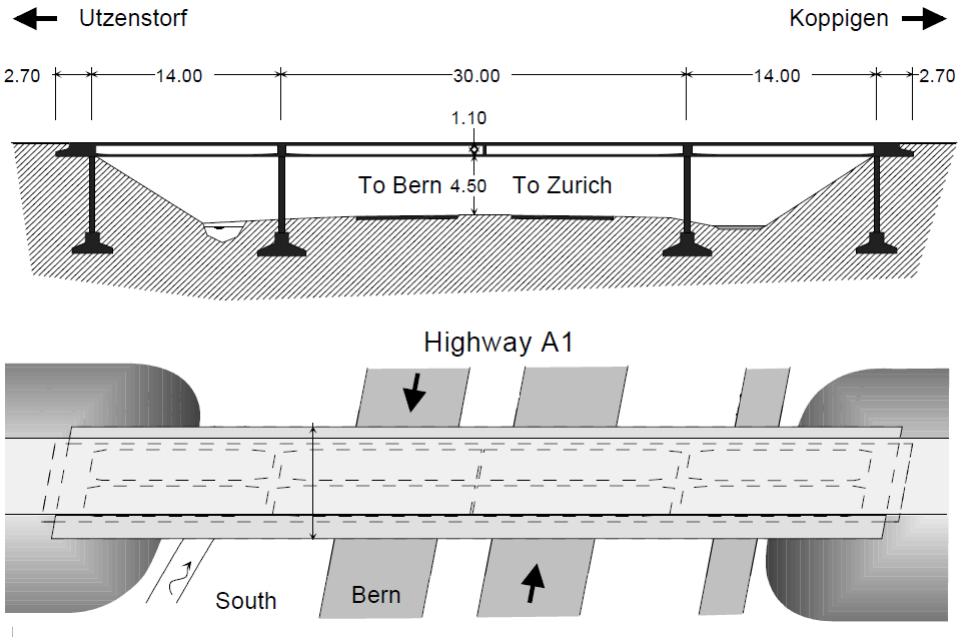


Figure 4.19: Elevation Plan and Overview of Z24 Overbridge [4].

According to the data acquisition strategy described in [4], 10 scans of environmental data were collected per hour and these are averaged to obtain the average hourly temperature value. In addition, 8 averages of 8192 samples at a sampling frequency of 100 Hz were collected and stored for the whole group of the 16 accelerometers. The main features for SHM were modal quantities extracted automatically by stochastic subspace identification (SSI) at regular intervals [46], hourly values. The corresponding modeshapes are not available for further analysis.

As mentioned previously, a series of non-destructive scenarios were performed on the bridge after 8 months of continuous monitoring. More information about these scenarios can be found in [4]. The natural frequency series are illustrated in Figure 4.20. In the left plot the normalised natural frequency data are shown, whilst in the right the original ones. From these two plots, it can be observed that the natural frequencies share common trends, while after 3500 observations a significant change on the second natural frequency can be observed (as designated by the vertical dashed red line).

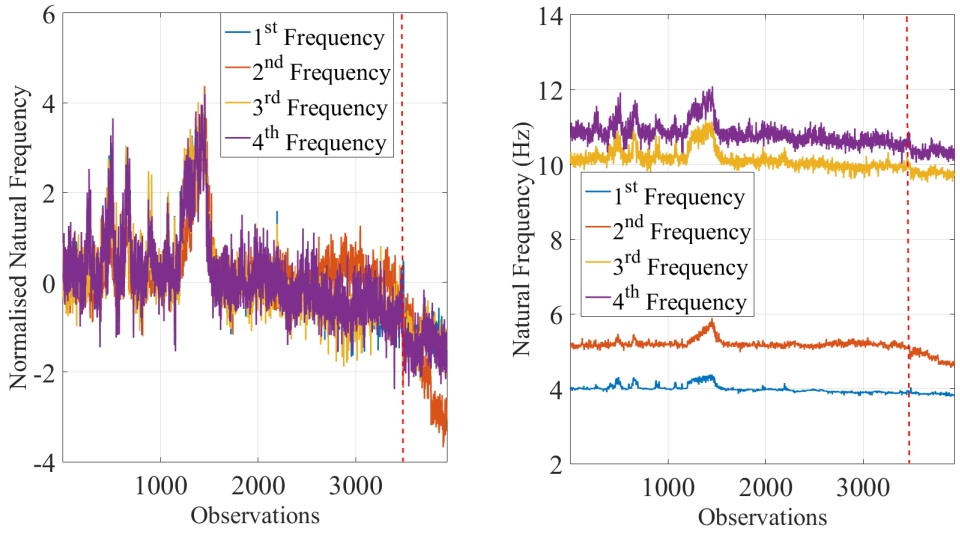


Figure 4.20: Normalised and original natural frequency time series of Z24; red line distinguish undamaged from damaged region.

4.4.3 Response Surface Model Results

The effect of air temperature on natural frequencies of $Z24$ can be explored using a response surface model. The air temperature series is illustrated in Figure 4.21. Here, all the data of natural frequencies and temperature are used inside the surface model and there is no provision for training and testing region, because the surface model aims only to explore the effect of temperature on natural frequency and not to be used for model predictions. The results of surface models with respect to normalised natural frequencies are illustrated in Figure 4.22, whilst Table 4.3 summarises the regression coefficients. It is important to mention that only temperature measurements were available for this analysis.

Table 4.3: Surface Model Regression Coefficients for Normalised Natural Frequencies with respect to Normalised Temperature.

Nat. Freq.	Temp. Coeff.
1^{st}	-0.7426
2^{nd}	-0.4169
3^{rd}	-0.7269
4^{th}	-0.7391

Looking at Figure 4.22 and Table 4.3, the main observation is that for the second natural frequency, temperature has the lowest regression coefficient (-0.4169). This is happening because this frequency series is influenced significantly by damage, as

one can observe after 3000 data points (Figure 4.20).

Another interesting result is that although air temperature can explain some of the variability observed on natural frequencies, the error of regression is significantly high and especially in specific parts of the plots, such as the regions of large artefacts, i.e. between 1200 and 1800 data observations. These regions have been associated with the effect of low temperatures on natural frequencies, assuming a nonlinear relationship between them [42, 46]. In particular, it was observed that a layer of ice was formed in the asphalt and structural boundary conditions changed, leading to an extreme increase in natural frequency values.

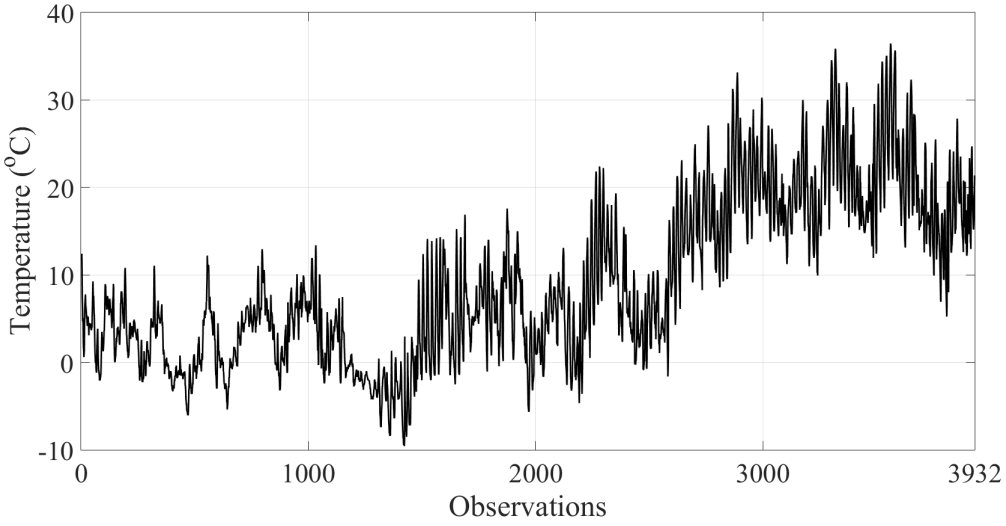


Figure 4.21: Z24 air temperature variation.

From the previous analysis on Z24 overbridge natural frequencies and especially from Figure 4.20, it can be said that the series share common trends. These trends can be associated with the long-term effect that temperature impose on the natural frequencies, as shown in Figure 4.22. However, it is clear that temperature is not the only factor affecting the variability of natural frequencies, something that cannot be investigated here due to the lack of available data from additional environmental and operational variables.

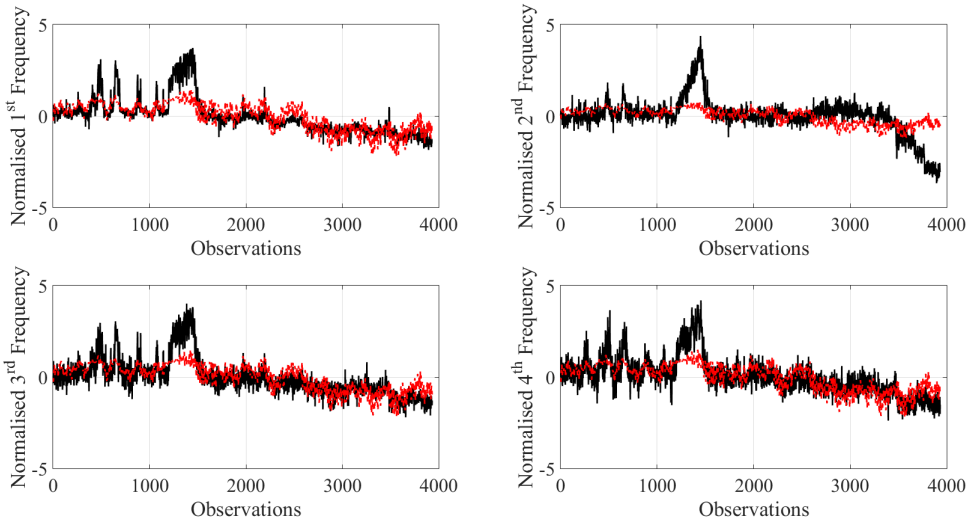


Figure 4.22: 1^{st} order surface model results (red colour) with respect to the normalised first four natural frequencies of Z24 (black colour).

4.5 Conclusions

In this Chapter, a description of two real-time monitoring systems of bridges, the available sensing systems and the recorded SHM data are presented. In particular, in the case of the Tamar Bridge emphasis is placed on the cable tension data and a preliminary analysis performed using simple plotting and the response surface model. From these, it was shown that the main EOV driving the variability of cable tensions is air temperature. In addition, two different types of tension series are examined pointing that they share inversely proportional trends, which can associated with temperature variability.

Furthermore, the data sets available from the monitoring campaign on the Z24 bridge were examined and more specifically the relationships between natural frequencies and air temperature. It was shown that the series share common trends, which in some extend can be associated with temperature. However, some of the variability observed in the natural frequency series, such as the large artefacts detected during winter period, are difficult to be explained in terms of linear regression and only based on the air temperature measurements.

The previous observations made and data sets examined, as well as the data sets of the laboratory truss bridge experiment presented in Chapter 3 will be used inside Chapter 5 in order to demonstrate the use of *Johansen's approach to cointegration*.

Chapter 5

Johansen's Approach to Cointegration for Data Normalisation

5.1 Introduction

As discussed in the previous Chapters, the use of cointegration for SHM is a data-based projection methodology, which lies inside the unsupervised statistical pattern recognition (SPR) techniques for SHM [7]. In the past it was employed for SHM and damage detection purposes in bridge monitoring applications (i.e. Tamar Bridge [2, 76]), laboratory experimental applications on aluminium plates and composites [77, 78], monitoring of dam structures [79] and offshore pipelines [80].

Cointegration is a multivariate statistical technique [74, 75], which requires for all series included to be of the same order of integration (I(1); nonstationary). The series introduced into cointegration should share common trends and have a long-term relationship between them, although in short-term such a relationship might not exist. This long-term relationship can be captured via a linear combination between series, which provides a lower order of integration ($I(0)$) signal, which is referred to as *cointegration residual*. In particular, using Johansen's approach to cointegration, the long-term common trends are projected into a different dimensional space and the corresponding residual is capable of novelty detection [2]. In the context of

SHM, the common trends shared by the SHM series are associated with the impact of EOVs (i.e. temperature, wind, traffic mass, humidity and others).

As mentioned above, the time-series introduced into Johansen's framework should be nonstationary (I(1)). Statistical tests originated from the domain of econometrics are commonly used for nonstationarity assessment. The most commonly employed are the *Augmented Dickey-Fuller* (ADF) [94, 95], the *Philip-Perron* (PP) [96] and the *Kwitkakowski etal.* [97] (KPSS) tests. These are based on the theoretical concept of *unit-root*, which means that the series involved are tested, under a null hypothesis, for the presence or not of a characteristic root. The unit-root test procedure is it quite straight-forward, as a simple autoregressive model (AR) is employed including the coefficients of the first lag of the original series and those corresponding to the lags of the first difference of the series. Then, the corresponding ADF statistic of the defined AR model is evaluated and compared with the corresponding critical ones [98]; depending on the confidence interval (significance level) selected. The ADF test is the first step of Johansen's approach to cointegration.

Furthermore, the next step of Johansen's approach is the use of a *vector autoregressive model* (VAR). The series are introduced in the VAR and the relationships between them can be obtained. In this model an *error correction term* is introduced able to *adjust/correct* the short-term errors of the VAR model in longer term achieving a long-term equilibrium. This equilibrium is a stationary process, which describes the linear combination of the series and is referred to as cointegration residual.

The current work follows the Johansen's approach to cointegration, which introduced in [76] for SHM applications. In brief, this idea for novelty detection is based on two steps. Firstly, a cointegration vector is established based on training data. This includes coefficients for each series, which multiplied with each observation of the series, can produce a stationary cointegration residual. Next, the remaining and future SHM data are projected on the previously obtained cointegration vector and the residual is monitored to detect any form of novelty. The most common practice to monitor the cointegration residual is using statistical process controls (SPC) [7], as the residual is stationary if cointegration holds (further information in Chapter 6).

The aforementioned procedure is presented and discussed, step by step, inside this Chapter. In particular, in the first section, the Johansen's approach to cointegration

is presented and discussed inside the SHM context. This includes discussion about the ADF test, the VAR model and the error correction term. In the next section, a theoretical discussion is made on how common trends are manipulated inside cointegration and what happens to the residual in the presence of damage. Finally, the SHM data of three different show cases (Tamar Bridge, $Z24$ and laboratory truss bridge) are used to demonstrate the application of Johansen's approach. The aim of the latter is to explore the interrelationships between the series used inside cointegration, as well as evaluate the effect of environmental and operational variations (EOVs) and damage on them. Finally, the last section includes the conclusions of the three aforementioned show cases.

5.2 Description of Johansen's Approach to Cointegration

In general sense, Johansen's approach to cointegration includes three main steps, which are (i) the ADF unit-root test, (ii) the VAR model and (iii) the reduced rank regression. These will be discussed in order to understand how, combined together, can be used for SHM purposes.

5.2.1 The ADF Test

The simplest form of ADF test is given in Equation (5.1), where ΔY_t describes the first difference of time-series Y_t, ρ is the root of the process, ΔY_{t-1} is the first lag-term of the first differenced time-series, e_t is a white noise term and t is the lag of the first difference ΔY_t. The main concept in which the ADF test performs, is the unit-root hypothesis. In particular, the root ρ can be either equal to one, which provides statistical evidence for nonstationarity, or is $0 < \rho < 1$ and the data can be considered as stationary.

$$\Delta Y_t = (1 - \rho)\Delta Y_{t-1} + e_t \tag{5.1}$$

Here, it should be mentioned that inside the context of econometrics, there is criticism about the the validity of the unit-root hypothesis concept, based on the fact

that the unit-root will never be found with complete precision and the statistical evidence provided for decision making, stationary or not, cannot always provide full confidence [96, 97, 99].

In Equation (5.2), a simple random walk (AR) is described, where Δ is the first difference of the time-series Y_t, ΔY_{t-i} is the value of the previous differenced lags for $i = 1, ..., p$, where p is the lag-length selected, multiplied by the appropriate coefficient (a_i). Finally, the term δY_{t-i} includes the potential unit-root, since $\delta = 1-\rho$ (potentially nonstationary term). In the case that a unit-root is present, $\delta = 0$, then the first difference ΔY_t is equal to a stationary process defined by the sum of lagged values of the first difference and the residual error term (e_t), which are also stationary. Two more ADF modifications are proposed in [74] by introducing two additional terms; the drift (c) and the time-trend (T), which can be used to describe the nonstationary series tested.

$$\Delta Y_t = \delta Y_{t-1} + \sum_{i=1}^{p-1} \alpha_i \Delta Y_{t-i} + e_t \tag{5.2}$$

The unit-root testing is based on a null hypothesis H_0, which states that the ADF regression has a unit-root $\rho = 1$ or $\delta = \rho - 1 = 0$. The null hypothesis H_0 is rejected if the obtained t-statistic is lower than the corresponding critical t-statistic [98]. The critical values considered in this work corresponds to 99% confidence interval. The t-statistic can be estimated using Equation (5.3), in which $\hat{\rho}$ describes the approximated unit-root, while σ_ρ describes the standard error of $\hat{\rho}$.

$$t_{ADF} = \frac{\hat{\rho}}{\sigma_\rho} \tag{5.3}$$

Furthermore, it is important to determine the most appropriate lag length of the ADF model. For this, the *Akaike Information criterion* (AIC) is employed [100]. The AIC statistic can be estimated by $AIC = nlog(\sigma^2_{e_t}) + 2p$, where σ^2 is the variance of the error term e_t, p is the lag length examined and n is the number of observations of the series under consideration. The procedure of lag-length specification starts with the estimation of the maximum lag number (p_{max}), according to Schwert [101] formula, $p_{max} = [12(n/100)]^{0.25}$, where n indicates the number of observations. Next, the ADF model is employed and run for multiple lag values, from zero to p_{max}. During each run, the AIC statistic is saved and the lag length corresponding to the

lowest AIC is employed as the most appropriate one.

5.2.2 Vector Autoregressive (VAR) Models

VAR models are a category of autoregressive models, which study the dynamic interrelationships between multiple variables. Typically, they are an extension of the univariate autoregressive model (AR) to multivariate time-series and are commonly employed in macroeconomics and financial time-series [102]. Inside the context of SHM and especially for bridges, there are not many attempts employing a VAR model for damage detection. However, some interesting work is presented in [103], where a VAR model is employed for damage assessment on the data provided by the response of an 8-degree freedom (DOF) simulation test. In addition, Mosavi et al. [43], employed a VAR model for damage detection and identification in the ambient vibration data provided by testing slender steel beams undergoing buckling. A VAR model was also employed by Li et al. [104] and used to track system changes during the vibration of a ship structure under real environments. It should be mentioned that more advanced VAR models have been employed also for fault detection and system identification purposes. A brief description of an one lag VAR model is given in A.2.

As in the case of ADF, the most appropriate lag length for the VAR model should be identified. This can be achieved using the AIC statistic, running multiple least squares fittings for different lag numbers. This procedure is similar to the one described for ADF model. In this case, the AIC statistic is provided in Equation (5.4), where n_t corresponds to the number of degrees of freedom, which is a function of the number of variables (n) and the number of lags (p; Equation (5.6)). Then, there is the likelihood ratio (L), which is given in Equation (5.7), where N is the number of observations and ω is the determinant of the squared VAR model error $(Y - \hat{Y})$ over the number of observation minus the lag number and multiplied by model order (Equation (5.5)).

$$AIC = \frac{2\,(n_t - L)}{N - p} \tag{5.4}$$

$$\omega = det \frac{(Y - \hat{Y})^2}{(N - p)p} \tag{5.5}$$

$$n_t = n(1 + n \times p) \tag{5.6}$$

$$L = \frac{(N-p)(n(1+log(2\pi)+log(\omega))}{-2} \tag{5.7}$$

All in all, from the VAR model specification two bits of information can be provided. Firstly, the most appropriate lag number that will be used to form the error correction model and secondly, a description of the linear interrelationships between the series involved in the cointegration analysis.

5.2.3 The Vector Error Correction Model (VEC)

The third step of Johansen's approach to cointegration is the formulation of the VEC model. The VEC model is a VAR model augmented by an *error correction term*, which adjusts the short-term errors of the VAR model. A VEC model is given in Equation 5.8, where ΔY_i is the vector of the first difference, [Π] is a coefficient matrix describing the long-run equilibrium between variables, $[B_j]$ is the matrix including the coefficients of the first difference lags, which also referred to as short-run impact matrix, Δy_{i-j} are the lags of the first difference and e_i is a normally distributed noise process ($e_i \sim N(0, [\Sigma])$).

$$\{\Delta Y_i\} = [\Pi]\{y_{i-1}\} + \sum_{j=1}^{p-1}[B_j]\{\Delta y_{i-j}\} + \{e_i\} \tag{5.8}$$

$$\{\Delta Y_i\} = [\alpha][\beta]^T\{y_{i-1}\} + \sum_{j=1}^{p-1}[B_j]\{\Delta y_{i-j}\} + \{e_i\} \tag{5.9}$$

The term $[\Pi]y_{i-1}$ is the only one in Equation (5.9), which potentially describes an $I(1)$ series. For ΔY_i to be I(0), $[\Pi]y_{i-1}$ should be also $I(0)$, which means that it contains the cointegrating relations if any exists. [Π] is referred to as correction matrix. The rank of [Π] can define the existence and the number of possible cointegrated linear equations. In particular, if the rank of [Π] is zero, implies that y_i is $I(1)$, VAR model includes unit-roots, and there is no cointegration existent. This leads to the reduction of VEC model to a VAR(p-1) model in first differences. The same hap-

pens also when the rank (r) of [Π] is full, which means that is equal to the number of variables (n). The only acceptable case is the reduced rank, which informs that y_i is $I(1)$ and there are $n - r$ available linearly independent cointegration vectors. If the latter occurs, then the most stationary vector should be selected to satisfy $[\Pi]y_{i-1} \rightarrow I(0)$.

In the last case, [Π] is rank deficient and can be written as the product of two matrices,$[\Pi] = [\alpha][\beta]^T$, where $[\beta]^T$ rows will span the row space of matrix [Π], while $[\alpha]$ denotes the effect of the cointegration vectors to the evolution of Δy_i. Hence, Equation (5.8) can take the form in Equation 5.9, where $[\beta]^T y_{i-1}$ is $I(0)$. Based on the previous, $[\beta]^T$ describes the desired cointegration vector.

The method followed to solve for $[\beta]$ is based on the product moment matrices, according to Johansen [74], from which the VEC model can be represented with different notation. Therefore, let $\{\Delta y_i\} = \{z_{0i}\}$, $\{y_{i-1}\} = \{z_{1i}\}$, $\{\Delta y_{i-1}\}^T, ..., \{\Delta y_{i-p}\}^T = \{z_{2i}\}$ and $[B_1], ..., [B_{p-1}] = [\Psi]$. Based on the above, the problem needing solution is:

$$\{z_{0i}\} = [\alpha]\,[\beta]^T\,\{z_{1i}\} + [\Psi]\,\{z_{2i}\} + \{e_i\} \tag{5.10}$$

The residuals $\{R_{0i}\}$ and $\{R_{1i}\}$ can be established according to the regressions, as described in Equations (5.11) and (5.12), while the corresponding four product matrices are given in Equation (5.13).

$$\{R_{0i}\} = \{z_{0i}\} - [C_1]\,\{z_{2i}\} \tag{5.11}$$

$$\{R_{1i}\} = \{z_{1i}\} - [C_2]\,\{z_{2i}\} \tag{5.12}$$

$$[S_{mn}] = \frac{1}{N}\sum_{i=1}^{N}\{R_{mn}\}\,\{R_{mn}\}^T, m, n = 0, 1 \tag{5.13}$$

Solving the generalised eigenvalue problem of Equation 5.14, the eigenvector (λ_i) can be obtained, which corresponds to available cointegration vectors. The most stationary cointegration residual corresponds to the cointegration vector exhibiting

the largest eigenvalue [76].

$$(\lambda_i)[S_{11}] - [S_{10}][S_{00}]^{-1}[S_{01}])\,\{\nu_i\} \tag{5.14}$$

Having the most stationary residual means that now the cointegration vector $\{\beta\}^T$ can be used for damage detection purposes. As mentioned in [7] the second axiom of SHM mentions that the novelty detection is based on the definition of two structural condition states, an undamaged and a potentially damage. This is achieved within cointegration by establishing the cointegration vector based on the undamaged data and then by projecting the remaining/future data on the cointegration vector. As described in [2, 76], the cointegration vector providing the most stationary cointegrating residual is commonly employed for damage detection purposes inside SHM. Selecting the most stationary cointegration residual, it is likely that the common trends shared between series are successfully purged. In other words, that the confounding effect of environmental and operational variations (EOVs) are projected out of the series.

5.3 Common Trends and Novelty

At this point, it will be useful to give some idea on how linear cointegration manipulates the series to find a cointegrating vector and eliminate common trends, as well as how the cointegration residual is affected when damage is present. This section aims to provide a theoretical background before moving to the application of cointegration into three case studies (Tamar Bridge, Z24 bridge and the laboratory truss bridge), which will follow in the next section.

To begin with, let's assume that, before applying cointegration, two signals (s_{1i},s_{2i}), which are described in Equations (5.15) and (5.16), share a common (deterministic) trend $g(t)$. The y's are some response measured variables of a linear system (for simplicity) which are stationary and Gaussianly distributed. The cointegration equation is given in Equation 5.17. From this Equation it can be observed that cointegration does not learn the common trends, but learns the correlation structure between the features and then cancels it by linear combination [105]. The linear combination provides a stationary residual (e_{1i}). In other words, this shows that cointegration can project out the common trends, independently of their pattern,

as far as they are common.

$$s_{1i} = \beta g(t) + y_{1i} \tag{5.15}$$

$$s_{2i} = g(t) + y_{2i} \tag{5.16}$$

$$z_{1i} - \beta z_{2i} = y_{1i} - \beta y_{2i} = e_{1i} \tag{5.17}$$

Now, supposing that damage occurs, with a severity d and assuming that the system stays linear, one potential scenario will given new governing Equations (5.18) and (5.19), where dashes describe damage case. Now, using the cointegration relation from the undamaged system (Equation 5.17), the cointegration relation can be obtained in Equation 5.20. Here, $\alpha_1(d) \neq \alpha_2(d)$. Even if this does not hold, a difference can still be detected as the dashed variables would give, potentially, a different stationary residual. The above can be expressed in simpler form in the bullet points below.

$$z'_{1i} = \alpha_1(d)\beta g(t) + y'_{1i} \tag{5.18}$$

$$z'_{2i} = \alpha_2(d) g(t) + y'_{2i} \tag{5.19}$$

$$z'_{1i} - \beta z'_{2i} = (\alpha_1(d) - \alpha_2(d) g(t)) + y'_{1i} - \beta y'_{2i} = e_{2i} \tag{5.20}$$

- Normal/Undamaged State: $\beta = y_{1i} - \beta y_{2i} = e_{1i}$, where $\beta = (1, -\beta)$.
- Potentially Damaged State: $\beta^* = y_{1i} - \beta^* y_{2i} = e_{1i}$, where $\beta \neq \beta^*$.

5.4 SHM Data for Johansen's Approach to Cointegration

Cointegration relations depend on the interrelations between the series and the effect that EOVs and damage impose on them. Although, cointegration can be performed without including measured EOVs, it is important to study how these affect the damage sensitive feature. The latter was performed in Chapters 3 and 4 for the laboratory truss bridge natural frequencies, the cable tensions of Tamar Bridge and the natural frequencies of Z24 bridge. However, the aim of this section is to study the interrelationships between the series that will be used inside cointegration. The main hypothesis that will be investigated here is that when two series have a clear negative or positive correlation between them can be successfully cointegrated, otherwise when the correlation is not very clear the cointegration residual it is not very informative. This will be investigated on the aforementioned three cases on bridge SHM data. For the purpose of this investigation, the features will be plotted against each other and the approximated cointegration relation will be studied. In addition, some damage cases will be examined in order to evaluate their effect to cointegration residual.

5.4.1 Tamar Bridge Cable Tensions

From the analysis of cable tension series performed in Chapter 4 it was shown that the main driver of cable tensions variability is the air temperature. In addition, it was assumed that cable tensions which share common trends, as well as inversely proportional trends can be used together in a cointegration analysis. For these reason, it was mentioned that the series of the first and second cable tension groups can be used together inside *Johansen's approach to cointegration*. However, the latter hypothesis should be investigated here. To do so, the relationships between cable tensions are examined and the cointegration analysis is performed to obtain the most stationary cointegration residual.

Cable Tension Relationships

As mentioned previously, it is important to show what exactly happens in the relationships between series of group 1 and 2. In particular, the correlation between two signals of group 1 and 2, as well as two signals only from group 1. The former

is depicted in Figure 5.1, from which a negative correlation between the series can be assumed.

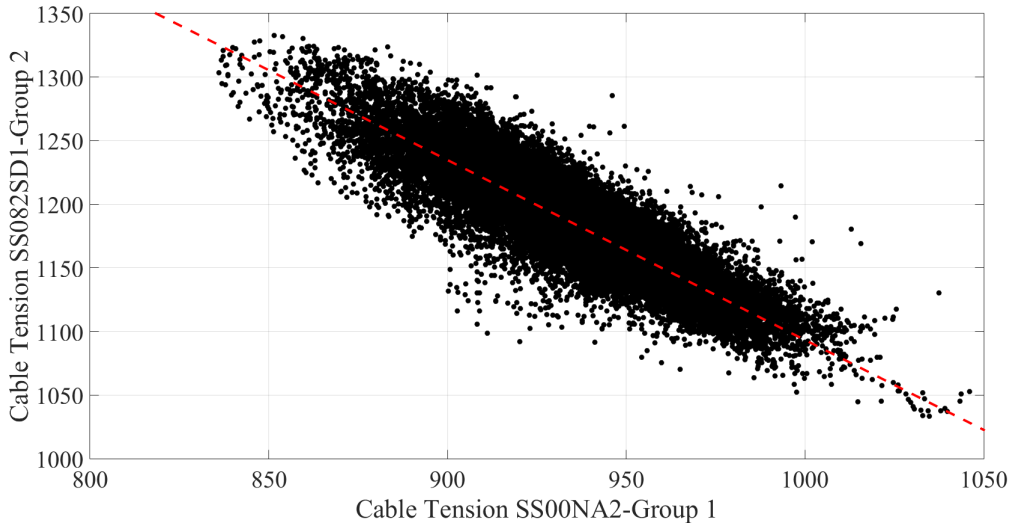


Figure 5.1: Correlation between cable tension signals of group 1 and 2 and best linear fit line (red colour).

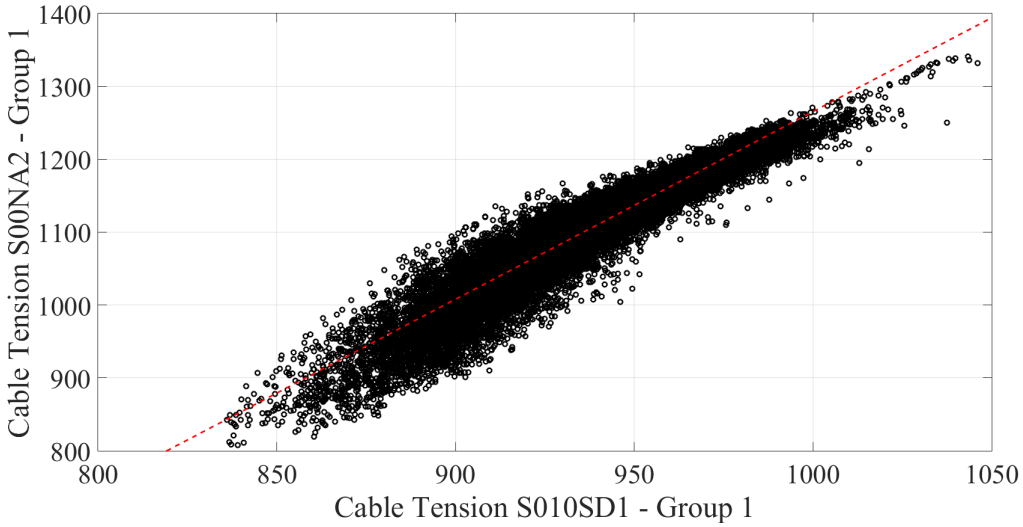


Figure 5.2: Correlation between cable tension signals of group 1 and best linear fit line (red colour).

On the other hand, the plot between the series of the same group, group 1, is depicted in Figure 5.2. From this a positive correlation between the series can be observed. In other words, the previously observed relationships between series are, in general sense, linear because all the tension series of group 1 and 2 are affected by temperature approximately in the same manner. The negative relationship observed between

group 1 and 2 series it is because the signals share inverse proportional trends associated with temperature, something that obtained previously from response surface model regression.

Cointegration Residual

Now, in order to demonstrate the use of *Johansen's approach to cointegration*, 12 tension series of groups 1 and 2 are plotted in Figure 5.3. The cointegration residual approximated, employing a training set of 12000 observations, and plotted in Figure 5.4. Each individual series included in the cointegration has 31570 data observations. The residual is stationary with the majority of observations to be inside the statistically determined thresholds, something that informs for the absence of novelty on the data. In addition, it can be observed the initial trends observed in the series were projected out. It should be mentioned that the stationarity of the cointegration residual depends on the number of lags used during the formulation of vector autoregressive model, as well as the length of the training region. Now, the next step is to see how the cointegration performs in terms of novelty detection.

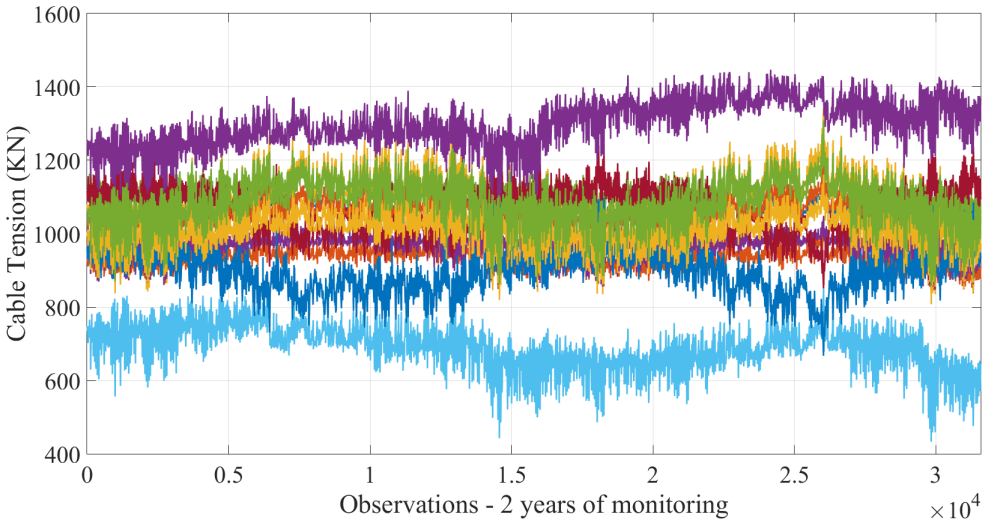


Figure 5.3: Tension signals of 12 cable tensions of group 1 and 2.

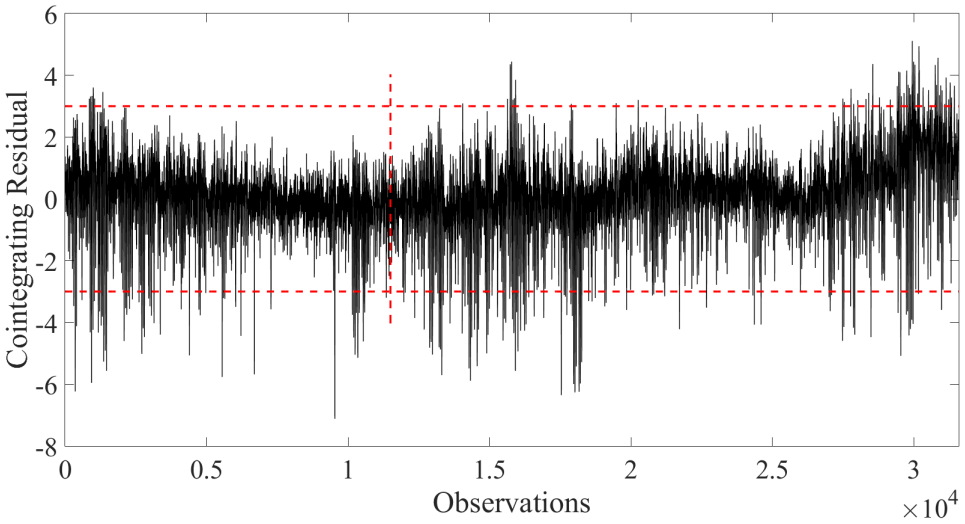


Figure 5.4: Cointegration residual for same number of tension signals from both groups. The horizontal red dashed lines represent the upper and bottom thresholds, while the vertical one describes the change from training to testing region. Threshold set at $\pm 3\sigma$ from the residual mean.

Novelty Detection

A case of an probably unexpected change in tension is investigated here in order to evaluate the novelty detection capabilities of cointegration residual. In particular, from the analysis performed on the cable tensions, the signal of SS048ND1 cable, which belongs to group 2, is found to be slightly distorted at a point around 16000 observations (Figure 5.5). The point of distortion is denoted with a vertical red dashed line. This signal is not involved inside the cointegration residual illustrated in Figure 5.4. However, for novelty detection purposes, the signal is introduced to the cointegration residual by substituting it with another signal of group 2.

The Johansen's approach to cointegration is re-performed providing the residual signal shown in Figure 5.6. From this it can be observed that the cointegration residual signal after 16000 observations becomes distorted and consequently the cointegration residual is able to detect the novelty. A control chart method can be used in such a case in order to detect abnormality more efficiently.

From the previous analysis, it was shown that in the case were multiple signals are available, here 12 cable tensions, the cointegration residual helped to understand that something happened with one of the tension signals. In addition, that the hypothesis that series sharing proportional and inversely proportional trends can be used successfully inside *Johansen's approach to cointegration*. In particular, the

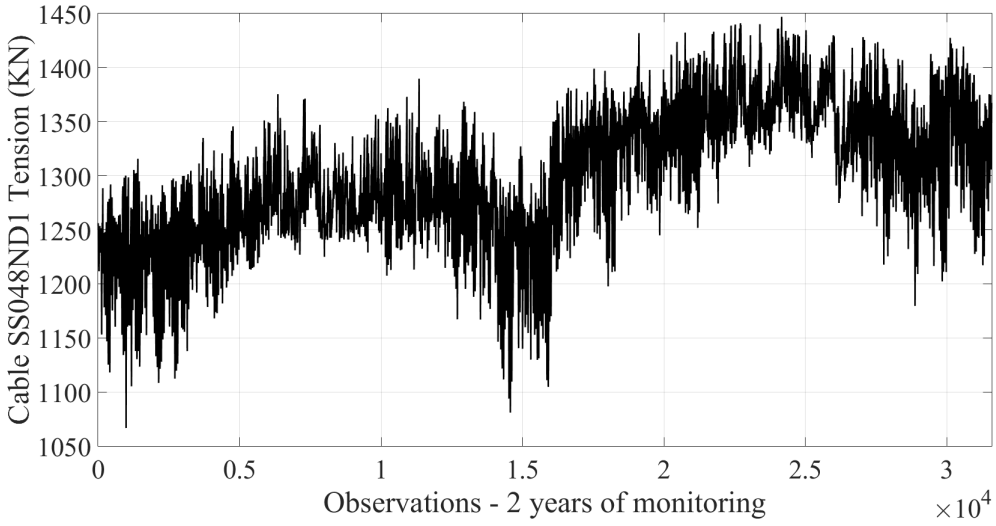


Figure 5.5: Tension signals SS048ND1 over the examined monitoring period.

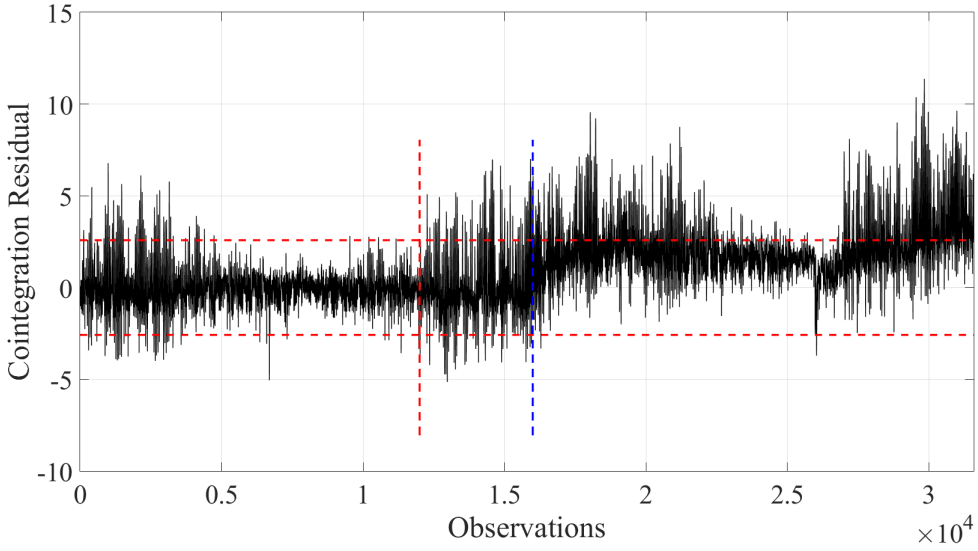


Figure 5.6: Cointegration residual including the SS048ND1 distorted signal. The horizontal red dashed lines represent the upper and bottom thresholds, while the vertical one describes the change from training to testing region. The vertical blue dashed line represents the introduction of damage in the signal. Threshold set at $\pm 3\sigma$ from the residual mean.

series of group 1 have a linear positive correlation between them, while the series between group 1 and 2 have a linear negative correlation between them.

5.4.2 The Case of Z24

According to the analysis performed in Chapter 4 on the natural frequencies of Z24 bridge, temperature observed to be the main driver of the variability observed on the frequency series over time. However, there were parts were the latter effect was not significant and more specifically the large artefacts observed during winter months (Figure 4.20). In addition, the second natural frequency shows lower dependency on air temperature, in terms of response surface model regression, because it is affected by damage significantly.

Knowing all the previous, the series can be used inside *cointegration*. However, before using cointegration, the relationships between natural frequencies should be examined to check for the existence of linear correlation. The natural frequency pairs are illustrated in Figure 5.7. These are categorised, in an ad-hoc manner, into three classes according to temperature: under and equal to 0 ^oC (cyan), between 0 and 15 ^oC (blue) and more than 15 ^oC (green).

From these, it can be said that there is a positive correlation between natural frequencies, in the normal state, as the value of frequency increases when temperature decrease. From literature [42, 46], the impact of temperature is the main driver of frequencies' variability, because at average one car passed over $Z24$ per day [4], while the effect of wind is limited. It should be also mentioned that the only available EOVs measures for the analysis are those of temperature.

Furthermore, looking into the plots, the data in red colour describe those recorded during destructive testing activities. It can be seen that natural frequency values decreased significantly forming a cluster, which shows greater distance from the mean value of data population of the normal state.

As mentioned above, although temperature variability is not capable to explain solely the variability of natural frequencies during the monitoring period, it was clearly observed that the natural frequencies exhibit always a positive correlation between them. This correlation is distorted in the case of the second natural frequency, when damage occurs (red data). The hypothesis tested here is that the signals are appropriate for cointegration, as they present correlation between them.

Cointegration Residual

As shown in Figure 5.8, the first 1000 observations are used for the training of

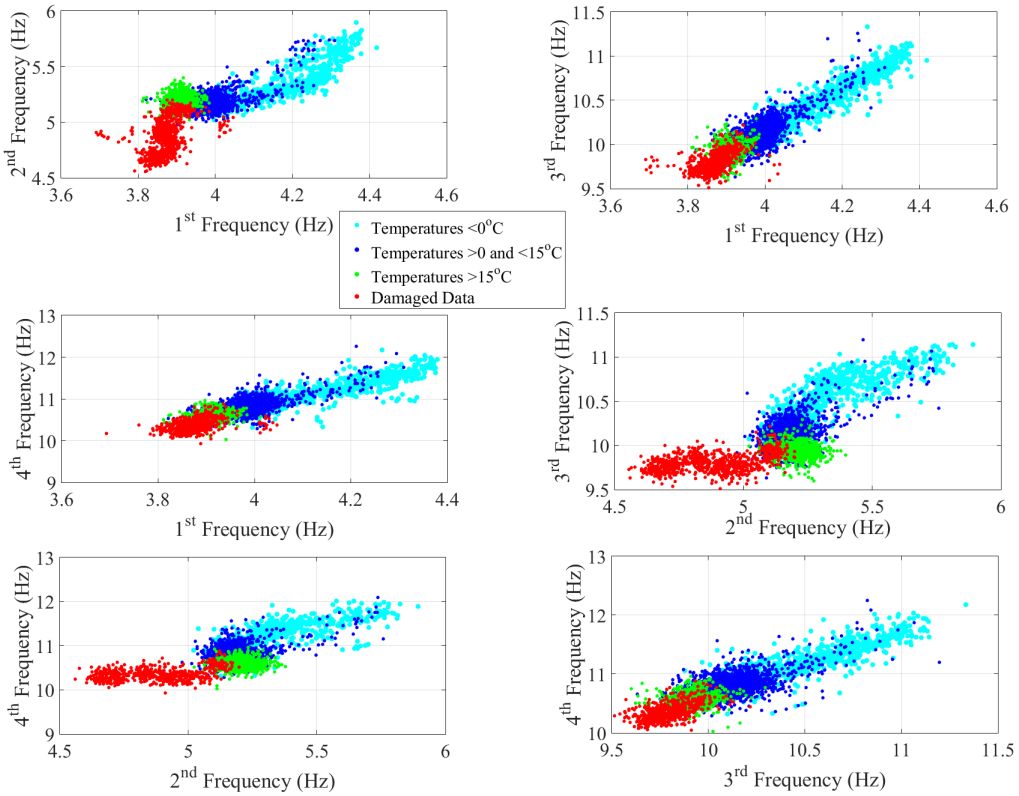


Figure 5.7: Frequency pairs of Z24 Bridge grouped according to air temperature.

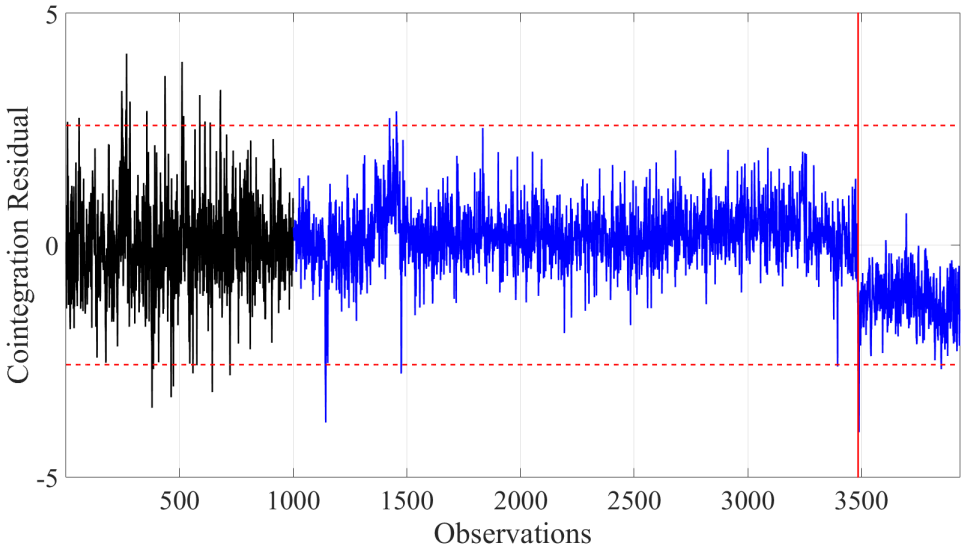


Figure 5.8: Cointegration residual of Z24 natural frequencies. Training region in black colour, while in blue colour is the testing set. Threshold: $\pm 2.58\sigma$ from signal's mean.

cointegration vector, whilst 3 lags are used for VEC model. The VEC model included all four available natural frequency series, as shown in Figure 4.22, and the approximated cointegration vector was: $\beta = (-3.0772, 0.6504, 0.8531, 2.2742)^T$. It should be mentioned that the series are normalised prior implementing cointegration. The damage begins at 3473 data points, as shown in Figure 5.8. According to Akaike Information criterion (AIC) [100], the most appropriate lag length is 6. The latter is used and the cointegration residual is approximated. This is shown in Figure 5.8, where the data in black describe the training region, those in blue are the tested/damaged data and the red vertical line shows the region where damage can be detected. A probable explanation here is that by including the larger artefact on the training set, the cointegration vector was trained to account for higher localised signal variability, that affected slightly the damage detection capability of the residual.

Starting from the training region, outliers can be detected. The number of these depends on the confidence interval adopted. In this case, the confidence interval corresponds to 99%, which described ± 2.58 standard deviations (σ) from the mean of the signal ($\pm 2.58\sigma$). This assumption can be seen as slightly conservative here and an alternative of $\pm 3\sigma$ can be employed, to decrease the number of outliers inside the training region. However, it should be mentioned that the outliers observed in the training regiomn did not hide any form of abnormality/ alarm cause and can be associated with the approximated linear cointegration vector.

Moving now into the testing region, it can be observed that the damage can be detected visually inside the residual after 3500 data observations. Here, it should be mentioned that the training region included the first two large artefacts observed in the natural frequencies before 1000 data observations, but not the largest of all. In particular, the effect of the largest artefact can be observed on the cointegration residual at the region between points 1400-1500, showing that the training set used wa not capable to completely eliminate the nonstationarity introduced by the largest artefact.

On the other hand, if the largest one of the artefacts is included inside the training region (greater length of training region; 2000 data obtained for training), the cointegration residual, as shown in Figure 5.9, provides a quite different result. Inside the training region outliers can be identified, however the impact of the largest artefact was eliminated. Inside the testing region also outliers can be detected, however without having any particular significance in terms of abnormality. The main obser-

vation here is that damage can be detected after 3400 data points, but one can say that by comparing Figures 5.8,5.9, the cointegration residual signal's change of mean value in Figure 5.9 is lower than that of Figure 5.8 inside the damaged region, where the training region considered included 2000 observations. In other words, extreme observations in data inside the training region can affect the damage capability of cointegration and therefore it is important to identify the sources of this variability as well as how one can overcome the associated problems.

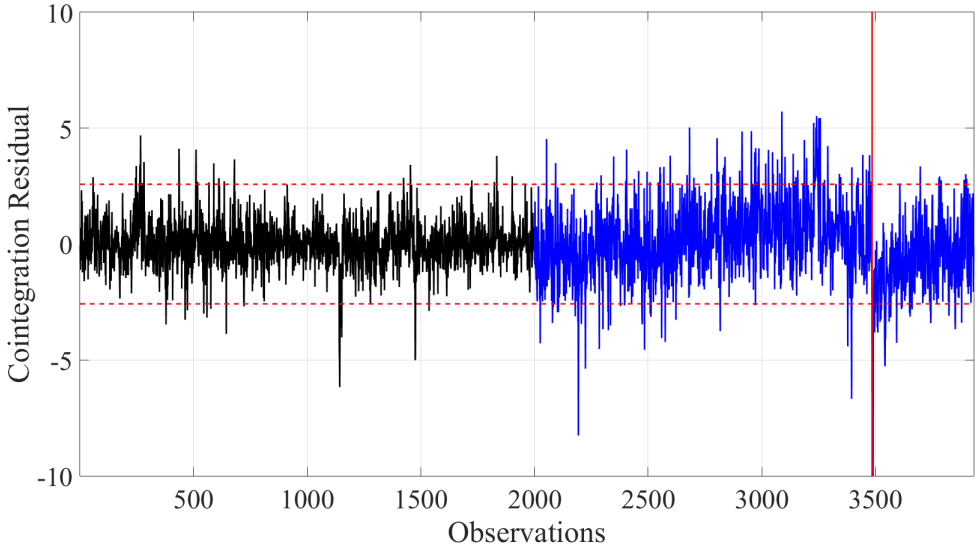


Figure 5.9: Cointegration residual of Z24 natural frequencies including also the largest artefact. Training region in black colour, while in blue colour is the testing set. Threshold: $\pm 2.58\sigma$ from signal's mean.

Furthermore, there are still two additional issues that should be considered to improve the damage detection of cointegration residual. More specifically, some variance can be observed inside the residual that should be removed in order to make the residual signal more stationary, whilst the detected damage is inside the statistical thresholds. These issues are investigated inside the context of non-linear cointegration.

Some initial ideas of using non-linear cointegration are summarised in [2], however the data of $Z24$ are examined more thoroughly in the work of Shi et al. [82]. In particular, in [82] it was described the extension of cointegration to a non-linear context, introducing the idea of a regime switching algorithm. This rearranges the frequency data according to temperature, introducing temperature-based breakpoints, which are tested according to ADF statistic. The result of Shi et al. [82] on Z24 is shown in Figure 5.10. The latter shows how the damage manifested on the cointegra-

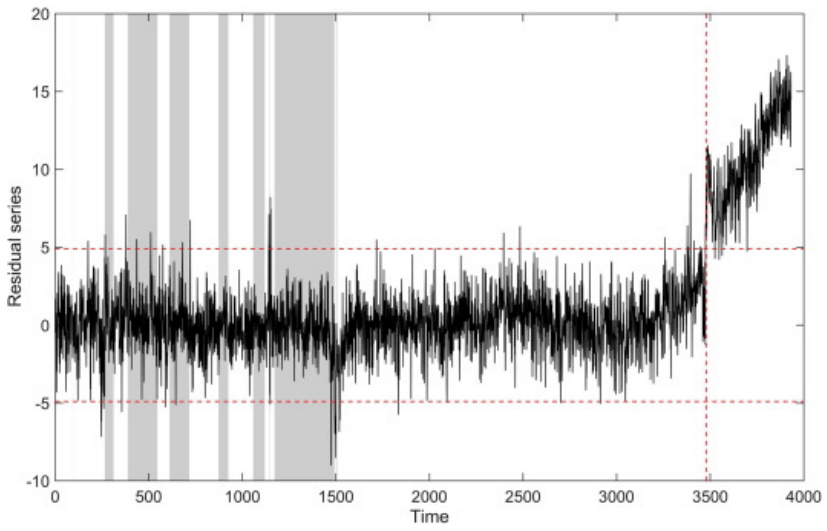


Figure 5.10: Regime switching cointegration results on $Z24$ natural frequency data by *Shi et al.* [24].

tion residual, improving the linear cointegration damage detection capabilities, by showing the damaged data outside the confidence intervals.

5.4.3 Aluminium Laboratory Truss Bridge

The same procedure as the one described on real-time *SHM* Bridge data (Tamar and $Z24$) is demonstrated here for the data coming from the laboratory truss bridge vibration test presented in Chapter 3. The plots of frequency pairs are illustrated in Figure 5.11. In particular, only the pairs of first frequency with respect to others are plotted here, as the results between the remaining pairs are similar. Looking on Figure 5.11, the first four plots describe the normal state relationships. The next four plots describe both the normal state and the first two damage cases, while the last four include also the last damage case.

From the first four plots, it can be observed that the natural frequencies increase as the temperature decreases. Looking into the four next plots, including the first and second damage (removal of the first and second bolt; Chapter 3), the natural frequency increases. As mentioned in Chapter 3, this is expected since the removal of the first two bolts affected only the mass of the structure.

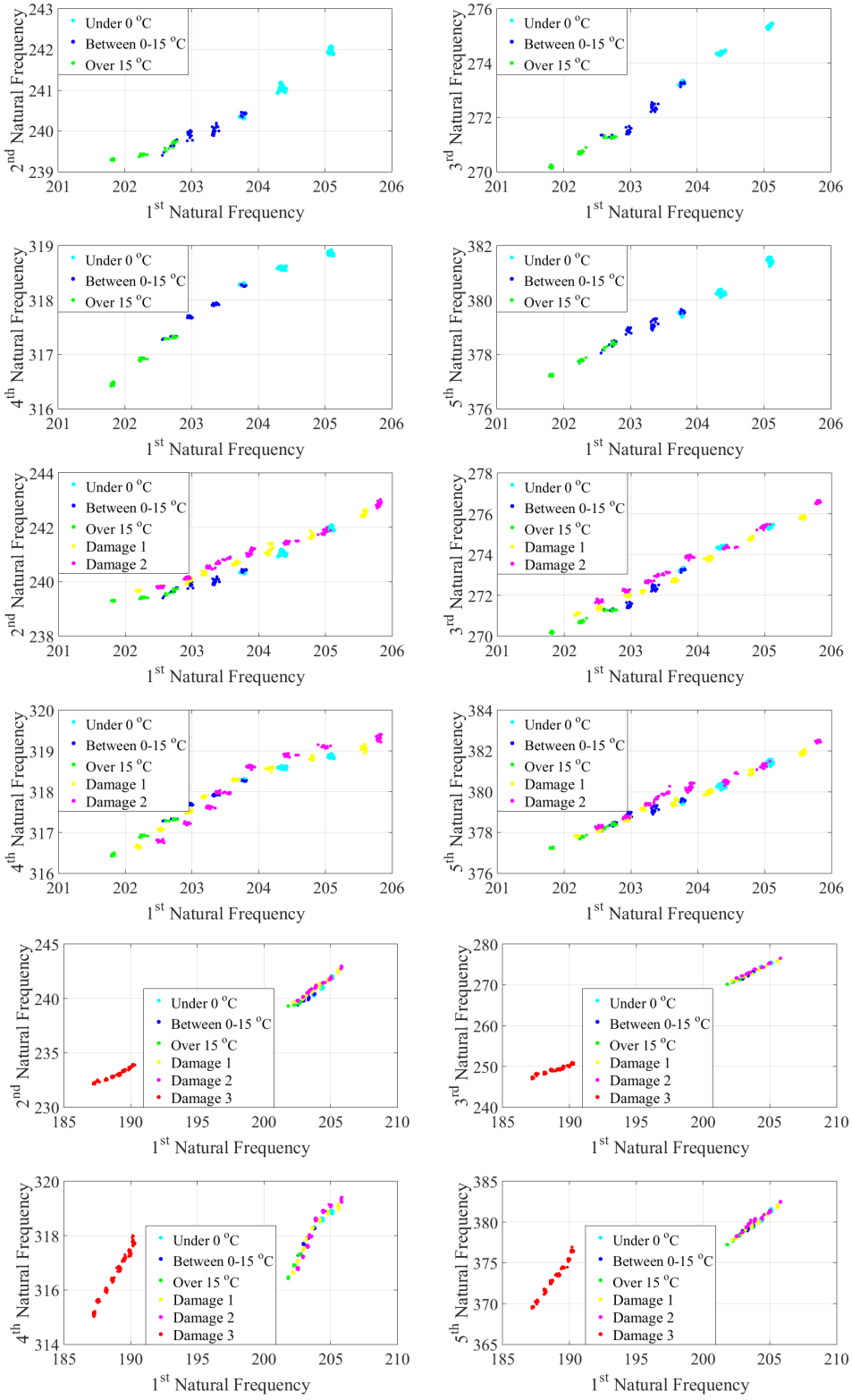


Figure 5.11: Natural frequency pairs for the laboratory truss bridge under different damage states.

In addition, the natural frequencies, associated with damage cases 1 and 2, are still increasing as temperature falls. Finally, in the last four plots, the damaged data (red colour) coming from the removal of the third bolt, show a significant decrease on natural frequencies. As mentioned in Chapter 3, the effect of the removal of the third bolt led to a significant stiffness decrease. Once again the damaged data are highly sensitive to temperature.

Comparing the last four plots, with the plots of the second damage an interesting observation can be made. In the case of second damage, the data are following the same relationship with that of the undamaged data, something that can affect the damage capability of *Johansen's approach to cointegration*. In contrast to the second damage case, in the third one, the damaged data show great distance from the normal data population.

Furthermore, a positive linear correlation between all natural frequency pairs is observed. Therefore, it is expected for cointegration to work successfully and provide a stationary residual, at least inside the normal region, capable for novelty detection. However, it is also expected to have some difficulty to identify damage on the first two damage cases, because the effect of damage is not clearly identified in the natural frequency pairs in contrast to the third damage case, since only mass is decreasing during the two first damage cases (a mass decrease of approx. 0.1% in the first damage case and 0.465% in the second). All these assumptions can be tested performing cointegration below.

Figure 5.12, the cointegration residual shows the normal, damaged and restored normal states, where each state consists of 160 data points. A control chart was set at $\pm 3\sigma$ from the normal's condition residual mean. As expected, using all normal condition data for training, the first 160 data points of the residual are inside the statistical thresholds. In damage state 1, novelty can be detected. The cointegration data corresponding to the first damage are close to normal state, however multiple outliers can be detected. As mentioned previously, at first damaged state an decrease of 0.1% of mass was expected due to bolt removal. For damage state 2, the damage can be observed as the majority of data fall outside the confidence intervals. Here an 0.465% of mass reduction was expected. It should be mentioned that one can expected a slightly greater change between damaged state one and two residuals, due to mass reduction which is much higher in the second damage.

For the third damage state, a more severe damage is observed, which corresponds to

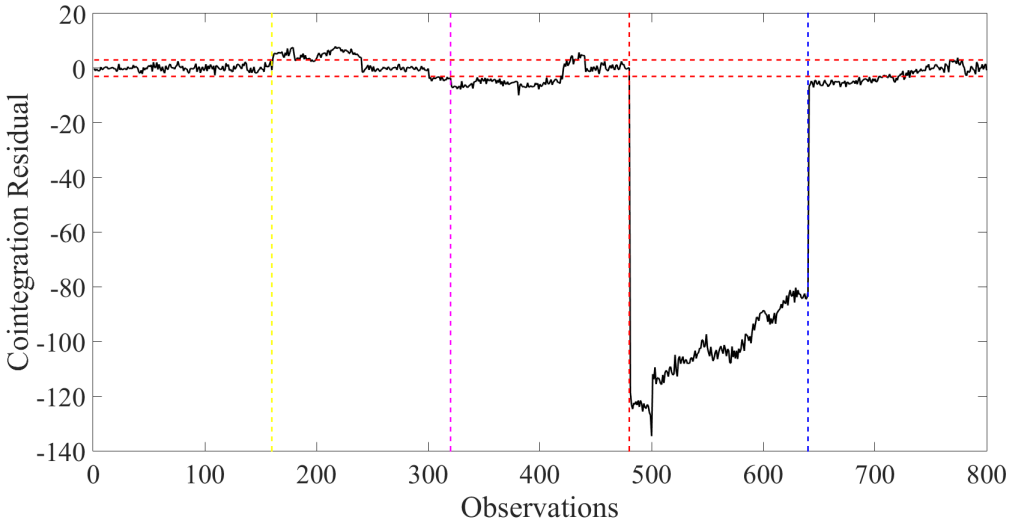


Figure 5.12: Cointegration residual for laboratory truss bridge data. The vertical yellow dashed line designates the end of training and beginning of damaged state 1, the magenta dashed line the end of first damage and beginning of the second, the red dashed line the end of damage 2 and beginning of damage 3 and finally the blue dashed line the end of third damage and the beginning of the system restoration region. Thresholds for this case were set at $\pm 3\sigma$ from mean are applied.

stiffness decrease occurring due to the removal of the third bolt. This led to the loss of contact between the lower chord and the internal diagonal of the truss. This result was as expected from the initial testing procedure. Finally, the restored condition shows that the system is slightly outside the statistical thresholds and then comes back to the normal condition.

An interesting result for comparison is provided in Figure 5.13. In this case, the first 80 data points of normal condition was set as training region, for cointegration vector training, whilst the remaining 80 data points as testing region. It is clear that inside the testing region and especially for recordings at temperatures under zero degrees (120-160 data points), a shift can be observed, however inside the statistical thresholds. One can argue here that this is the consequence for not considering a full temperature cycle (25 to -10^oC) in order to train the cointegration residual more accurately. It is true that a greater normal condition, i.e. including two or more full temperature cycles could have provided sufficient data for the normal testing condition.

However, the cointegration residual for the first damage state, shows that this can be considered as a normal condition as all points are inside the statistical thresholds.

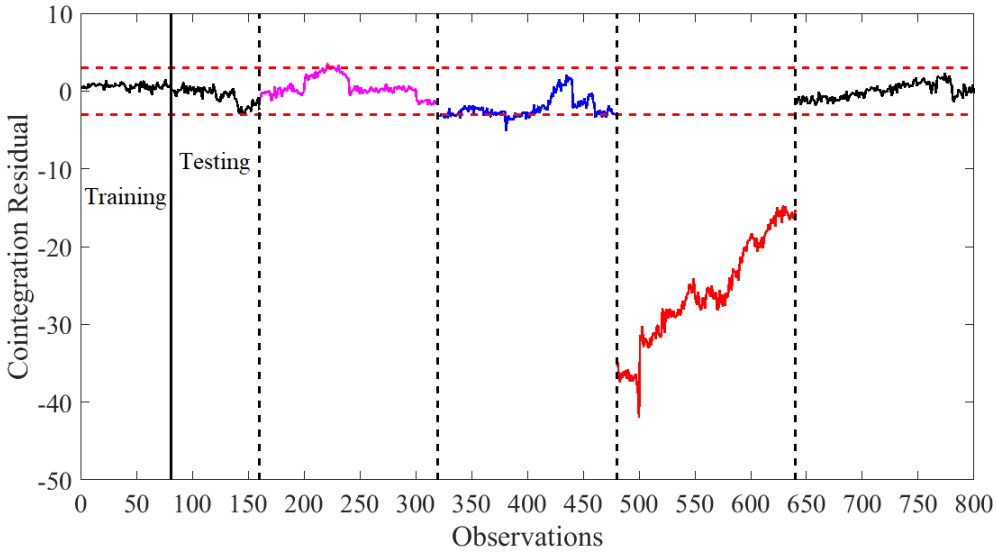


Figure 5.13: Cointegration residual for laboratory truss bridge data. The vertical black line designates the training from the testing normal condition data, whilst the remaining dashed vertical black lines designate the change of states. The data points in black describe the normal condition, those in magenta the first damage, the blue ones the second damage, the red ones the third damage and finally the last black data are the restored normal condition data.Thresholds for this case were set at $\pm 3\sigma$ from mean are applied.

In the second damaged state, outliers can be detected to the cointegration residual, which can justify the mass reduction due to bolts and angle bracket. The third damage state, as expected, shows that the system stiffness decreased significantly. Finally, the restored normal state shows that all data observations of residual are inside the statistical thresholds informing there is no damage in the system.

From the comparison of these two cointegration residuals it was observed that the second one including both training and testing region provided more reasonable results with respect to the initial assumptions made during the experimental process. It was also shown that the inclusion of more data inside the training region does not always lead to a more informative cointegration residual, as a more informative cointegration vector identified including less data observations of the normal state in this case.

From the laboratory test, it can be said that damage identified in all cases tested and that *Johansen's approach to cointegration* worked successfully. In addition, seems that small-scale structures are more sensitive to temperature, even for temperature

changes occurring at $\pm 0.5^oC$ level. These led to fast changes (rate effects) on the natural frequencies, which can be associated with small changes in system's underlying physics. Such small changes observed inside all damage states and violated the cointegration relation established for the normal state.

5.4.4 Conclusive Remarks from Bridge Data

From the analysis performed in Section 5.4, some important conclusions can be drawn. Starting from the case of cable tension signals of Tamar Bridge, it was shown that air temperature is the main driver of their variability, while also signals sharing inverse proportional trends can be used together inside *Johansen's approach to cointegration* by providing a stationary cointegration residual. In addition, an example is provided where cointegration can be used for novelty detection when employing multiple signals. Furthermore, in the case of Z24 natural frequency data, it was shown that the impact of EOVs on the examined real-time bridge data manifests itself in different manner than that of damage. In particular, it seems that the effect of temperature on the data, in the majority of cases, exhibits a deterministic pattern, since frequencies tend to increase in low temperatures and decrease in higher. However, this pattern might not be always visible due to the sensitivity of features to other EOVs (i.e. Tamar vehicle mass [2]). In addition, the importance of training set selection discussed providing an example on the Z24 natural frequency data. In particular, the cointegration residual was observed to be more damage sensitive when the training region include only the first two large artefacts of natural frequency and not the largest one. Moreover, from vibration analysis and damage simulation experiment conducted on the laboratory truss bridge, it was shown that small-scale structures' response is more sensitivity to EOVs than real-time structures. For the damage cases examined, it seems that the existence of significant damage can lead to the projection of damage data in a space exhibiting greater distance from the normal state data. Finally, it was observed that even in case of damage, the damaged feature keeps, more or less, its initial temperature sensitivity (frequency increase as temperature falls).

5.5 Conclusions

In this Chapter, the three steps (ADF, VAR and VEC models) included inside *Johansen approach to cointegration* are presented. After that, a discussion was made about the way that common trends are manipulated inside cointegration and what happens in case of damage. In particular, it was shown that cointegration does not learn the common trends but learns the correlation structure between the features and then cancels it by subtraction. Furthermore, the SHM data coming from the Tamar Bridge, the $Z24$ Bridge and the laboratory truss bridge experiment are explored and used to demonstrate the application of *Johansen's approach to cointegration*. More specifically, the interrelationships between SHM series are examined, focusing on how these are affected by EOVs (i.e. temperature) and damage, as well as the impact that these have on the cointegration residual. From all cases it was shown that cointegration performs well in terms of damage detection. An additional interesting observation made was that when a linear negative (inverse proportional trends) or positive correlation is existent between the series considered from cointegration, it is possible to obtain a stationary cointegration residual. Furthermore, it is important to identify the most appropriate training region for cointegration. An example given here on the data of Z24 bridge. Further improvements to cointegration damage detection capability can be achieved employing non-linear cointegration (*Shi et al.* [82]) approaches. Moreover, it was shown that in the case of a small-scale truss bridge, the effect of temperature is greater on natural frequencies and especially during small temperature variations ($\pm 0.5^oC$). The latter can lead to fast changes in frequency variation, which are referred to as rate effects [105]. These were imposed by EOVs and subsequently can change the underlying physics of the system, leading to a direct impact on cointegration residual, as the normal state equilibrium is violated (Figure 5.12).

The main aim behind all the aforementioned was to test the initial hypothesis made, which is that in the presence of linear positive or negative correlation between series, the linear cointegration analysis, as described inside *Johansen's approach*, can be applied successfully. The latter, in all bridge SHM data cases examined previously, has been associated with the impact of temperature on the SHM series, which led to the introduction of time trends on the series. According to the description of *Johansen's approach to cointegration* [74, 75], the series in order to be cointegrated successfully should share common trends and be nonstationary and more specifically

of first order of integration I(1). The test used to judge if a signal is stationary or nonstationary, both inside *Johansen's approach to cointegration* and here, is the Augmented Dickey-Fuller (ADF) test [94, 95]. Although, this is used commonly inside Econometrics and used previously inside SHM, it will be useful to investigate it further and understand how exactly works and what attributes of each signal counts in order to make the judgement. In particular, the latter is attempted in the next Chapter, in which an interpretation of ADF statistic is provided based on basic ideas of signal processing and dimensional analysis.

Chapter 6

On Stationarity and the Interpretation of the ADF Statistic

The last Chapter demonstrated how many of the features that are commonly used are nonstationary in nature. On the other hand, this Chapter considers the nature of stationarity of a time series or signal, and how it may be quantified. It is argued that a subjective assessment is as effective as one based on mathematical definitions, if one actually has finite samples of data, and that such an assessment is fundamentally based on the number of cycles of the dominant periodic component visible in the sample. It is shown by dimensional analysis that one of the most often-used measures of stationarity, the Augmented Dickey-Fuller (ADF) statistic, supports this hypothesis. This analysis can be useful for anyone concerned with time-series analysis and nonstationarity.

6.1 Introduction

The concept of *p*rediction is central to science and engineering, also inside economics and social sciences. Especially in engineering applications such as SHM, which is the main scope of this thesis, the procedure of *p*rediction becomes difficult when the physical laws governing a system change over time. In many situations of interest

in engineering, one does not know the exact physical laws and is forced to try and infer some or all of them from measured data, which essentially is the problem of system identification [106]. Therefore, it is important to know if a system changes over time in order to model and predict its response more sufficiently. The latter is associated with the concept of nonstationarity, which originates from the analysis of stochastic processes.

Suppose the set of observed signals is a vector time series $x(t)$ governed by a stochastic process X_t. The process is *strictly stationary* if the joint probability density $p_t(x(t))$, is independent of time i.e.

$$p_t(x(t)) = p_{t+\tau}(x(t+\tau)), \qquad \forall \tau \tag{6.1}$$

where the variables (t, τ) can be discrete or continuous. Although this definition is straight-forward from a mathematical perspective, its application in engineering is more complex, because it is framed in terms of multivariate density functions, which are difficult to estimate from data [107, 108]. Thus, other indicators of stationarity are needed, which will be based on quantities that can be observed inside data samples.

In the recent past, the idea of stationarity has become central to another engineering discipline, that of *Structural Health Monitoring* (SHM) [7]. The basic idea of SHM is to infer from measured data, whether a system or structure has become damaged during operation. For various reasons (discussed in detail in [7]), SHM is often implemented in practice by trying to determine if the measured data changes its nature at some point (i.e. when the structure moves from the undamaged state to a damaged state). Such a change is clearly an instance of nonstationarity; however, the problem is complicated by the fact that signals from structures may be nonstationary for entirely benign reasons. To give a concrete (no pun intended) example, consider a bridge to be the structure of interest. Suppose that the measurands/features of interest are the first few natural frequencies; these are known to be sensitive to damage because reduced structural stiffness (say, as the result of cracks) will cause the natural frequencies to be reduced. Unfortunately, these same frequencies are also sensitive to the ambient temperature of the bridge, the traffic loading etc. These benign operational and environmental variations are known as *confounding influences* and represent one of the main technological problems in SHM implementation. Various methods have been proposed for the elimination

of confounding influences [7]; one of the most powerful is based on the concept of *c*ointegration [76, 109].

The theory of cointegration arose in the domain of econometrics, where it was mainly used in order to analyse and understand common trends in economic and financial time series. However, one application of the theory is to combine multiple non-stationary time series in such a way as to produce a stationary residual; this is the application that proves to be useful for SHM. The idea is that confounding influences in structural response measurements often manifest as common trends between signals i.e. the daily variations in temperature of a bridge will cause slowly-varying (compared to the dynamics) common trends in estimated natural frequencies; cointegration allows combination of the signals in such a way that the result is stationary, and sensitive only to changes due to structural damage (or previously unseen benign effects). One of the necessities of the cointegration technique, is a means of deciding if a signal is stationary or not, and this is often based on a hypothesis test using specialised statistics – the most common being the ADF statistic [94, 95]. In fact, the ADF statistic works in the abstract sense by detecting unit roots in the set of characteristic roots of the time series of interest. An important objective of this paper is to provide a simple interpretation of the ADF statistic and show that it makes contact with engineering intuition about the nature of stationarity. In fact, it will be argued that the decision on whether a time series is stationary or not is a matter of subjective judgement.

The layout of this is as follows: in the first section the nature of stationarity is discussed, as well as a hypothesis on how a human expert might judge stationarity. In the second section, it is demonstrated that the ADF test statistic is a function of the dominant frequency of a mono-harmonic signal, its length and its signal to noise ratio, via a dimensional analysis. The third section is the results section, which includes the simulations performed in order to illustrate the aforementioned point. Finally, the results will be discussed and conclusions on the use of ADF inside SHM context will be drawn.

6.2 Practical Tests for Stationarity

As mentioned in the introduction, the mathematical definition of a *strictly stationary* stochastic process is quite straightforward. However, it is framed in terms of

multivariate density functions, something that is difficult to be estimated from data. As the main aim is to perform a stationary assessment on the engineering signals obtained from SHM, the concept of stationarity should be formulated in terms of quantities which can be estimated more easily. Inside SHM, damage detection can be performed on series exhibiting stationarity using statistical process control and other control charts [110]. In such a case the definition of stationarity in *weak sense* [1] can be useful. According to [111], a signal can be considered as *weakly stationary* if the mean of the process is constant i.e.,

$$m(t) = E[X_t] = m \quad (constant) \tag{6.2}$$

(where E denotes the expectation operator) and the auto-correlation function of process depends only on time differences, i.e.

$$E[X_s X_t] = \phi_{xx}(s,t) = \phi_{xx}(s-t) \quad \forall s,t. \tag{6.3}$$

The second of these conditions can be replaced by the requirement for a constant covariance.

Although this definition about stationarity is helpful, one can only compute sample estimates for $m(t)$ and ϕ_{xx} from $f$inite data samples. As expected, the estimates of the mean and variance from a finite sample will be constant and one can track the mean and variance over a sequence of moving windows. A minor issue is the determination of the window's width. The main issue is that engineering signals are not stationary, in most of the cases, something that makes their assessment according to mathematical determination of weakly stationarity very complex. More specifically, the SHM signals can be considered as *periodically-correlated* or $c$yclostationary [111]. A second-order stochastic process X_t is periodically-correlated if for all s,t,

$$m(t) = m(t+T) \tag{6.4}$$

and,

[1] Detection using a control chart depends critically on the tails of the signal probability density, and these are independent/insensitive to the central statistics controlled in weak stationarity.

$$\phi_{xx}(s,t) = \phi_{xx}(s+T, t+T) \tag{6.5}$$

for some *period* T and there are no smaller values of $T > 0$ for which these conditions hold. Even in this case, the application of these criteria to a finite data sample is difficult.

To justify the aforementioned point an illustration can be considered. For this the estimated fundamental natural frequency series, Figure 6.1, of the Z24 Bridge is used, describing a whole year of monitoring [4]. Looking at Figure 6.1, one can assume that the signal of the frequency can be decomposed into two components: a high-frequency one, which is associated with its noisy part, contributed to the variance in frequency estimation and a lower-frequency one, which can be associated with the large artefacts resulting from very low temperatures occurred during monitoring. Looking at Figure 6.1, an analyst, based on his/her engineering judgement, can conclude that the natural frequency series is nonstationary. Considering, that the cold periods in the region concerned can occur at roughly the same points each year, the response of the bridge can be simulated for 50 years by repeating figure 5.1, 50 times, as depicted in Figure 6.2. In this case, the analyst can conclude that due to the repetition of the annual periodic component, the signal of 50 years can be considered as cyclostationary. Although, the aforementioned two judgements are clearly subjective, however there is a high likelihood to be done by an expert.

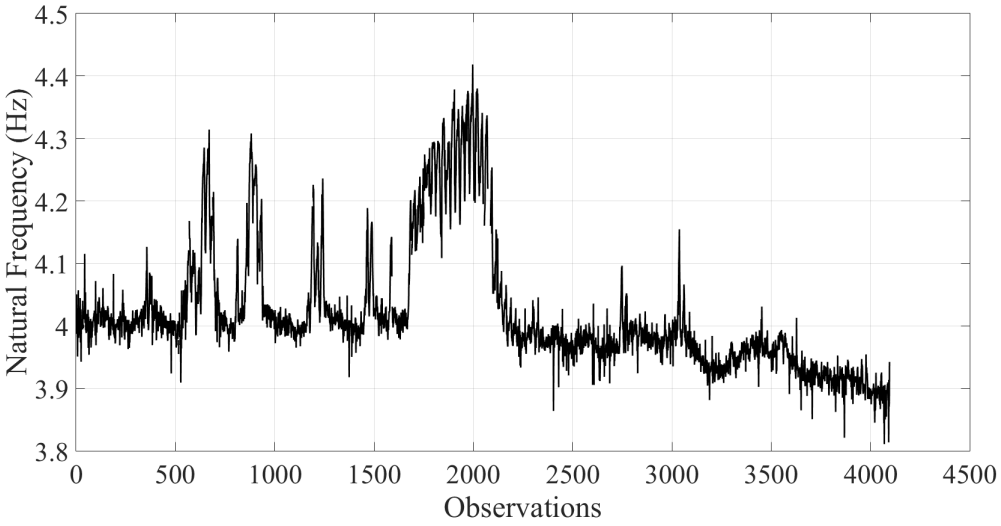


Figure 6.1: Fundamental natural frequency of the Z24 Bridge corresponding to a single year of monitoring.

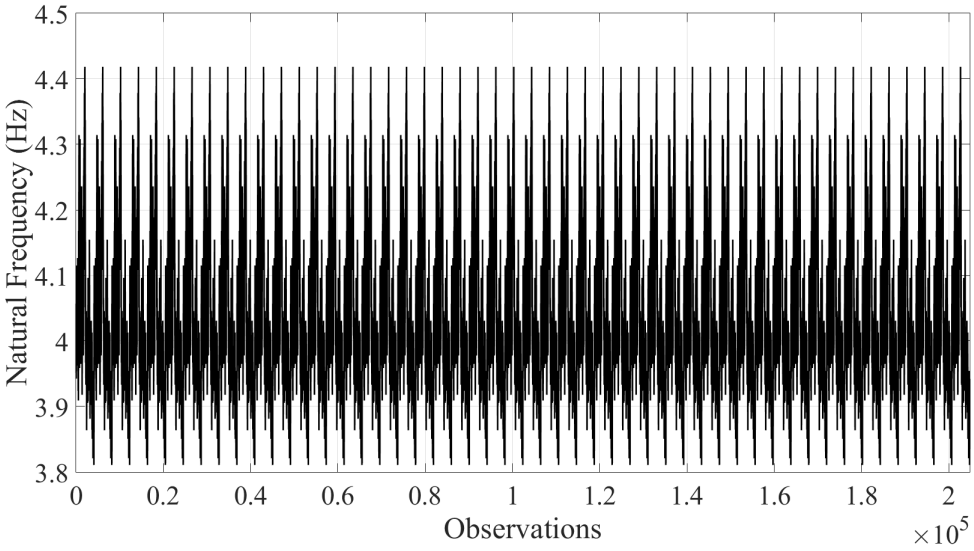


Figure 6.2: Fundamental natural frequency of the Z24 Bridge, the single year repeated 50 times to simulate measurements over a 50-year period.

The problem in SHM is that one wishes to detect damage from signals like the natural frequencies by detecting changes from a stationary response using tools like control charts. Our current best approach is to look at features/data measured over some finite window and to try and make a judgement as to their stationarity; if they are stationary, it is valid to use control charts etc; if they are nonstationary, one needs to use cointegration first etc. It has already been noted that it is difficult to apply mathematical definitions of stationarity to finite samples of data, and it could be argued that the subjective judgements or engineers/experts might be just as reliable. However, engineers and scientists (rightly) distrust subjectivity as a rule, and they have produced statistic tools in order to try and render the judgement objective.

Inside Econometrics there are several tests that can be used in order to assess if a given signal is stationary or not. Some of the most commly used are the ADF test [94, 95], the PP [96] and KPSS [97] tests. These tests follow the concept of *u*nit root hypothesis, which described in depth in Chapter 4. In the current Chapter the ADF test will be further investigated, because this forms a part of Johansen's approach to cointegration and also the results between the aforementioned econometric tests are expected to be similar. It is useful to add that inside the context of Econometrics there is criticism about the validity of unit root hypothesis [97, 99, 112], which is mainly based on the fact that a unit root will never be found with complete precision

in order to judge with confidence if a signal is stationary or not.

The existence of proxies such as ADF statistic can be an excellent indicator for testing the nature of a SHM signal, if scientists and engineers trust ADF indications. To do so, the indications of these proxies should be in line with their intuitions about signals. Hence, the question now is, what is the basis of expert intuition on the nature of a finite sample of a time series. Here, it seems that the main criterion can be the number of periods of the dominant frequency observed in a sample. According to the example presented for the first natural frequency of Z24, an analyst presented with data from one period or less, assuming that noise can be identified easily, can judge the signal as nonstationary, while in case of multiple periods of the dominant frequency can be judged as cyclostationary or stationary. Considering the latter as the basis of expert judgement, proxies such as ADF can be tested if they respect this basis.

In econometrics literature the main considerations affecting unit root tests are: the sample size, the sampling frequency and the overall time duration of the sample [113–118]. These observations will be captured in the simple dimensional argument presented in the following section.

6.3 Dimensional Analysis of the ADF Statistic

6.3.1 Analysis

For the purpose of the analysis of this section, some simplifying assumptions are made about the signals under consideration. The signal examined will be,

$$x(t) = A \sin\left(\frac{2\pi t}{T}\right) + \epsilon(t) \tag{6.6}$$

which represents a sine wave with amplitude (A), cyclic frequency $(2\pi/T)$, where T is the period and ϵ is a zero-mean white Gaussian noise sequence with variance σ_ϵ^2. The term N describes the number of observations of the sample, which is sampled every Δt seconds. Hence, the sampling frequency will be $f_s = 1/\Delta t$, while the the dominant frequency will be $f = 1/T$. The *normalised frequency* corresponding to f will be defined by $f_n = f/f_s$.

According to [119], the ADF statistic (t_{ADF}) can be calculated by estimating the root ($\hat{\rho}$) of an AR process and dividing it by the standard error of that estimate (σ_ρ). The latter is given in equation (6.7). As the t_{ADF} is by definition dimensionless, it can be expressed as function of any other appropriate dimensionless group, $\pi_2 \dots \pi_n$, defined for the problem of interest, according to Buckingham-Pi theorem [120] (Equation (6.8)).

$$t_{ADF} = \frac{\hat{\rho}}{\sigma_\rho} \tag{6.7}$$

$$t_{ADF} = \pi_1 = f(\pi_2, \dots, \pi_n) \tag{6.8}$$

Under the previous assumption, the quantities of interest that can be considered are: the dominant period of the signal T, the sampling interval Δt, the number of sample observations N, the signal variance σ_y^2 and the noise variance σ_ϵ^2. Despite the fact that there are five variables under consideration, only two dimensions are represented: time $[T]$ and length (amplitude) $[L]$. This means that two primary magnitudes for the dimensional analysis are needed, which are the noise variance σ_ϵ^2 ($[L]$) and sampling interval Δt ($[T]$). On the other hand, three secondary variables or dimensionless groups are possible, formed from T, N (already dimensionless) and σ_y^2. Thus, a valid set can be formed as: $\pi_2 = \Delta t/T = f/f_s$, $\pi_3 = N$ and $\pi_4 = \sigma_y^2/\sigma_\epsilon^2$. It follows from the Buckingham-Pi theorem that,

$$t_{ADF} = f(\frac{\Delta t}{T} = \frac{f}{f_s}, N, \frac{\sigma_y^2}{\sigma_\epsilon^2}) \tag{6.9}$$

where the first variable is simply the normalised frequency of the dominant periodic component, the second is the number of observations (N) of the series examined and the third is a signal-to-noise ratio (SNR). Equation ((6.9)) identifies an exhaustive set of variables that the ADF statistic can depend on. To examine the dependence of t_{ADF} on each of the aforementioned variables, simple simulations using the signal of equation (6.6) will be performed and their results will be investigated.

6.3.2 Simulations and Results

Based on the previous dimensional analysis assumptions, three groups of simulations are designed to examine the dependency of t_{ADF} on the number of sample observations, the normalised frequency of the dominant periodic component and the signal-to-noise ratio. To begin with, the first group of simulations is concerned with the effect that the number of observations has on the t_{ADF}. In particular, to investigate the latter, SNR is set at 100 (a low noise amount is necessary otherwise t_{ADF} conditioning issues may arise when fitting the AR models), the dominant period T and ΔT are fixed by generating cycles of sine wave with 128 points per cycle. This leaves only N to vary, and this is accomplished simply by concatenating k cycles of the sine wave, from 1 to 100. Figure 6.3 shows the sample with 10 cycles.

Figure 6.4 shows the ADF statistic plotted against the number of cycles of the dominant periodic component. In agreement with the intuition that greater stationarity is supported for mode periods of the sine wave, the ADF statistic falls monotonically with number of cycles. This simulation motivates three points:

1. Because the statistic falls monotonically, there is a critical number of cycles k_{crit} beyond which the signal is judges stationary. In the case shown, this number of cycles is around 18, as indicated in Figure 6.4. (This is actually a function of the number of points per cycle; however, the objective here is to show how the ADF statistic varies with each of the main dimensionless groups in turn, holding other variables fixed. The quantitative behaviour of the statistic-i.e. monotonic decrease- is the same for different number of points per cycle.)

2. Instead of using $\pi_3 = N$, it is arguably more meaningful to use $\pi_3 = N\Delta t/T$ as this simply counts the number of cycles of the dominant period of the sample. This is the quantity which was identified above as the prime mover in intuition about stationarity. The ADF statistic falls monotonically with this quantity as expected

3. Here it is important to explain why the previous observation is happening. In simple words, the t_{ADF} is proportional to an AR coefficient estimate and therefore adding the data in the form of more identical cycles leads to a monotonic decrease of the standard error, equation (6.7), (until hitting a Cramer-Rao bound [121, 122]), which however does not change the underlying AR process.

This means that the t_{ADF} grows in magnitude (i.e. becomes more negative) and judge the sample more stationary.

Therefore, the basic relation now is,

$$t_{ADF} = f(\frac{f}{f_s}, \frac{N\Delta t}{T}, \frac{\sigma_y^2}{\sigma_\epsilon^2}) \tag{6.10}$$

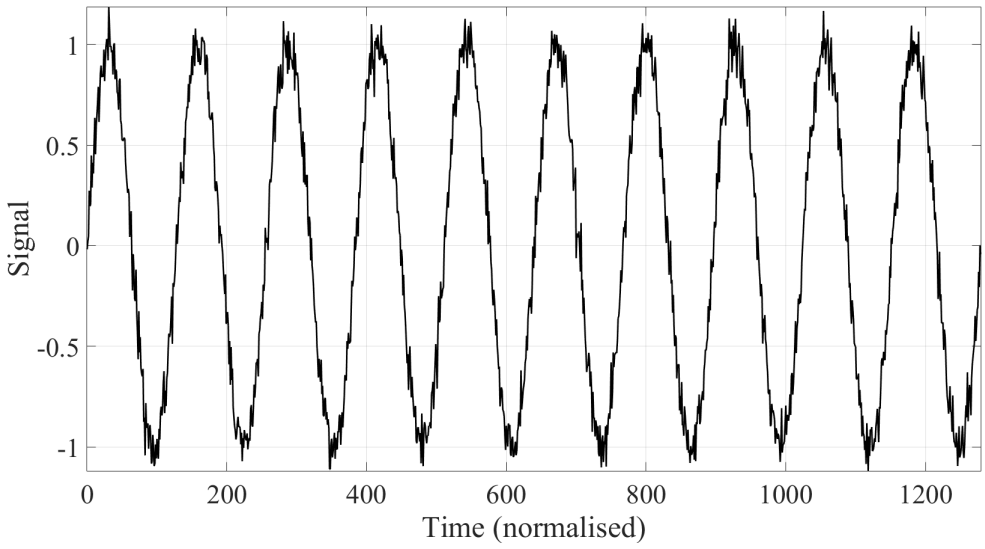


Figure 6.3: 10 cycles of the slightly noisy sine wave.

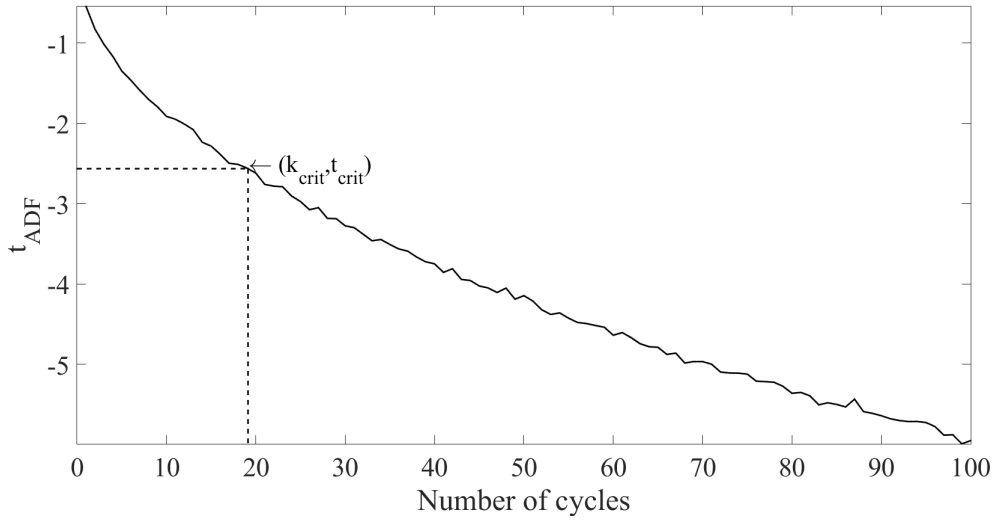


Figure 6.4: ADF statistic as a function of the number of cycles of the dominant periodic component in a sample.

The second set of simulations is concerned with the dependence that t_{ADF} has on f/f_s. Clearly one can increase the number of periods of the dominant frequency in a sample of fixed length by increasing that frequency or by decreasing the sampling frequency. Under the hypothesis expressed above, the signal would then appear more stationary and one would expect the t_{ADF} to decrease.

In the next simulation, the SNR is fixed as before, and the number of points N, is fixed at 104. The sampling frequency of sample is varied by moving by moving from one extreme when all 1024 points form one sine wave cycle, to the other extreme when 256 cycles are present, sampled at 4 points each (i.e. at half the Nyquist limit). Figure 6.5 shows that the ADF statistic decreases monotonically as the sampling rate decreases (or conersely as the dominant frequency increases) and this is once more consistent with the hypothesis that the ADF statistic essentially counts the number of cycles/periods of the dominant frequency component in a sample.

The last simulation reveals something rather interesting. For a fixed N, the ADF statistic falls monotonically with increased normalised frequency of the dominant component. This implies that, corresponding to a given critical ADF statistic t_c, there is a critical normalised frequency f_c such that, signals with normalised frequencies greater than f_c will be judged stationary, while those with $f < f_c$ will be judged nonstationary. For $N = 1024$ the critical ADF statistic is -2.56, and this corresponds to a critical normalised frequency of 0.015 (Figure 6.6). Given the sampling frequency in any given situation, the critical frequency in Hz is easily derived from the critical normalised frequency, which is a function of N.

Although it proved useful to redefine π_2 to count the number of periods of the dominant harmonic component, there is actually another interesting effect associated with varying N. This effect is revealed by varying the number of points in the record N, but fixing the signal so that it only contains one period of the harmonic component. Figure 6.6 shows what happens to the estimated ADF statistic in this case. The arguments based on engineering intuition would suggest that an expert faced with a single period of the waveform would judge it to be nonstationary, regardless of the number of sample points. In fact, this idea is supported by the ADF statistic, up to a number of points N_{crit}, beyond which the ADF statistic judges the signal to be stationary. While this is initially counter-intuitive, it does make sense. As N increases, keeping only one cycle of the harmonic present, the sampling frequency increases monotonically and the normalised frequency of the harmonic decreases monotonically. In the limit as f_n tends to zero, one has a constant signal,

which one would clearly have to judge as stationary.

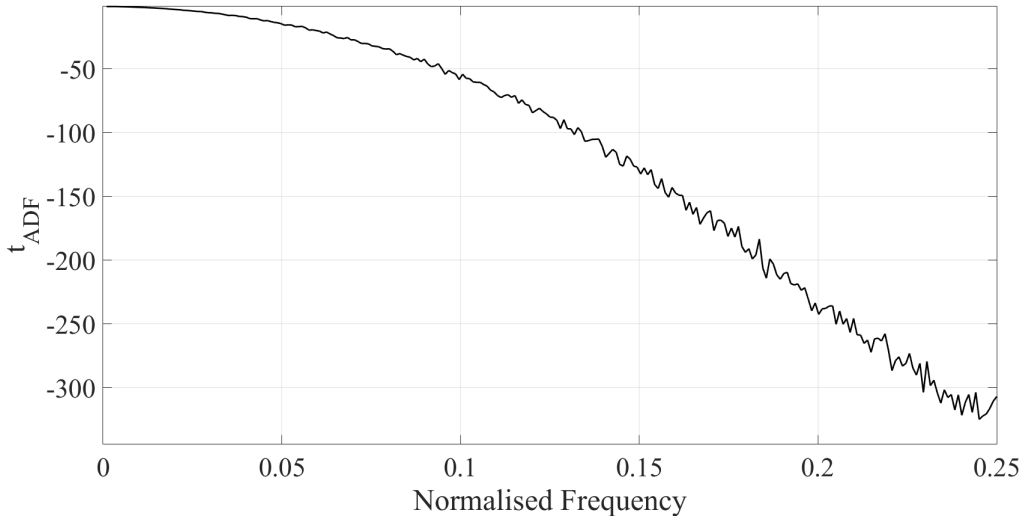


Figure 6.5: ADF statistic versus normalised frequency.

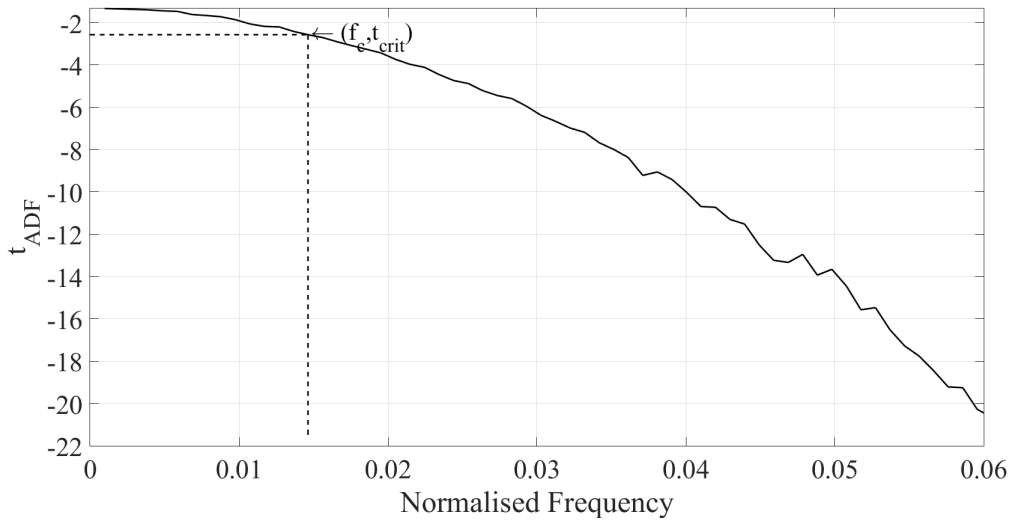


Figure 6.6: Zoom of ADF statistic versus normalised frequency, indicating presence of critical normalised frequency.

Finally, the last simulation tries to identify the effect that the amplitude of the signal has on the ADF statistic. Due to the fact that ADF is dimensionless, there is no direct dependency with signal amplitude. However, the latter can be investigated through the variation of SNR, $\sigma_y^2/\sigma_\epsilon^2$. For the signals presented here, the latter corresponds to the ratio of signal over the noise power. It is expected that when the signal is dominated by noise, the ADF statistic will provide inference for stationarity, while if the periodic component dominates, inference for nonstationarity. In other

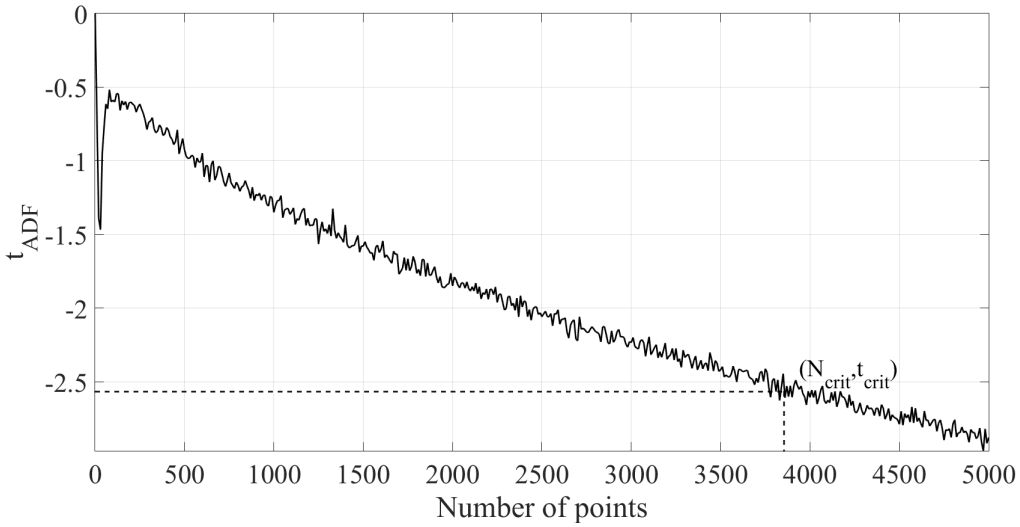


Figure 6.7: ADF statistic versus N, keeping a single cycle of the dominant frequency, indicating presence of critical number of samples beyond which the signal is judged stationary.

words, it is expected that stationarity will increase in monotonic manner when the SNR increases. The latter expectation is supported by the results shown in Figure 6.8. Of course, one could also define a critical normalised frequency with respect to SNR. The latter case occurs when the signal is a noisy sine wave.

Of course, one could also define a critical normalised frequency with respect to SNR; however, this would be less interesting as it would apply only to the very restricted case of the signal being a noisy sine wave. In the general case, the ADF statistic will depend on the relative sizes of periodic components and the noise components. If periodic components with frequencies below the critical frequency dominate other components, one would expect the ADF statistic to judge the signal stationary. If the signal is a mixture of a narrowband random process and white noise, it might be possible to extend the arguments presented here by defining the 'dominant frequency' in terms of the peak in the spectrum or as the frequency centroid for example.

6.3.3 Discussion

The results so far support the hypothesis that the ADF statistic essentially counts the number of cycles of the dominant periodic component visible in the sample; it falls monotonically as this quantity increases. A second result of the analysis is

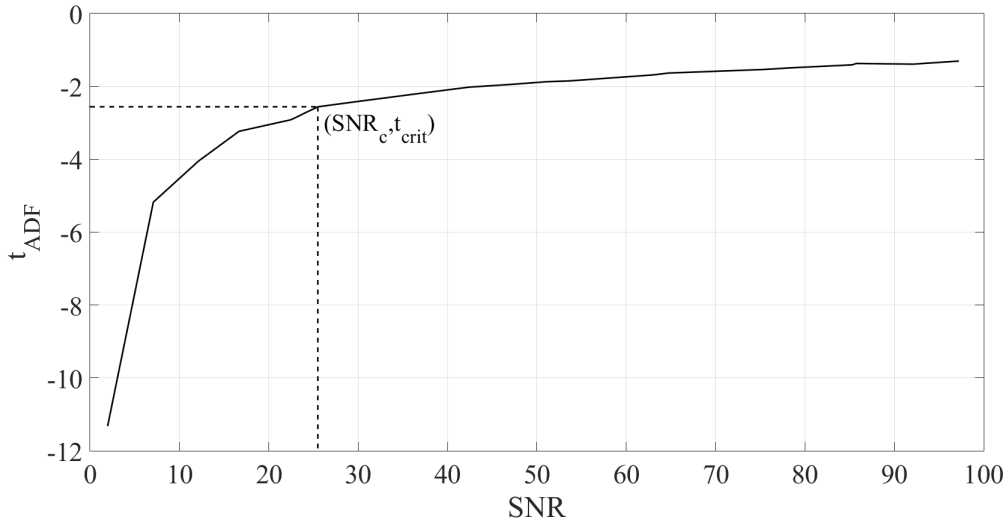


Figure 6.8: ADF statistic versus signal-to-noise ratio.

that there exists (at least for monocomponent signals + noise), a critical frequency f_c corresponding to the critical value of the ADF statistic for a fixed number of samples. This point requires a little more discussion, due to the fact that there is more than one critical ADF value. The critical value depends on the type of model assumed for the underlying process [119]. To simplify matters, the discussion here will be restricted to the Dickey-Fuller (DF) statistic, rather than the ADF, but the point is the same. In the DF test, three common models are considered:

1. Simple AR model without constant (c) or time-trend (T): this represents a random walk with no drift and its structure is given by,

$$y_i = \rho y_{i-1} + \varepsilon_i \tag{6.11}$$

   where ε_i is the residual sequence.

2. AR model with constant (c) but no time-trend (T): the model structure is given by,

$$y_i = \rho y_{i-1} + c + \varepsilon_i \tag{6.12}$$

3. AR model with both constant (c) and time-trend (T): the model structure is given by,

$$y_i = \rho y_{i-1} + c + Tt + \varepsilon_i \tag{6.13}$$

   where both c and T are constants.

Each of these models will generate a different t_c, and the same is true of the ADF test, which really only adds certain AR terms to DF model structures. So, for the situations above where the samples contain $N > 500$ points, and at 99% confidence, the ADF t_c values for the corresponding types of models (no constant, no trend etc.) are (1) -2.56, (2) -3.437, (3) -3.967. None of these models seem to completely fit the situation here where the deterministic components of the time series are periodic; however, model (1) has been chosen for its simplicity. The different choices do not affect the main conclusions relating to the monotonic behaviour of the ADF statistic, but they will have a small affect on the numerical value of f_c.

One can also make a general observation in the context of SHM. Consider the case when one is measuring the dynamic response of a structure. In the absence of any confounding influences, this would usually be stationary (assuming stationary excitation). If the structure suffers a damaging event, this usually has the effect of reducing the structural stiffness and thus the natural frequencies. In this case, the response would have marginally fewer periods of the dominant frequencies than for the undamaged structure measured over the same time window. With the interpretation of the ADF statistic in this section, it would follow that the ADF statistic of the response would increase after damage, although this is likely to be a small effect unless there are major changes in the frequencies i.e. large damage. The use of the ADF statistic as a damage sensitive feature has actually already been proposed [78].

One of the more interesting outcomes of the analysis above, has been the determination of a critical frequency for a signal f_c, associated with the critical value of the ADF statistic. It follows that one might be able to separate a given signal into stationary and nonstationary components by separating the frequency components above and below f_c.

6.3.4 Example on SHM signal

From all the aforementioned analysis performed and the discussion made on the interpretation of ADF t-statistic, it was clear that ADF essentially counts the number of cycles of the dominant periodic component in order to make the judgement if the signal is stationary or not. In addition, it was shown that a critical frequency f_c in Hz can be found for each signal, if the sampling frequency f_s is given for a signal, which can be derived as function of the number of observations. To demonstrate the latter for a real-time bridge signal, the first natural frequency of Tamar Bridge will be used.

For the purpose of this analysis, the measurements of one day of the Tamar bridge natural frequency are used. The signal is shown in Figure 6.9, where x-axis includes the hours of monitoring, and describes one cycle of the signal. The corresponding dominant frequency can be found, using the Discrete Fast Fourier Transform (DFFT) [123] and here is 1 Hz.

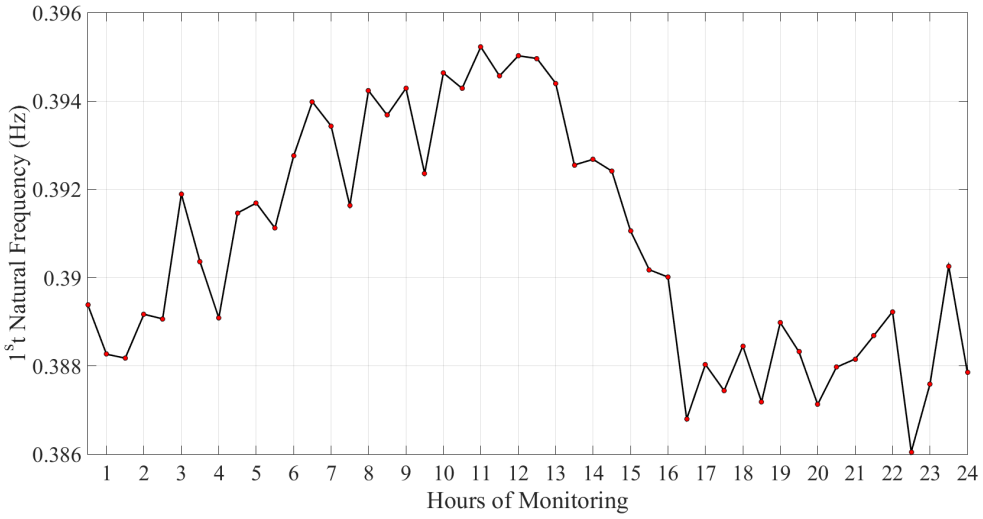


Figure 6.9: One day signal of the fundamental natural frequency of Tamar Bridge representing 48 half-hourly measures.

The assumption made here is that the sampling frequency includes 48 observations, as each observation on Tamar Bridge represents an half-hour measurement and consequently 48 observations represent one day of monitoring. In other words, the normalised frequency of the signal is 1/48. The question that this analysis will attempt to answer is; after how many hours the signal is expected to become stationary and what is the normalised frequency at that point. Knowing that, the

critical frequency f_c can be obtained.

In particular, using the ADF test on the signal shown in Figure 6.9, a t-statistic of -0.943 is obtained, which is higher than the critical t-statistic. The t_c here is -2.6135 and represents the 1% significance level. Now, increasing the cycles of the signal by adding successively cycles, the t_c is reached around five cycles. In particular, Figure 6.10. In particular, the critical number of cycles (C_c) is 4.6 cycles. Then, the t_{ADF} can be plotted with respect to normalised frequency, as shown in Figure 6.11. From this, the critical normalised frequency f_c is 0.0948, employing linear interpolation.

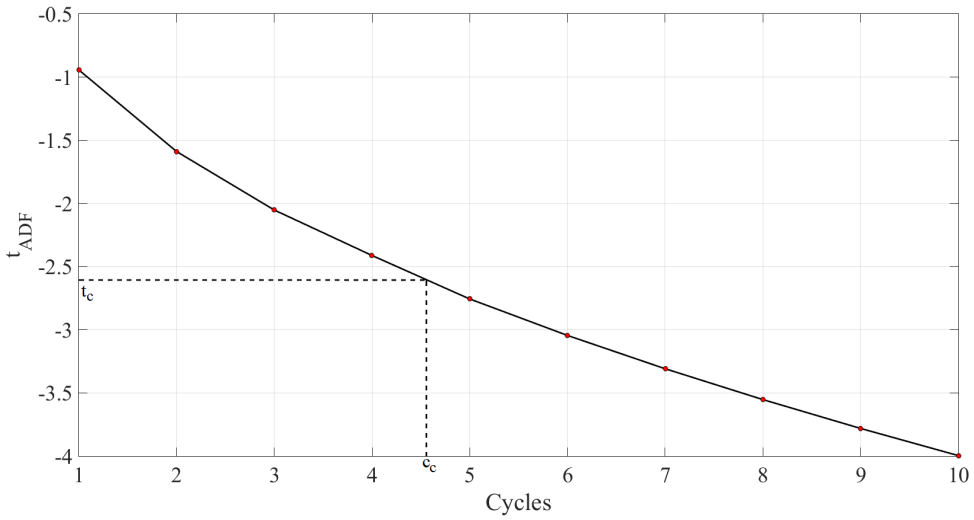


Figure 6.10: Calculated t-statistic versus the number of cycles.

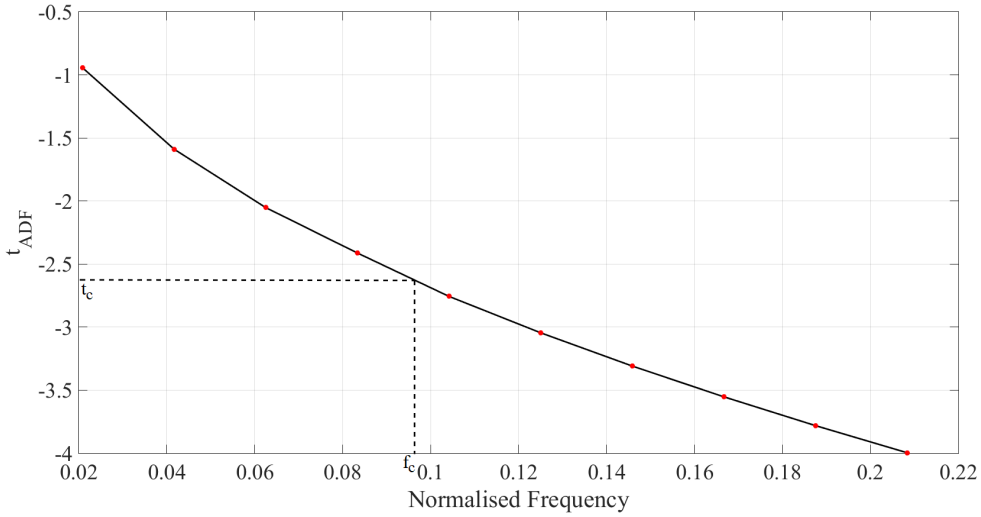


Figure 6.11: Calculated t-statistic versus normalised frequency.

Furthermore, the number of observations after which the signal becomes stationary

can be found based on critical normalised frequency, which is $N = 220$. The latter, in this case, is also the critical frequency F_c, because $N \times \Delta t/T$, where Δt is 1 and also the dominant period (T) is unity, as dominant frequency (f) is 1. Therefore, F_c is 220 Hz. In terms of hours monitoring, after 110 hours the signal will become stationary.

6.4 Conclusions

For various reasons, relating to applications in engineering like structural health monitoring (SHM) and system identification, this Chapter speculates on the nature of stationarity and how it can be quantified. It is briefly argued that mathematical definitions are of little practical use on finite samples of data and that a subjective assessment by an expert may be as effective. A hypothesis is presented that experts would base their assessment on how many cycles of the dominant periodic component are present in the sample. If this hypothesis holds true, then any statistics used for nonstationary assessment would need to support it. This Chapter shows, using dimensional analysis, that the ADF unit root test, which one of the commonly employed nonstationarity tests inside econometrics, follows the latter hypothesis. It should be a simple matter to apply similar dimensional arguments to other measures, such as PP and KPSS. The results of this Chapter should be of interest for engineers, econometricians and anyone else concerned with time series analysis where stationarity assessment is of great interest.

After proving this interpretation about ADF statistic and how a signal can be judged as nonstationary or not, it is important to refer to different way that can be used to monitor the cointegration residual. Although, cointegration is a damage sensitive features, there are tools, such as statistical process controls that can be used to monitor the stationary cointegration residual. These will be presented in the next Chapter.

Chapter 7

Cointegration Residuals and Novelty Detection

7.1 Introduction

In the previous Chapters the *Johansen's approach to cointegration* was introduced, discussed and demonstrated based on three bridge case studies (Tamar, Z24 and the aluminium truss bridge). In addition, a simple interpretation of the ADF test for SHM series nonstationarity assessment was presented, based on basic signal processing concepts and dimensional analysis. From the aforementioned, it was observed that SHM signals from structures in operation are generally nonstationary and need to be rendered stationary in order to monitor optimally them for novelty/damage detection. In particular, SHM signals are nonstationary due to benign reasons contributed to the impact of EOVs. Cointegration can be used successfully to project out this benign nonstationarity where variables are linearly related. The main feature used for damage detection when employing linear cointegration is the cointegration residual.

According to *Axiom II* for SHM [7], damage detection is based on the comparison between two structural response states, an undamaged and potentially damaged one. Hence, to establish a linear cointegration relationship requires an initial training set taken from the undamaged state. Based on [7], the training set should be long enough to obtain a stationary residual, but up to a point, where the feature is still

damage sensitive. An example of such a case is given in Chapter 4 on the natural frequency data of $Z24$ bridge.

As the cointegration residual is a stationary time series, different methods can be used to monitor how the residual changes over time. A common practice to track the evolution of a stationary time series over time is by the use of statistical control (SPC) charts [7].These have been used in the past in the context of SHM, as well as in combination with Johansen's approach to cointegration [2, 7, 105, 124, 125] .

This Chapter specifically studies the available means of monitoring the cointegrated residual in order to explore their suitability and success for damage detection. Commonly only a standard control chart is used to monitor the cointegrated residuals in SHM. The Chapter will provide a review of the general process control methodologies employed inside SHM literature, which are combined with cointegration residual used for novelty detection purposes. In order to discuss the aforementioned methods and compare their results, a simple mass spring system is used to generate natural frequency series both for undamaged and damage conditions. In addition, the experimental data introduced in Chapter 3 will also be used to explore the available possibilities. Finally, conclusions will be made as to which options are most suited for residual monitoring.

7.2 Data Generation: Mass-Spring System

A simple mass-spring system is developed and simulated in Matlab in order to generate series of natural frequency data. The system is illustrated in Figure 7.1 and consists of six masses and seven springs. Solving the eigenvalue problem, the six natural frequencies can be obtained. In order to simulate the variation of natural frequencies over time, the stiffness of each spring is considered as a function of temperature ($k_i = 4 - 0.05T$ for $i = 1, \ldots, 7$). This equation used to simulate the stiffness of each spring was selected because it was tested and provided nonstationary results representing satisfactorily the expected variability of stiffness. The variability of spring stiffness is provided using the air temperature field taken from the monitoring campaign of the Tamar Bridge (Figure 7.2). Of course one can amend this and generate more or less stationary variables. It should be also said that a linear behaviour between temperature and stiffness assumed here, as the simplest one that can exist, which can be also used to explain more easily the future results.

In total, six frequency time-histories have been generated with 16000 observations each (Figure 7.3). At 10500 data observation, a decrease in stiffness K_4 was introduced to simulate damage. Different damage percentages are introduced in order to compare the methods.

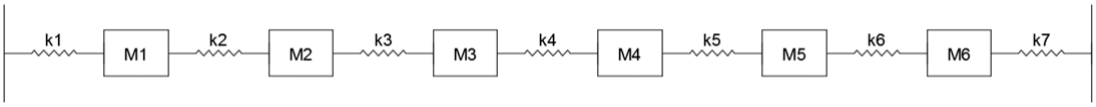


Figure 7.1: Mass-spring-stiffness system.

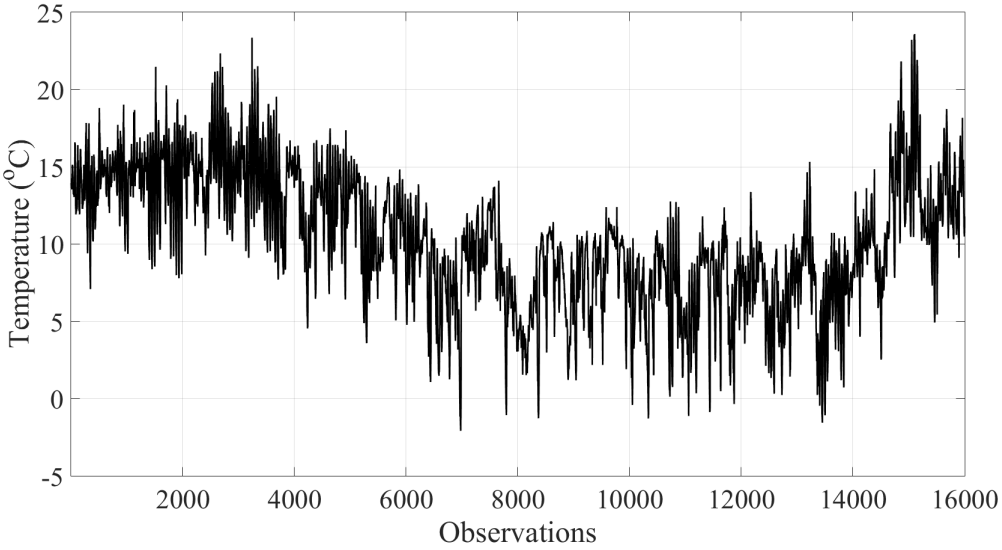


Figure 7.2: Temperature variation of Tamar Bridge for 8 months of monitoring (01/07/2007 to 31/03/2008).

7.3 Monitoring the Cointegration Residual

The aim of the cointegration methodology within SHM is to find the most stationary combination of the variables of interest. The idea is then to infer a structural change from a variation in this combination (residual), i.e. the cointegration residual can be used to draw inference about a system's underlying physics and how this is changing over time. The available methods for monitoring the cointegrated residual can be sorted into two main categories: (i) outlier analysis and (ii) nonstationarity assessment. In the former, measures of data distance are used to identify data population exhibiting different characteristics from normal ones. The most commonly

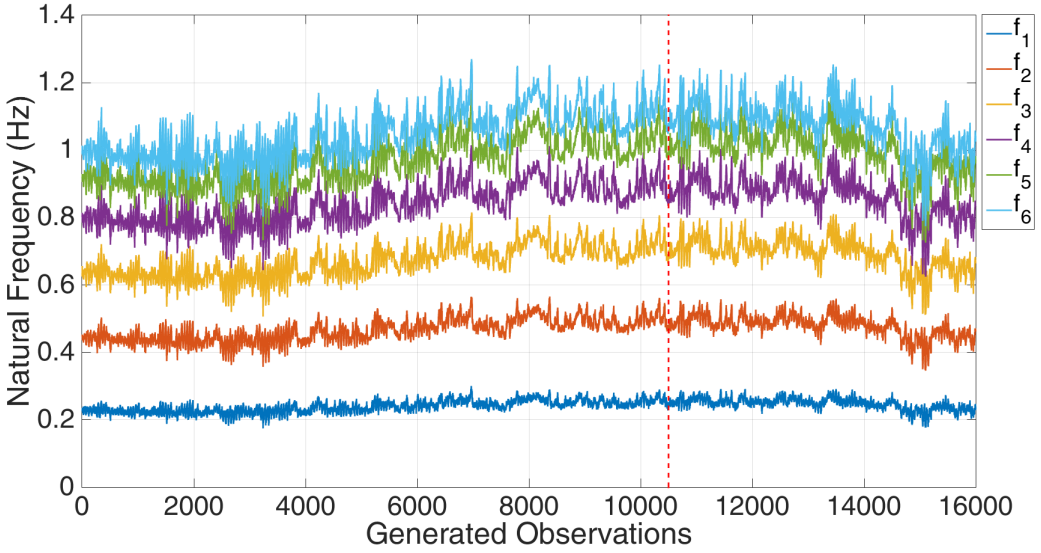


Figure 7.3: Natural frequency series generated using the mass-spring system and the temperature recording of Tamar Bridge for a 5% stiffness reduction introduced in spring four (K_4).

employed measures of distance are those of *Mahalonobis* and *Euclidean distance* [7]. On the other hand, for non-stationarity assessment, the moments of residual, such as the mean, the variance or the nonstationarity of residual using *unit root tests*, are tracked over time. The most commonly employed measures of the second category are the *X-bar* (mean) and *S-control* (variance) charts.

As the methods mentioned above all track the change in some way, the question of which to use may seem trivial. However, as this Chapter will demonstrate, a number of different behaviours in cointegrated residuals have been observed and these patterns may lend themselves to one monitoring method over another. The different behaviour observed stems from the nature of the approach taken to cointegration for SHM; one is searching for the *most* stationary combination of variables, however stationarity is not guaranteed, the residual, even within the training set, may not be *strictly stationary*, but still of use for damage detection [126].

For the purposes of this Chapter, the X-bar and S-bar chart methods for monitoring the cointegration residual will be used and discussed, employing the synthetic natural frequency data generated from the previously introduced linear mass-spring system (section 7.2) and the laboratory data discussed in Chapter 3. Different methods to monitor the cointegration residual are compared and discussed. Finally, the use of the ADF t-statistic for cointegration residual monitoring will be discussed

and compared with the previous two control charts.

7.3.1 Tracking Methodologies

Control Chart ($\mu \pm 2.58\sigma$)
An important issue faced when building a novelty detector is how the appropriate thresholds (alarm levels) can be specified. In particular, [7] provides evidence that novelty detection threshold can be based on extreme value statistics (EVS). The upper and lower threshold approximated using Gumbel and Weibull empirical and cumulative distribution function (CDF) for maxima and minima. From this it was concluded that at least 99.5% of the data should be enclosed by the upper and lower thresholds, which in terms of data in a normal distribution distribution normality can be expressed as a number of standard deviations (σ) from the mean value (μ). More specifically, 99.54% corresponds to $\mu \pm 2\sigma$, while 99.73% corresponds to $\mu \pm 3\sigma$. The threshold adopted for the analysis here is $\mu \pm 2.58\sigma$, which represents the 99% confidence interval.

X-bar Chart
X-bar is a basic statistical process control method focusing on the idea of suppressing noise by averaging the process and of making the summed variable more Gaussian (central limit theorem) [110],[7]. The signal is regarded as a Gaussian noise sequence and one can set alarm level or control limits at $\mu \pm 2.58\sigma$ for training set data. In abstract terms, X-bar tracks the mean of the signal, which is the first moment of stationarity of signal and observes how this change over time. As mentioned in the introduction, for use with cointegration this method requires the signal to be stationary, at least inside the undamaged state, in order to be able to detect the presence of novelty or in other words signal nonstationarity, which can effect the mean of the signal over time. In particular, the X-bar chart calculation used in [124] will be followed here in order to monitor the cointegration residual (stationary signal). In that case, the X-bar chart was used for the assessment of tilts' series provided by the monitoring of a footbridge located at UK's National Physical Laboratories (NPL).

S-bar Chart
In contrast to X-bar, S-bar control charts assess the variance of the signal (${\sigma_y}^2$). As in the case of X-bar, S-bar chart also looks into a sequence of data moving

windows using a predetermined window's width. For the S-bar chart, the upper (UCL) and lower control limits (LCL) are obtained, according to [7], as described in equations (7.1) and (7.2), where $X_{(p,n)}$ describes the p^{th} quantile of chi-squared random variable, n is number of observations of each window. $\bar{S}$ is the mean value of variance ($\sigma_y{}^2$) of all observations corresponding to the normal state, whilst α represents the significant level (i.e. 0.01 for 1%).

$$UCL = \bar{S}\sqrt{\frac{X_{1-a/2,n-1}}{n-1}} \tag{7.1}$$

$$LCL = \bar{S}\sqrt{\frac{X_{a/2,n-1}}{n-1}} \tag{7.2}$$

The previously tested statistical process control charts (X-bar and S-bar), are the most commonly employed inside SHM. However, further control charts, such as cumulative sum (CUSUM), exponential weighted-moving average (EWMA), Hotelling or Shewhart T^2, can be used in the same manner. A review of these aforementioned control charts can be found in [7]. It should be mentioned here that for damage detection purposes inside SHM, the cointegration theory has also been employed in combination with multivariate EWMA (MEWMA) in [127].

7.3.2 The Suitability of Different Tracking Methodologies

Figures 7.4 and 7.5 show the cointegration residuals of the simulated data with 5% and 20% damage introduced respectively. In both of these Figures, the statistical threshold was set at $\mu \pm 2.58\sigma$, where μ is the mean of the normal/training state, and corresponds to 99% confidence intervals.

From these Figures one can see that the residuals appear stationary in the training period, in fact, they appear to be white noise sequences. In this case, an outlier analysis is an obvious candidate, as new data can be compared to the stationary distribution of the data in the training set. As the cointegration residual is a single series, an univariate outlier analysis is selected here; the Euclidean distance. This is a discordance measure, where the potential outlier is compared against some objective threshold to evaluate the statistical likelihood of coming from an specific distribution.

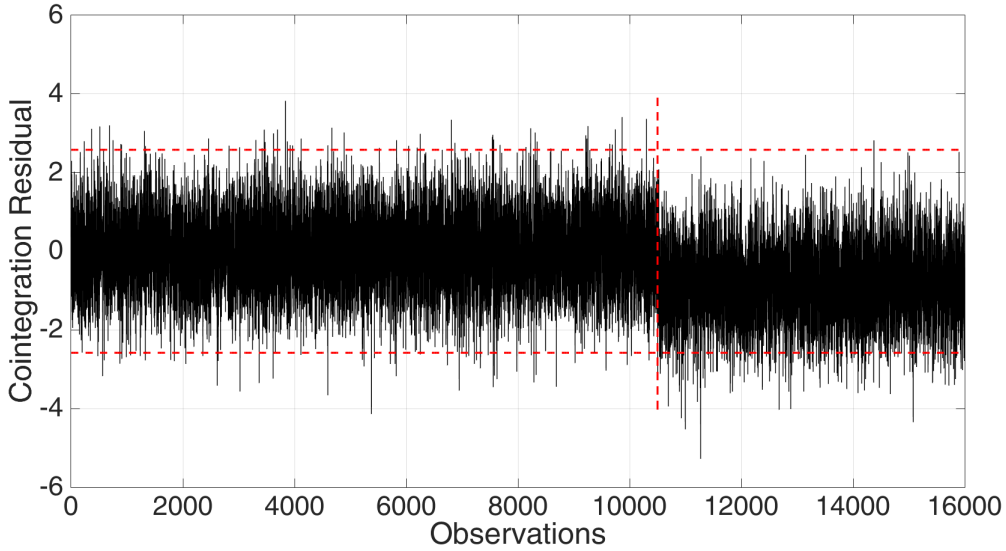


Figure 7.4: Cointegration residual for 5% damage introduced in spring stiffness K_4. The vertical red dashed line corresponds to the point where the damaged introduced. The threshold was set at $\mu \pm 2.58\sigma$.

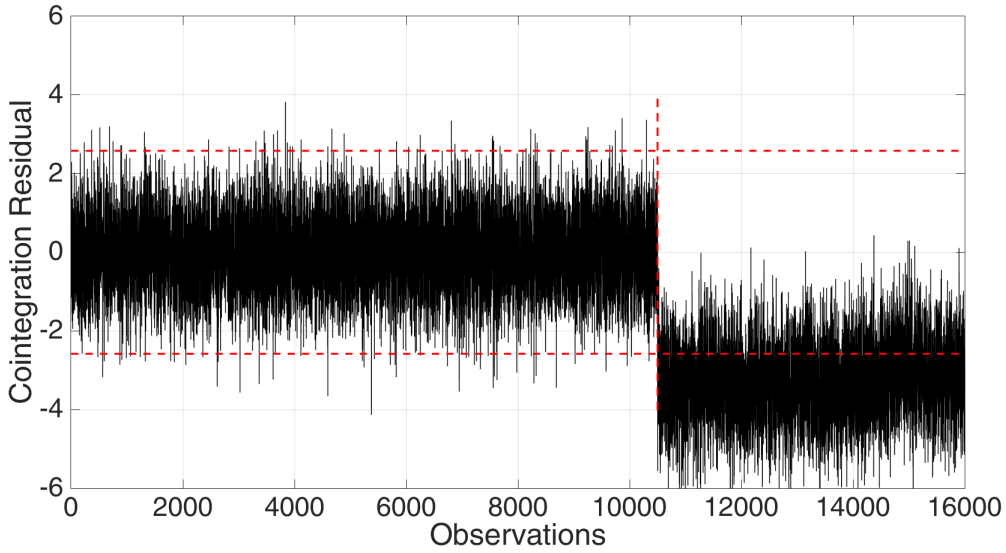


Figure 7.5: Cointegration residual for 20% damage introduced in spring stiffness K_4. he threshold was set at $\mu \pm 2.58\sigma$.

The deviation statistics used here is given in Equation (7.3), where x_ζ is the potential outlier, $\bar{x}$ is the mean of the training set of the residual, s is the standard deviation of the training set, whilst z_ζ is the novelty index. The thresholds of the outlier analysis depends on both the number of points of the signal and the number of dimensions of the problem being studied [7]. As here the univariate analysis is employed, one

can assume a Gaussian normal distribution and a confidence interval of 99%, which corresponds to $\mu \pm 2.58\sigma$.

$$z_\zeta = \left| \frac{x_\zeta - \bar{x}}{s} \right| \tag{7.3}$$

As can be seen clearly from Figures 7.5 and 7.6, an outlier analysis can easily detect the mean shift that results from a change in stiffness. It is interesting to point that inside the normal condition of the structural response (1-10500), outliers can be detected, which can be assumed as lase positives, as there are no inference of damage and probably inform the existence of local changes of temperature that led to increase of natural frequencies.

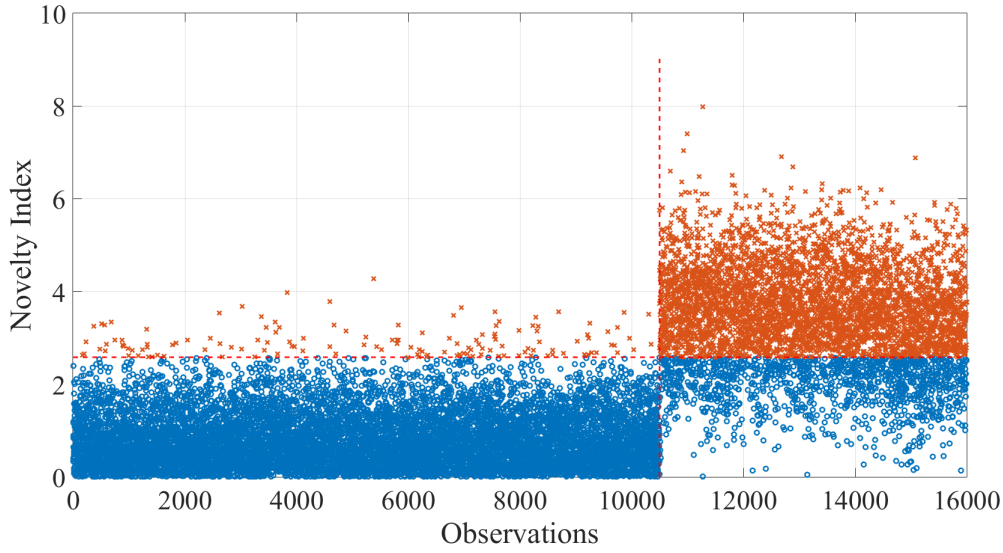


Figure 7.6: Outlier analysis for 20% damage introduced in spring K_4. The blue circled points described the inliers, while the orange crosses the outliers. The threshold was set at $\mu \pm 2.58\sigma$.

However, some potential damage behaviour could be undetectable with this kind of analysis. For example, if damage introduced a repeating trend with a constant mean and an amplitude smaller than the error bars\threshold, this would go undetected, a reduction in variance of the residual would also go undetected. Such an example is shown in Figure 7.7. In these cases, the alternative methods discussed in Section 6.3.1 may be of more use. For this simple simulated case, the results with these alternative methodologies are shown below.

More specifically, the X-bar chart is demonstrated here on the cointegrated signal

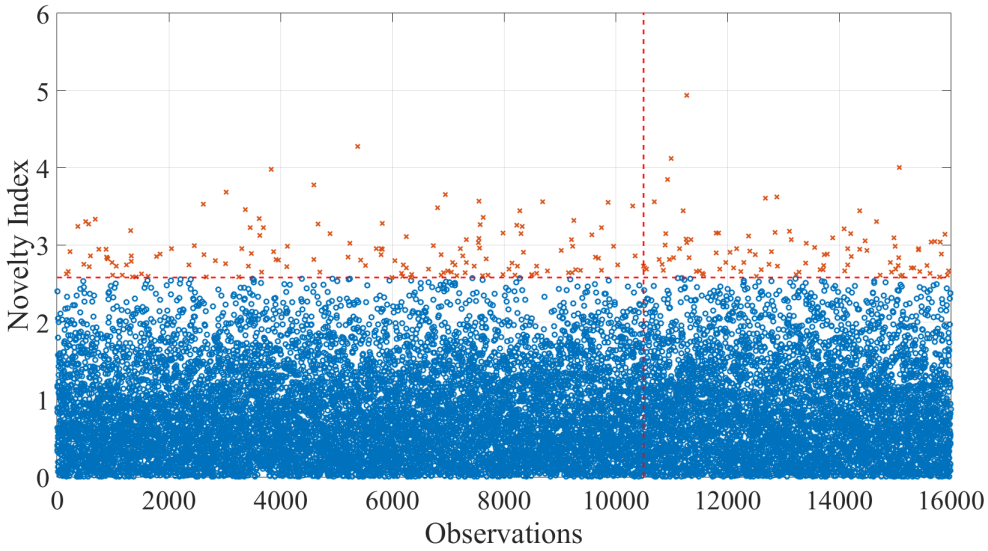


Figure 7.7: Outlier analysis for 2% damage introduced in spring K_4. The blue circled points described the inliers, while the orange crosses the outliers. The threshold was set at $\mu \pm 2.58\sigma$.

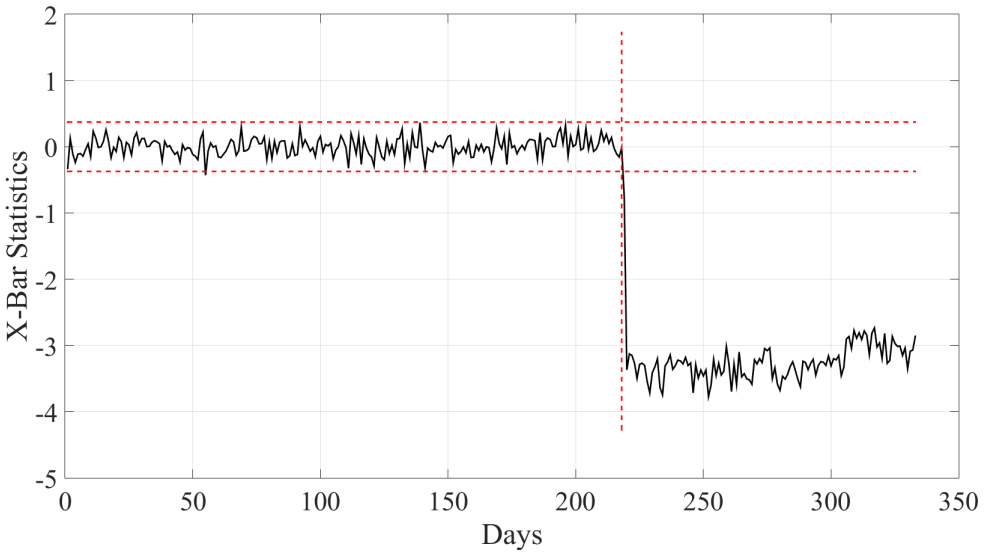


Figure 7.8: X-bar Chart for a 10% damage introduced in K_4. The horizontal red dashed lines represent the thresholds at 99% confidence interval.

of the mass-spring system. As shown in Figure 7.8, X-bar tracks the mean of the signal, considering daily data (48 half-hour observations) and the thresholds are set at $\mu \pm 2.58\sigma$ from the original signal's training data (first 2850 observations). The total data are 15984 and corresponds to 333 subgroups of 48 observations each (one day each). As one can observe from Figure 7.8, at 219^{th} subgroup/day, the

mean value of the signal changes significantly, something that informs the presence of nonstationarity in the original natural frequency signals. The response of the structure after the 219^{th} day continues to vary, however around the increased mean value.

Furthermore, the results obtained by the S-bar chart analysis on mass-spring cointegration residual are depicted in Figure 7.9.

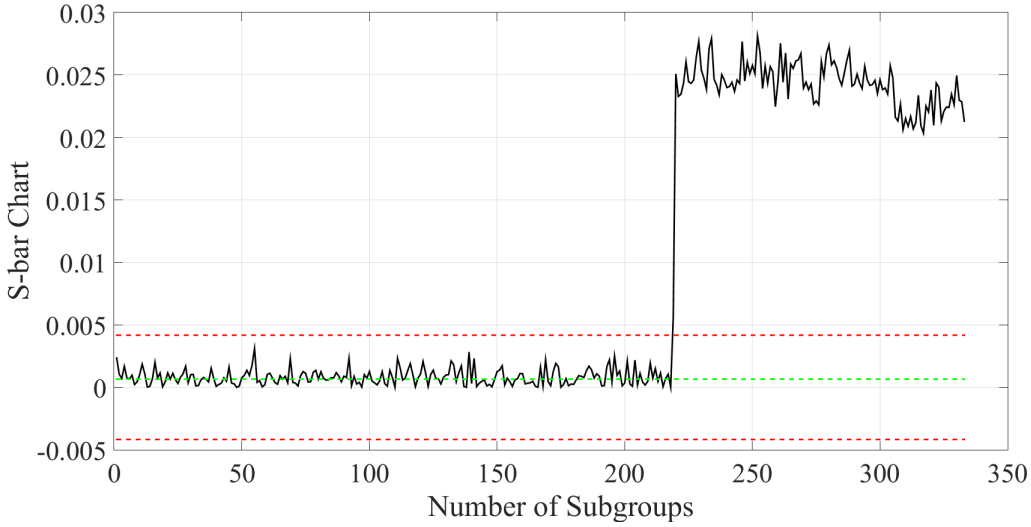


Figure 7.9: S-chart for 20% damage introduced in K_4 spring. The horizontal red dashed lines describes the thresholds at 99% confidence intervals, while the horizontal green dashed line is the standard deviation of the standard deviations calculated.

This shows that the introduction of a significant amount of damage in the cointegration signal leads to introduction of nonstationarity in the cointegration residual and consequently changes the signal's variance.

As mentioned previously, the cointegrated residual of the data provided by the simple mass-spring system was perfectly stationary, because the natural frequency signals share exact common trends. Now, to consider a more realistic situation the case of laboratory truss natural frequencies will be examined employing the X-bar and S-bar charts. In particular, Figure 7.10 shows the cointegrated residual of the experimental data introduced in Chapter 3. Here one can see a more typical example of residual behaviour, where some structure remains, and even in the training set, one could not say that the residual was strictly stationary. Nevertheless, the residual clearly shows sensitivity to damage and is, therefore, useful for SHM. The key question of interest here is whether nonstationarity tests might be more suited for tracking

this kind of imperfect residual, or whether it is more suitable to consider a range of control charts.

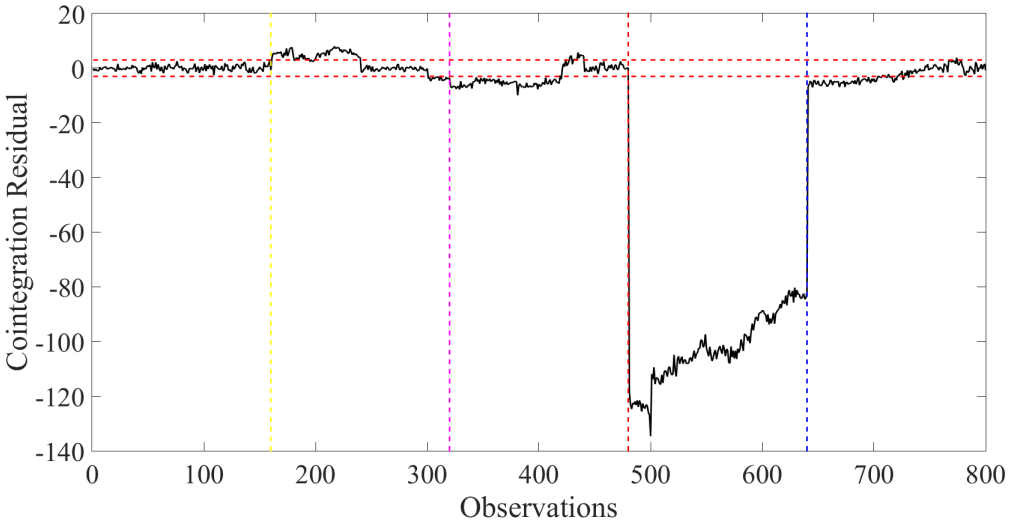


Figure 7.10: Cointegration residual for laboratory truss bridge data. The yellow dashed line distinguish the normal state from that of the first damage, the magenta one the first from the second damaged state, the red the second from the sever third one and finally the blue dashed line distinguish the third damage from the restored state.

To begin with, in Figure 7.11 the X-bar chart of truss bridge cointegration residual is illustrated. For this chart, the mean is considered for every 20 data observations. The latter represents every temperature step, which includes 20 recordings of natural frequency. The zoomed-in version of the Figure is also provided in Figure 7.12. From both figures, it can be observed that the damage is successfully detected for all cases.

Next, the S-bar is employed in order to track down the variance of the signal over time. This is illustrated in Figure 7.13. Here, the variance is considered for every 20 observations. An interesting observation is that for the first two damage cases, the variance did not change much in contrast to the third damaged case, where variance changes significantly. In addition, the variance seems to be relatively stable during the restored state, where observations are close to the normal state.

In other words, from the analysis performed on the cointegration residual of the laboratory truss bridge experiment employing X-bar and S-bar charts, it was observed that X-bar chart is more sensitive to detect the damage than the S-bar chart. Therefore, it can be said that the removal of bolts affected the mean value of the residual more than its variance. In the next section the use of Augmented Dickey-

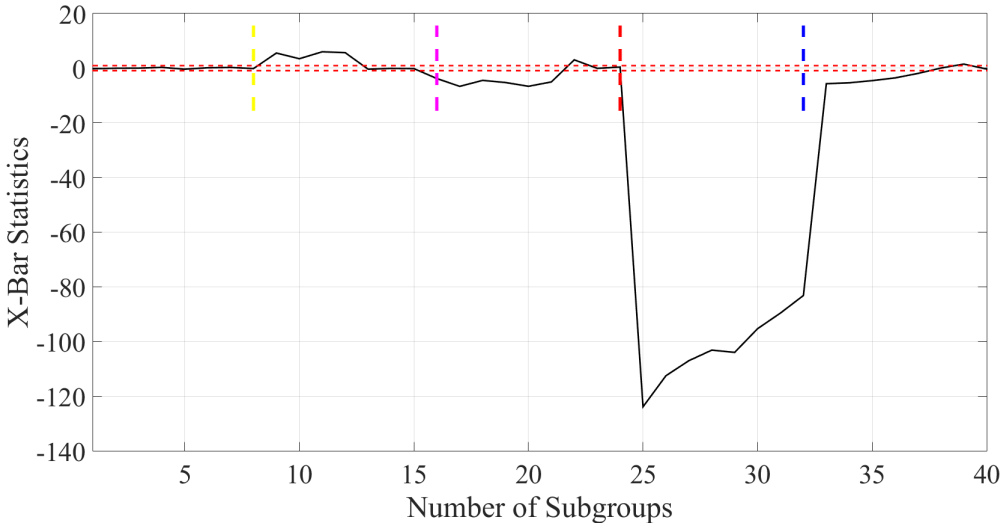


Figure 7.11: X-bar chart considering the mean for every 20 observations.

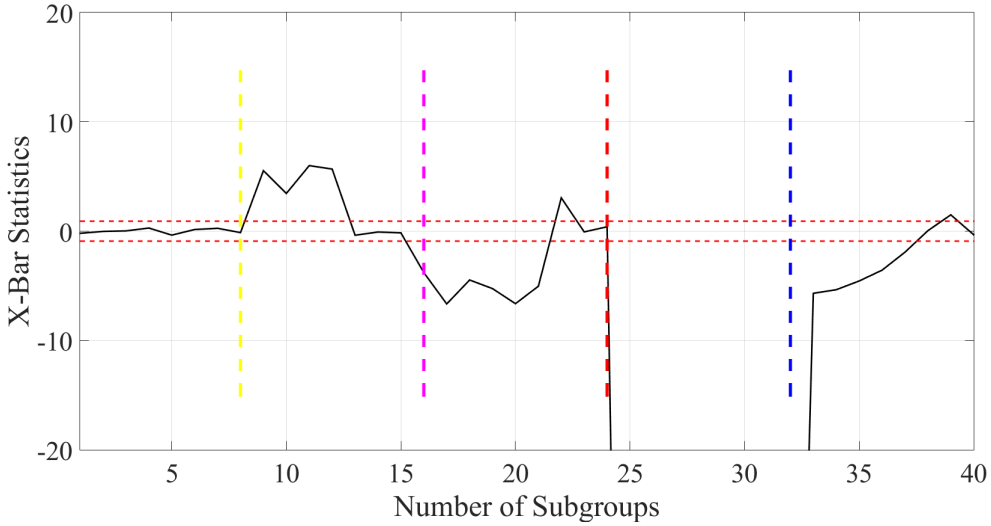


Figure 7.12: Zoomed in version of X-bar chart considering the mean for every 20 observations.

Fuller (ADF) statistic as a novelty measure is demonstrated and discussed in both the synthetic mass-spring system data and the natural frequencies of the laboratory test.

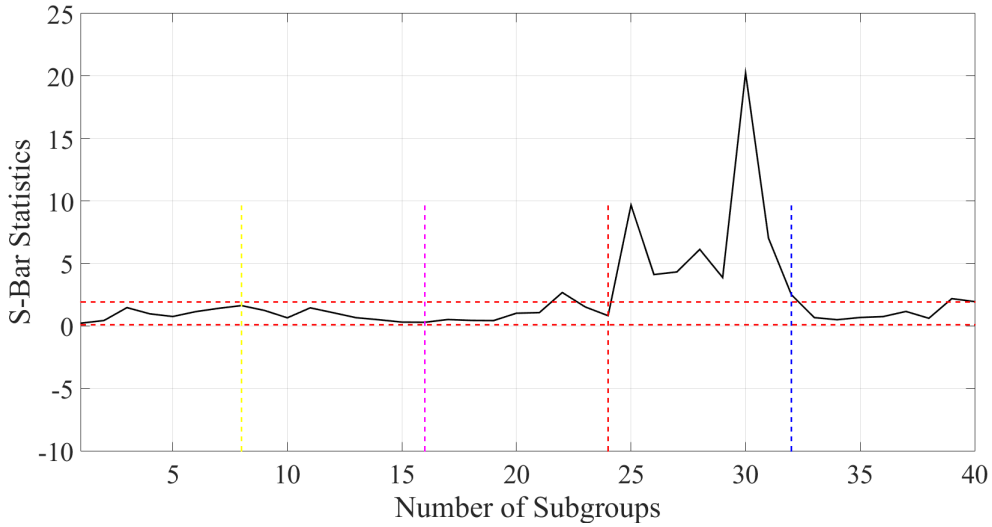


Figure 7.13: S-bar chart considering the mean for every 20 observations, 40 subgroups.

7.3.3 Non-stationarity Testing

Besides control charts, nonstationarity proxies can be used to identify the presence of nonstationarity in the cointegration residual. In the field of Econometrics the most commonly employed nonstationarity test is the ADF unit root test, which was described in detail in Chapters 5 and 6. Some other alternatives for the ADF are Phillips-Perron (PP) [96] and the Kwiatkowski et al. (KPSS) [97] tests. In general sense, the results provided from all three aforementioned tests are quite similar in terms of nonstationarity indication. From these, only the ADF have been used in the past inside SHM literature for monitoring the cointegration residual. In particular, the use of ADF t-statistic for monitoring the cointegration residual and damage severity assessment was proposed in [78, 128]. More specifically, cointegration was used for EOVs impact elimination on the data provided by a lamb-wave test performed in an rectangular (200mm $\times$ 150mm) aluminium plate of 2mm thickness, where a small hole was introduced to simulate damage. The ADF t-statistic was used to assess both the pre-cointegrated and cointegrated data, showing that the initial nonstationarity of the signals is eliminated using cointegration and nonstationarity introduced when the plate is damaged.

The ADF t-statistic method for cointegration residual monitoring is demonstrated here, on the mass-spring generated natural frequencies and natural frequencies of laboratory truss bridge. To do so, a sequence of moving windows is chosen with

a window width of 200 observations each. For each window the ADF t-statistic is calculated and plotted. The thresholds, upper and lower, are calculated at $\mu \pm 2.58\sigma$, from the normal state data (first 2850 observations). The results of this analysis are depicted in Figure 7.14.

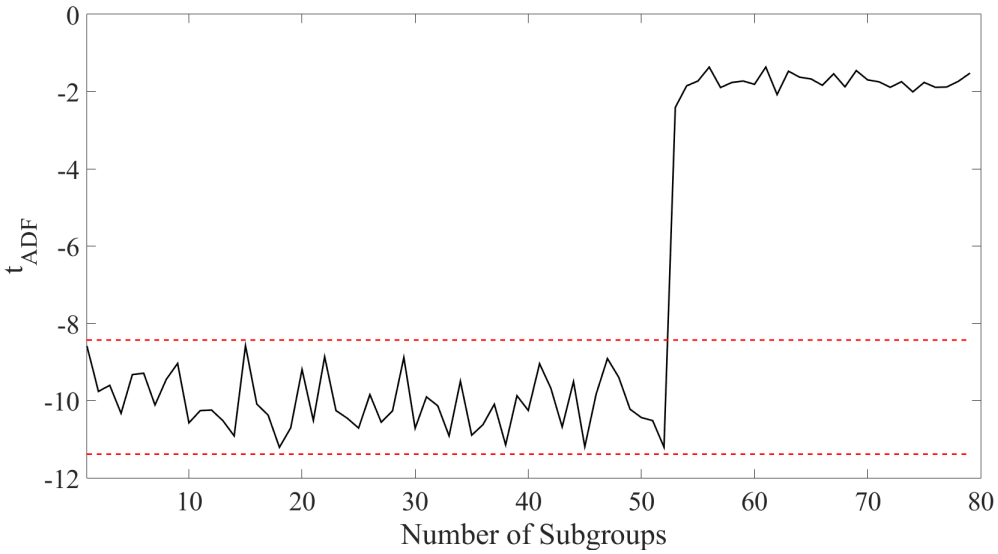


Figure 7.14: ADF t-statistic for the mass-spring system residual.

In particular, Figure 7.14 shows how the ADF t-statistic changes over the undamaged and damaged windows, for a reduction of K_4 spring's stiffness of 10%. Inside the first 51 windows the ADF t-statistic stays inside the calculated thresholds. However, when the damage is introduced, the t-statistic increases. This is because nonstationarity introduced to the signals and consequentially to the cointegration residual. From this example, one can say that the ADF provides similar results to those provided by X-bar and S-bar charts in the previous section. However, there are some issues related to ADF t-statistic that should be discussed.

More specifically, to illustrate and discuss them, the cointegration residual provided from the laboratory truss bridge experiment should be used. The ADF t-statistic employed for windows including 20 observations each of the cointegration residual and the results are depicted in Figure 7.15. From this plot it can be observed that t-statistic provides inference about the different states of damage in the system, however one can understand that t-statistic is a weak measure of damage. In particular, from the use of ADF t-statistic as a novelty measure, by testing different damage cases and different windows' widths, the observations and discussion are provided below:

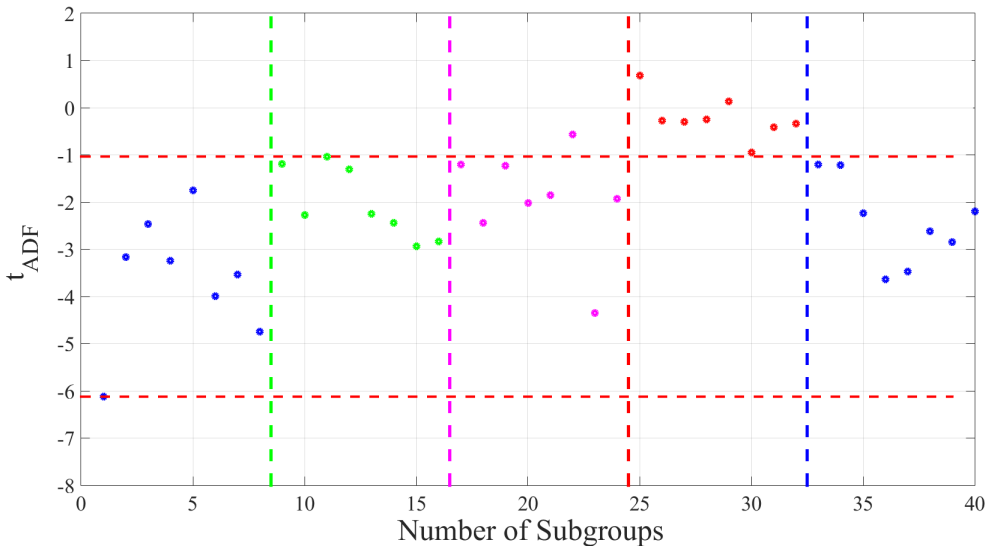


Figure 7.15: ADF t-statistic for the cointegration residual of the laboratory truss bridge natural frequencies. The points in blue describe the normal and restored condition,in green the first damaged state, in magenta the second one and in red the most severe third one.

More specifically, three main points/drawbacks that have been observed employing ADF t-statistic as novelty measure are:

- In cases of rate effects, such as temperature variation in small-case structures, as in the laboratory experiment, high variability can be observed on t-statistic. In particular, the use of t-statistic for a sequence of moving windows with small width, showed that high variability can be observed both during undamaged and damaged condition, leading to outliers inside the normal state.

- The results of ADF t-statistic shows that there is nothing specific happening in the damage case 1 (green), while also a little information in terms of damage severity can be extracted for the observations of the second damaged state (magenta).

- In the damage case 3 (red), which is the most severe damage case of all, outliers are observed, but the amount with which t-statistic increased in not very informative. As a matter of fact, from further testing using different damage amounts, it was found that ADF t-statistic is not increasing linearly with the cointegration residual estimation, especially for positive t-statistic values.

- Finally, its is important to mention that the damage sensitivity of the ADF t-statistic depends a lot on the procedure of lag-length specification, which depends on the number of observations used for each window. The latter effects can be very problematic and sometimes highly time-consuming to deal with.

All the above points, show a specific drawback that ADF t-statistic has as a novelty measure. This is the fact that the value of t-statistic represents the mean value between the components of a signal. Assuming the simplest decomposition, a signal can have a high-frequency and a low-frequency component. The former is associated with the noisy component of the signal and the latter with the more nonstationary part. Therefore, in the presence of low levels of damage, the nonstationarity will be introduced in the signal, however its effect will not be captured in full by ADF t-statistic because the high-frequency part is still inside the cointegration signal and probably dominates its variation. In other words, ADF t-statistic is damage sensitive measure, but seems to be a *weak* one. This observation fits with interpretation provided in Chapter 5 for the ADF t-statistic.

7.4 Conclusions

In this Chapter, different ways used in the past in SHM literature to monitor damage sensitive features are demonstrated for cointegration residual and discussed. Firstly, the residual is assessed to identify the possible outliers overcoming statistically determined thresholds (control chart). From this analysis it was shown that the number of outliers increases linearly with the effect of damage introduced. Then, the cointegration residual was averaged, over a sequence of moving windows, employing the X-bar statistical process control (SPC), using daily data, which showed that the mean changes significantly in the presence of novelty/damage. Next, the variance of the cointegration signal was tracked over a sequence of moving windows, which also observed to be damage sensitive. Finally, the use of non-stationarity tests, such as that of ADF t-statistic is demonstrated and discussed. Although the t-statistic showed that it can be used in some cases in order to monitor the cointegration residual, its sensitivity is in question. This is because ADF t-statistic represents the mean of the high and low-frequency components of the signal and therefore if the effect of non-stationarity is not significant enough to dominate the signal, the ADF

t-statistic is affected significantly by the noisy part exhibiting low damage sensitivity capabilities.

All in all, from the aforementioned analysis it is clear that the use of SPC charts requires stationarity of the signal under consideration. This can be achieved using the Johansen cointegration, projecting out the confounding impacts of EOVs. The SPC charts has the capability to identify the existence of non-stationarity on the cointegration signal and successfully monitor residual's progression over time. This non-stationarity is the consequence of novelty existence i.e. damage or the effect of rapid varying temperature effects (small-scale structures/experiments).

Chapter 8

A Principled Multiresolution Approach for Signal Decomposition

8.1 Introduction

In Chapter 6, a simple interpretation of the ADF statistic is proposed based on dimensional analysis. A byproduct of the analysis was to show that there existed a critical frequency f_c for given sample parameters, such that any signal components containing only frequencies below f_c would be judged nonstationary by the ADF test. This result is used here to motivate the definition of a simple, but principled, decomposition method for signals, which can resolve them into their stationary and nonstationary components. This decomposition is based on the multiresolution analysis, which facilitates the orthogonal wavelet analysis.

The layout of this chapter is as follows; firstly, a brief description of the analysis methods including the MRA, the ADF test and autocorrelation functions is given. The second and third sections provide an illustration of the method on the Z24 bridge data. Finally, the conclusions are presented.

8.2 A Signal Decomposition based on Wavelet Analysis and the ADF Test

Based on the analysis presented in [129], it can be assumed that one can decompose a signal into its stationary and nonstationary components by separating out the components with frequency content above and below a critical frequency f_c, established using the ADF test criterion. However, in the previous analysis the existence of the critical frequency was only demonstrated for the case of a harmonic signal corrupted by noise; in the general case, the critical frequency would require a more complicated analysis. It will be shown in this section that the decomposition idea can be implemented in the context of wavelet analysis, without an *a priori* calculation of f_c. This will prove important in the context of SHM.

The issue in SHM, is that there will generally be nonstationary (periodic) components in time series at more than one frequency. If one considers the example the Z24 Bridge, Figure 8.1 shows the fluctuations in natural frequency associated with seasonal variations in temperature over a period of one year. However, other environmental and operational variations (EOVs) generate nonstationarity at other time scales e.g. day-night variations in temperature and traffic loading will create nonstationarity with a daily period; patterns of traffic loading will induce weekly variations as a result of the different loads during the working week. SHM signals will thus require the more general decomposition method alluded to earlier. As mentioned in the introduction, it is necessary to remove all the confounding influences caused by benign EOVs, so that any potential signatures of damage can be detected.

Previous work on confounding influences and cointegration has exploited multiresolution analysis (MRA) and wavelet analysis [130], as a means of decomposing time series into frequency bands so that the EOVs at different timescales were exposed [105]. In particular, a scale-by-scale analysis performed in [105] considered the effects of using cointegration at different time scales. It was shown that the most nonstationary time scales, associated with lower-frequency signal components, manifested greater damage sensitivity than the higher-frequency ones, due to the fact that the onset of damage leads to a stronger reassertion of nonstationarity than in the original signal and correspondingly to a greater effect on the cointegration residual.

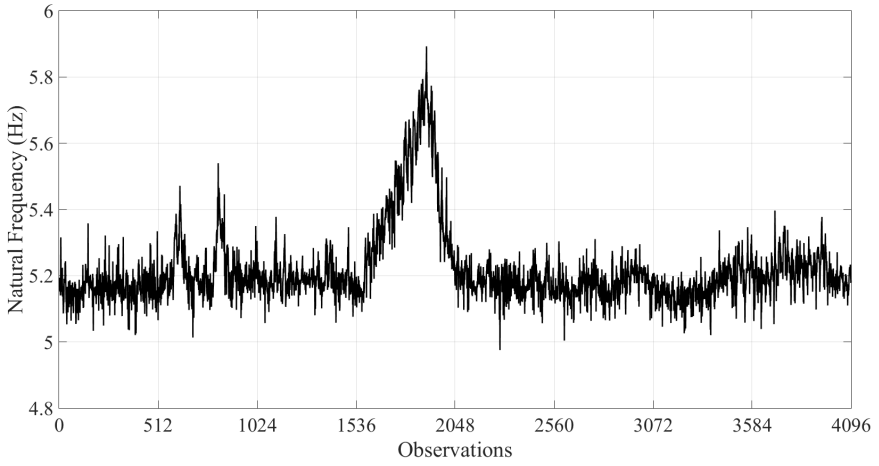


Figure 8.1: Second natural frequency of Z24 Bridge over a single year of monitoring.

The *level decomposition* used in [105], will be used here to motivate the required decomposition into stationary and nonstationary components. In fact, by making use of autocorrelation functions, it will prove quite simple to decompose signals – in a principled manner – into four main components: (i) a mean, (ii) a stationary part, (iii) a non-stationary part and (iv) a noise component. In essence, the main idea will be to obtain the MRA signal levels and then assess them using the ADF test to identify those that can be judged as nonstationary and those as stationary. In this way, a *critical level* can be found and used to decompose the original signal into a nonstationary and stationary part. This critical level (denoted L^*) is based on the ADF statistic associated with the 99% confidence interval, which corresponds to a critical t-statistic of -2.56. The second step in the method proposed is use *autocorrelation functions* (ACFs) in order to test the MRA levels and identify those that are *delta-correlated*, and can therefore be considered to be noise. In order to explain the decomposition, it will be necessary to summarise a little wavelet analysis.

8.2.1 The Orthogonal Wavelet Transform and MRA

As is now well known, the wavelet transform is a linear transform of the function $x(t)$,

$$x_k^m = \int_{-\infty}^{\infty} x(t)\psi_{m,k}(t)dt \tag{8.1}$$

into a translation (k) and scale (m) domain [130]. Unlike the Fourier transform, the expansion basis is in terms of functions localised in time, and this allows the wavelet transform to represent nonstationary signals, among its other benefits. If the set of scale and translation parameters are reduced to a discrete set, one obtains the *d*iscrete wavelet transform. A common choice for the discrete parameters is the dyadic framework in which, $a_j = 2^j$ and $b_{j,k} = k/2^j$. In this framework, the expansion basis functions are defined by,

$$\psi_{m,k}(t) = 2^{m/2}\psi(2^m t - k), \qquad (m, k \in Z) \tag{8.2}$$

where ϕ is termed a *mother wavelet*. The mother wavelet must satisfy certain properties e.g. it must be localised in the variable t, and this can be ensured by using functions ψ with compact support for example. The conditions on the mother wavelet allow many families of wavelet functions, but the most important families are those which have additional orthogonality properties,

$$< \psi_{m,k}, \psi_{n,l} >= \delta_{mn}\delta_{kl} \tag{8.3}$$

where $<,>$ is the standard inner product,

$$< h, g >= \int_{-\infty}^{\infty} h^*(t)g(t)dt \tag{8.4}$$

where the $*$ denotes complex conjugation. The term δ_{mn} represents the *Kronecker Delta function*, which is equal to unity when $m = n$, while zero for $m \neq n$.

With the extra condition on the wavelet functions, the transform in equation (8.1) becomes the *orthogonal wavelet transform* (OWT). The great advantage of the OWT is that it allows the definition of extremely efficient algorithms for computation, for example Mallat's pyramidal algorithm, which can calculate the expansion coefficients in $O(N)$ time, where N is the number of sample points in the time series of interest. There are various families of orthogonal wavelets that can be used; however, by far the most commonly used are the *Daubechies* families [130]. There do not exist closed forms for the Daubechies mother wavelets; instead, they are computed in terms of scaling functions ϕ defined by recursion relations [131],

$$\phi(t) = \sqrt{2} \sum_{k=0}^{2r-1} c_k \phi(2t - k) \tag{8.5}$$

where different sets of coefficients c_k yield different families. The order of the family is controlled by the integer r; mother wavelets of higher order are smoother than those of lower order (i.e. more differentiable), but the low-order families have more compact support. The lowest-order mother wavelet in the Daubechies family is the Haar wavelet which essentially yields a square-wave expansion basis.

Once the wavelet coefficients have been computed, one can construct a decomposition in terms of *wavelet levels*,

$$x_m(t) = \sum_k x_k^m \psi_{m,k}(t) \tag{8.6}$$

where the sum is over all translation parameters at a common scale (which is roughly 1/frequency), and thus provides a decomposition in terms of frequency band-limited time series. The different levels potentially capture different low-frequency components corresponding to EOVs, so the level decomposition is used here as the basis of a new decomposition into stationary and nonstationary parts.

The idea behind the decomposition is quite simple. Having established the level decomposition, one simply applies the ADF test at each level in order to establish if a given level time series is stationary or not. It is clear that one level will capture the critical frequency f_c, so that level – denoted by the index L^* – has a frequency band covering the transition from stationary to nonstationary. Summing the levels up to, and including level L^* will determine the nonstationary component of the original signal (as dictated by the ADF test) and the sum over the remaining levels will yield the stationary component. By applying the ADF test level-wise, one does not need to pre-determine the critical frequency f_c.

As before, for the assessments in this paper, the 99% confidence interval on the ADF test is used, which corresponds to a t-critical value of -2.56.

The decomposition proposed here can be made a little more informative, with a little more effort. First of all, the mean of the signal can be extracted before the wavelet decomposition; secondly, one can make a further separation within the stationary component in order to identify a 'noise' component.

8.2.2 Autocorrelation Functions (ACFs)

According to [132], for a stochastic process X_t for a finite sample, with mean μ and sample variance ${\sigma_x}^2$, which are both time-independent, the autocorrelation function (ACF) ϕ_{xx}, can be defined which determines the degree of dependence between any two samples in the process (signal). For stationary signals, the ACF depends only on the time interval τ between the pair of values considered and not their individual positions in time,

$$\phi_{xx}(\tau) = \frac{E\left[\left(X_t - \mu\right)\left(X_{t+\tau} - \mu\right)\right]}{\sigma_x^2} \tag{8.7}$$

where E again denotes the expectation operator.

Now, if one defines 'noise' as a component within the signal with *no* temporal correlations (i.e. one cannot predict any future values from past values), one can test for this property using the ACF. A signal will be termed noise, if it is *delta-correlated* i.e.

$$\phi_{xx}(k) = \delta_{k0} \tag{8.8}$$

with the index k now representing the sample index in terms of discrete time. With the normalisation given in equation (8.7), one can compute confidence intervals for a zero result; the intervals corresponding to 99% confidence are $\pm 2.54/\sqrt{N}$.

In order to establish the noise component in the signal, one computes the ACF for any levels in the stationary component (stationarity is required in the definition of the ACF) and moves into the noise component any levels which are delta-correlated.

This completes the description of how the new decomposition is defined. To summarise, the steps are as follows:

1. Remove the mean from the signal $\{x_i : i = 1, \ldots, N\}$. The mean μ is the first component.

2. Compute the wavelet transform, and determine the wavelet levels. (Note that this may require truncation or zero-padding in order to have 2^M points for the OWT, if N does not satisfy this condition already.)

3. Apply the ADF test level by level and determine L^*.

4. Resum the levels corresponding to the stationary and nonstationary parts.

5. If desired, one can calculate the ACFs for the stationary levels and separate out the noise component.

The procedure will now be illustrated on the second natural frequency and the air temperature of Z24 Bridge.

8.3 Second Natural Frequency of Z24 Bridge

As the motivation for this work came from the SHM context in the first place, the case study will be taken from data acquired in a major SHM campaign based on the monitoring of the Z24 Bridge. The second natural frequency from the bridge has already been encountered in Figure 8.1. The Z24 Bridge (overpass) was constructed in Switzerland in the beginning of the 1960s in order to link Koppigen and Utzenstorf by passing over the main national Highway A1 (Lausanne-Zurich). The bridge was superceded by a new construction in the late 1990s, and authorisation was given for an SHM campaign before its demolition in the autumn of 1998. The exercise, which included the introduction of realistic but controlled damage, is described in [4]. Various sensor signals were acquired, these included measurements of the ambient environmental variables like air and deck temperatures, humidity etc. The main features for SHM were modal quantities extracted automatically by stochastic subspace identification (SSI) at regular intervals [46]. The signal examined here will be the second natural frequency series.

From a first look at the signal (Figure 8.1) one can discern two major components. There is a high-frequency component, which is essentially the uncertainty associated with the SSI estimates; one would expect this component to be stationary. There is also a clear low-frequency trend in the time series that reflects the EOVs experienced by the bridge over the year. The most dramatic features in the signal are the peaks, which are associated with stiffening of the bridge asphalt layer during sub-zero temperature conditions in the winter months [46].

Moving now to the decomposition, Figure 8.2 shows the levels from the wavelet/MRA analysis. The Daubechies 4^{th} order wavelet was used as a compromise between smoothness and support. In total, 12 levels were extracted. Note that levels zero and one both seem to be copies of the mother wavelet; in fact the sum of the two

levels carries the mean of the original signal, according to the implementation of the algorithm, which was taken from the book *N*umerical Recipes [133]. Furthermore, the algorithm is applied in such a way that the levels are numbered to increase with increasing frequencies. It is immediately obvious from the figure, that the trend behaviour – from the EOVs – is captured in the lower levels.

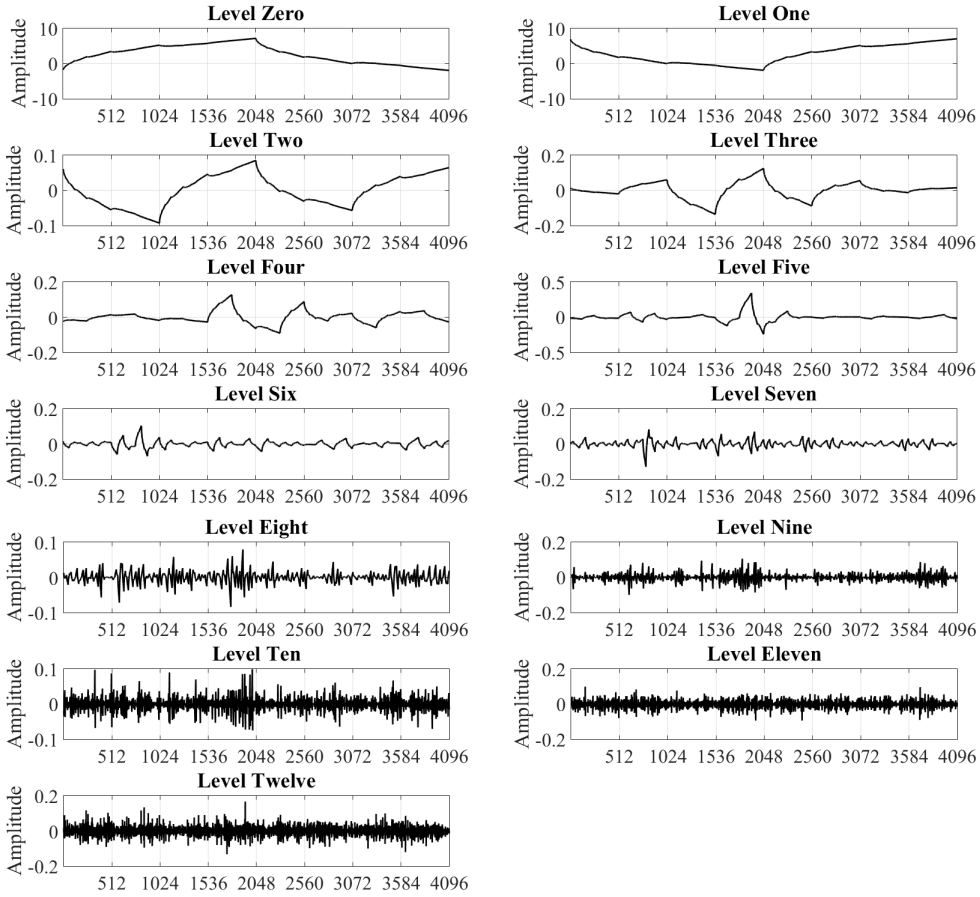


Figure 8.2: Wavelet decomposition/MRA levels for the Z24 second natural frequency.

The ADF test was then applied to each level in turn, and the statistic values obtained are summarised in Table 8.1. At the 99% confidence level applied here, the critical t-statistic is -2.56; this means that the critical level here is $L^* = 5$, and the sum of levels 0 to 4 is determined as the nonstationary component of the signal.

In addition, the discrete Fast Fourier (DFFT) can be used on the obtained MRA levels [134]. These are depicted in Figure 8.3 and show that in lower-frequency levels there is only one dominant frequency, significant low one, which dominates

the signal. As the number of levels increases, the energy is distributed in more than one frequencies inside DFFT spectra, something which provides evidence for the presence of stationarity in the signal.

Table 8.1: T-statistics corresponding to different wavelet/MRA levels.

Level	t_{ADF}
0	-0.025
1	0.705
2	-0.1929
3	-1.103
4	-1.501
5	-2.724
6	-5.021
7	-9.030
8	-15.177
9	-25.852
10	-42.085
11	-87.96
12	-93.58

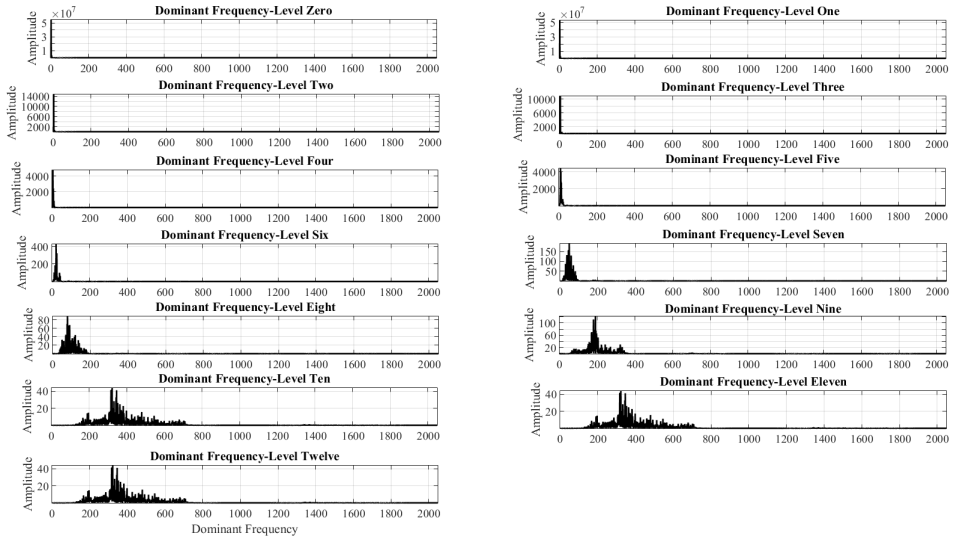


Figure 8.3: DFFT response spectra for the MRA levels.

The final stage of the decomposition is to compute the ACFs for the stationary levels, so that one can separate out the delta-correlated levels as a noise component. The ACFs are given in Figure 8.4 (the ACFs for the nonstationary levels are also shown here for comparison); the red dashed lines in the figures are the 99% confidence levels. In this case, levels 10-12 are considered sufficiently delta-correlated and are separated out as a noise component.

The final decomposition is shown in Figure 8.5. There is a clear separation into the nonstationary and stationary components.

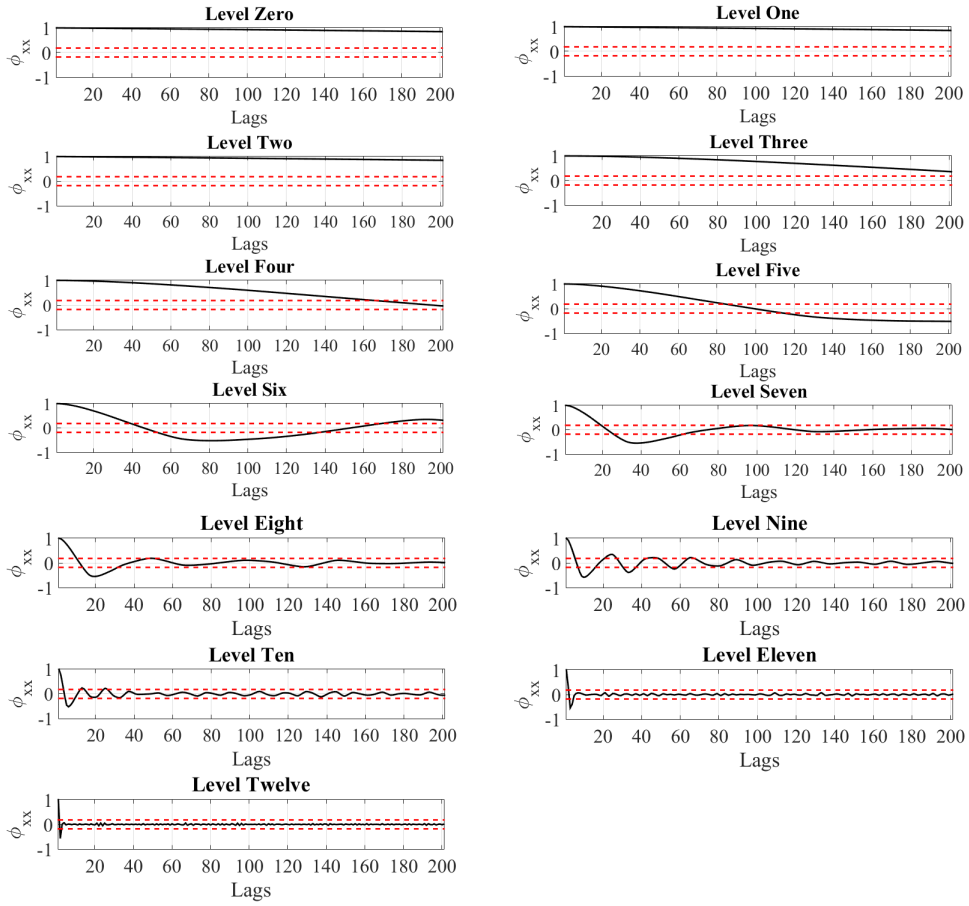


Figure 8.4: Autocorrelation functions corresponding to each wavelet/MRA level. The red dashed lines indicate the 99% confidence interval.

The further separation into a stationary and a noise component also appears meaningful, with the noise component resembling more the Gaussian white noise sequence one might associate with noise; however, one should recognise that the component is the result of essentially high-pass filtering the original signal, and therefore cannot be truly white.

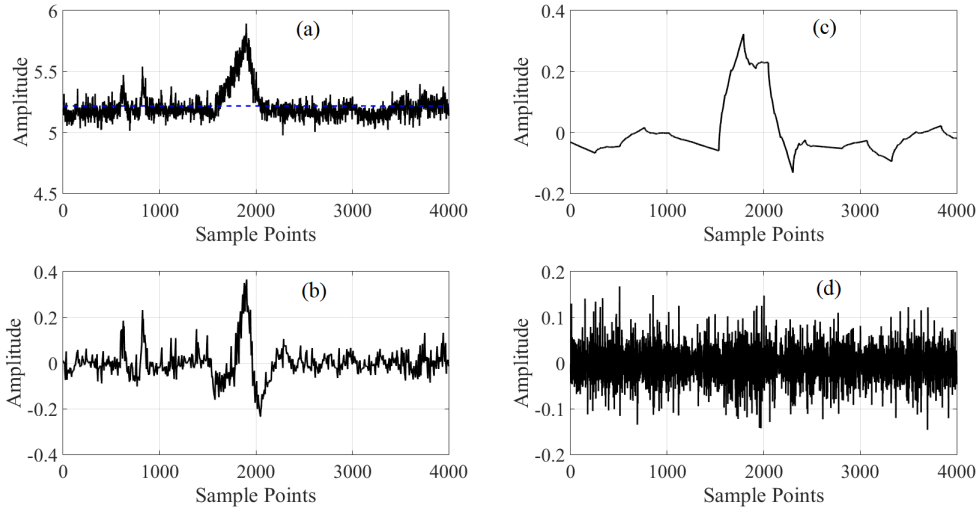


Figure 8.5: Decomposition of the Z24 second natural frequency: (a) original signal; (b) nonstationary component; (c) stationary component; (d) noise. The mean of the original signal is indicated as a blue dashed line in (a).

8.4 Z24 Air Temperature

In the same way as done previously for the second natural frequency, the air temperature of Z24 is analysed. The series of air temperature is illustrated in Figure 8.6, showing that in the first 1500 observations some cyclic repetition is present, which stops during winter months and especially inside February $(1500 - 2000)$, where many observations found to be under 0^oC. After the winter period and more specifically after 2048 observations, the air temperature signal exhibits an increasing trend until reaching its maximum of 33.4 0^oC. From the latter observations it is expected that a trend and an cyclic part are present in the signal and characterise the variation of temperature. To further investigate the temperature signal, the MRA is employed.

The MRA is based once again on a $DB4$ wavelet and the results can be observed in Figure 8.7. In total 12 MRA levels are obtained from the MRA, with lower levels to represent the most nonstationary components and the high-frequency ones the most stationary. In order to distinguish the nonstationary from the stationary levels, the ADF test employed and the MRA levels are assessed. The results of ADF test are shown in table 8.2. From these concluded that the critical MRA level (L^*) after which the levels are considered stationary is level 4. Once again the 99% confidence

interval employed, which corresponds to a t-critical value of -2.56.

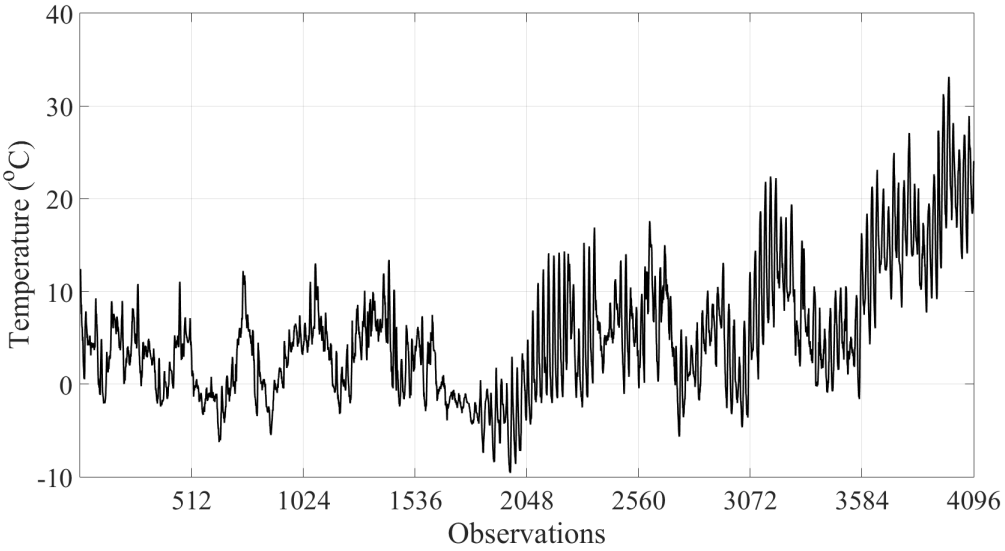


Figure 8.6: Z24 air temperature variation.

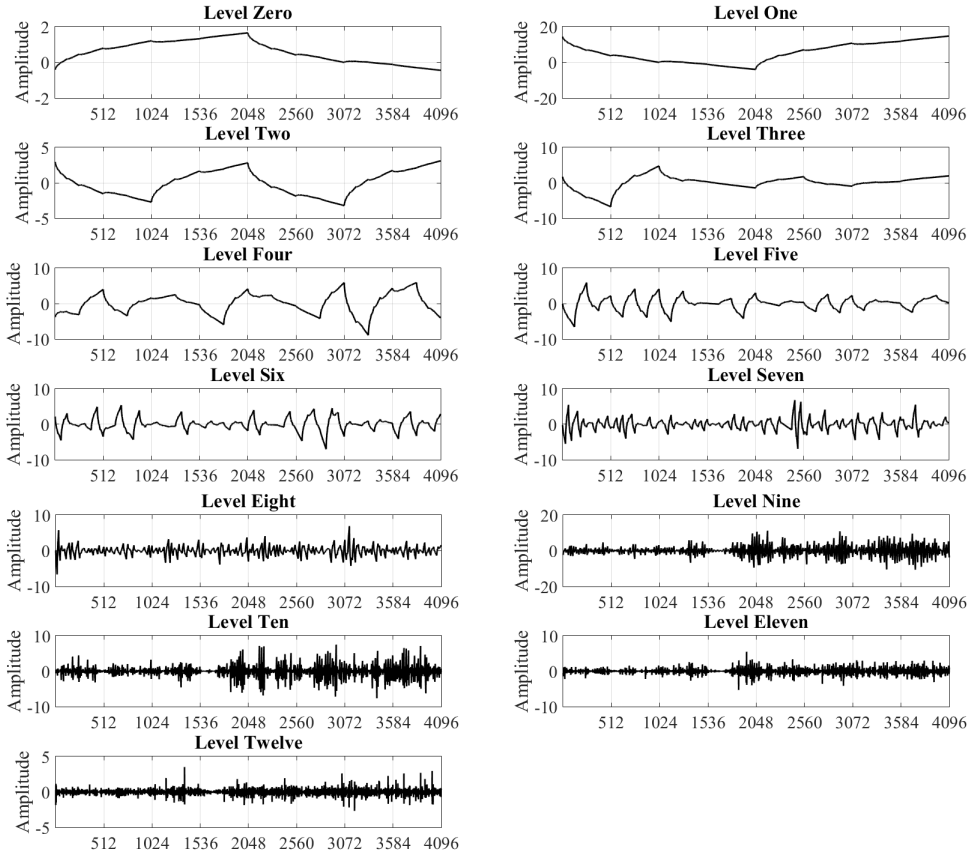


Figure 8.7: MRA decomposition of Z24 air temperature series.

Table 8.2: T-statistics corresponding to different MRA levels for Z24 air temperature.

Level	t_{ADF}
0	-0.026
1	0.705
2	-0.044
3	-0.713
4	-1.441
5	-3.312
6	-4.559
7	-9.197
8	-14.685
9	-25.314
10	-38.741
11	-80.559
12	-95.317

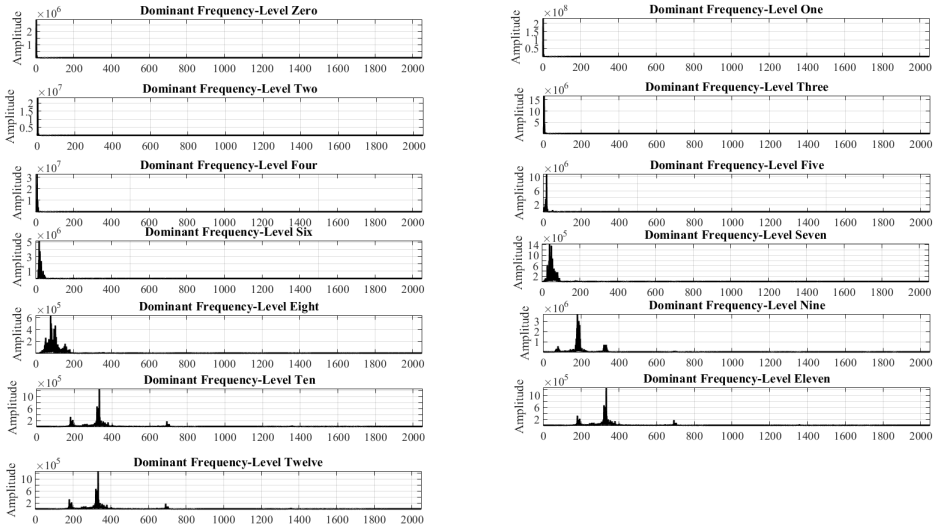


Figure 8.8: DFFT analysis results for the obtained MRA levels.

In addition to ADF test, the DFFT analysis also performed, as shown in Figure 8.8. The latter shows that as the number of MRA levels increase, the energy of the signal is distributed to more frequencies, something informing for the presence of stationarity.

Finally, to identify the noisy part of the signal, the ACF analysis is performed. This is shown in Figure 8.9. According to ACF analysis, the last two MRA levels (level eleven and twelve) can be considered as *delta-correlated* and can be used in order to form the noisy part of the signal. Also here, one can argue that the noisy signal is not exactly a white noise one.

Based on the aforementioned, the original signal of $Z24$'s air temperature can be decomposed into four parts; a nonstationary, an stationary, a noisy and an mean part. All the aforementioned parts are illustrated in Figure 8.10, where blue dashed lines used to indicate the mean value of each part.

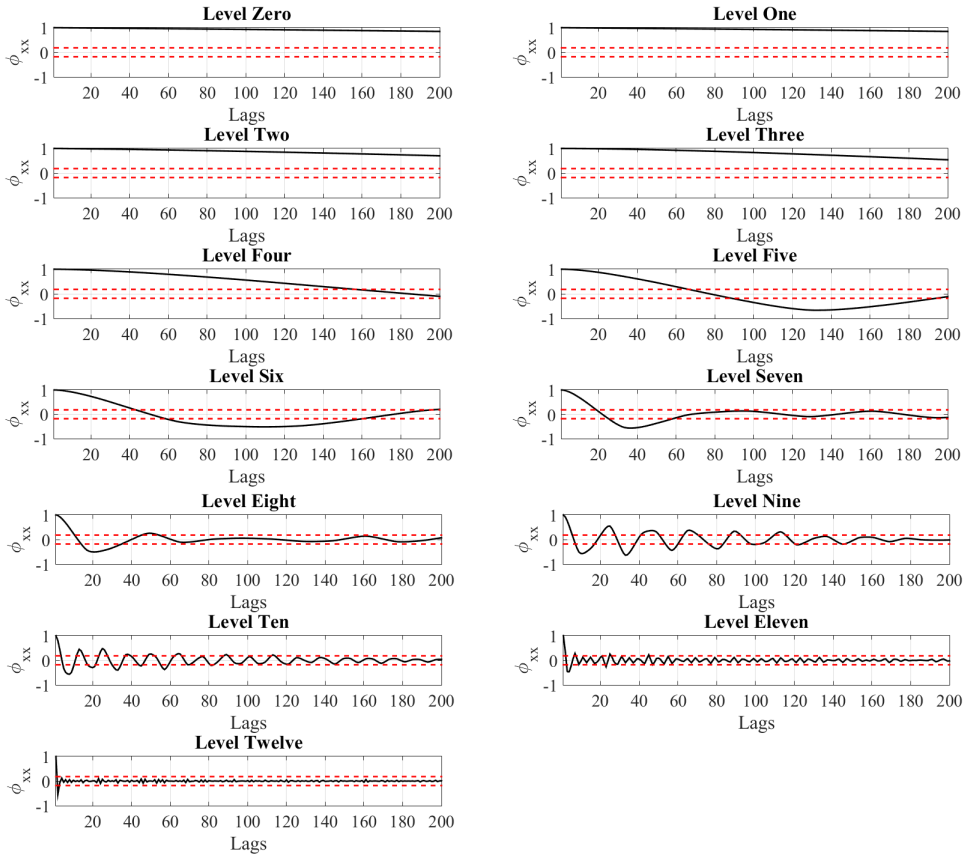


Figure 8.9: ACF analysis on the MRA levels of Z24 air temperature signal.

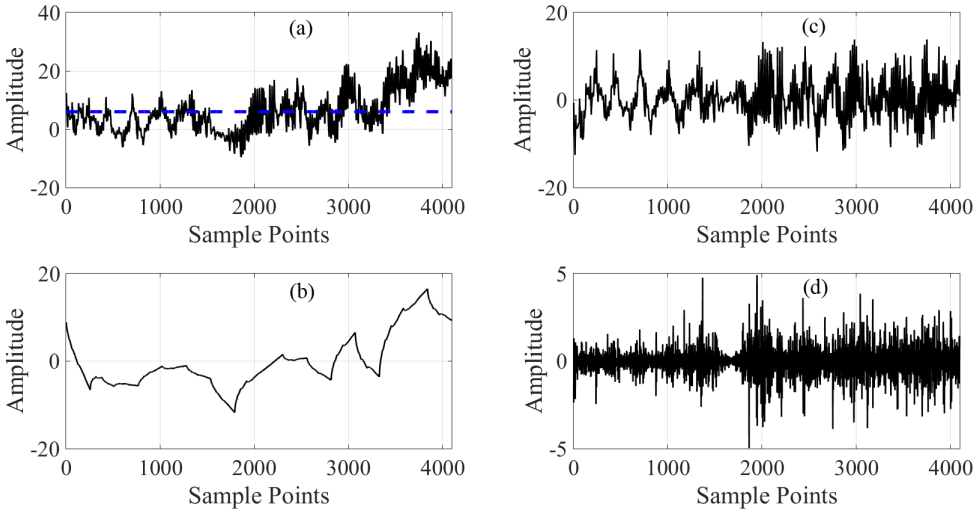


Figure 8.10: Decomposition of temperature signal: (a) original signal; (b) nonstationary component; (c) stationary component; (d) noise. The mean of the signal is indicated as a blue dashed line.

8.5 Conclusions

In this Chapter, a new decomposition method is presented, able to separate stationary and nonstationary components of signals. The basic idea is that a wavelet/MRA is used to separate the time series of interest into a set of frequency band-limited wavelet levels. The levels are then individually tested for stationary by using the ADF statistic; the stationary levels are re-summed to give the stationary component of the original series and likewise the nonstationary levels. It is shown that one can also separate a 'noise' component from the stationary component by using autocorrelation functions (ACF). The 'noise' component is not a Gaussian white noise, but is distinguished by the fact that it is delta correlated, so that future values of the component are independent of past values. The decomposition method is demonstrated successfully on signals from a SHM campaign on a real structure – the Z24 bridge.

Despite the SHM context selected here, the results of this Chapter are quite general and should be of interest, not only to engineers, but also to econometricians, or anyone seriously concerned with time series analysis where stationarity is an issue.

Chapter 9

Summary and Conclusions

9.1 Importance of SHM for Civil Infrastructure (Bridges)

Structural Health Monitoring (SHM) can provide numerous measures that can be used to develop diagnostic tools in order to capture and evaluate the response of a structural system over its normal condition, as well as under the consequences of unpredicted and undesired events. The main history of SHM originates from the field of aerospace and automobiles, however in the last decades its use extended in civil infrastructure works and especially on large-scale works, such as dams, bridges, off-shore structures, tall-buildings, ports and others.

The main motive behind the use of SHM in greater extent today, comes from the need of society to invest in infrastructure, develop economically and efficiently manage its current assets. As all the previous are subjected to budget considerations, each society needs to develop the mechanisms and tools in order to perform educated decision making and prioritisation of investment. The latter can come as a consequence of the efficient management and evaluation of the existent infrastructure. Inside the aforementioned framework, the role of bridges is of significance importance, due to the fact that bridges are a vital part of civil infrastructure, as they are the main link between roads, gap spans over water and connect cities/countries between each other. The role of bridges is very important for the economic growth of a society, as bridges and roads are the main mean of trading, goods' delivery and

people's transportation.

Although in the past two decades there has been an increasing trend to investigate the use of SHM within the context of bridge engineering, the cases that have been applied in practice are not many. One of the main reasons behind this is the fact that authorities, companies and owners, which are responsible for bridges, are obliged to find ways to efficiently manage a significant/large bridge stock and due to continuous shrinkage of governmental budgets, more cost-effective and less time-consuming solutions are selected, such as visual inspection, sufficiency rating and assessment in terms of current standards of practice [135].

In fact, all of the previously mentioned cost-cutting decisions aim to prioritise budgets and provide a general solution for the majority of bridges. However, these shortages can have a significant long-term impact and cannot always assure the longevity of a structure. For example, the periodic monitoring cannot provide concrete results and conclusions on the effect that environmental and operational impacts have on the structure over time. In addition, the practice of specific standards for assessment, i.e. BD21/01 of DMRB [135], based on development and the evaluation of FE models cannot always provide assurance that the best decision is made for the structure. This becomes even worse in the case that assessments are not based on in-situ observations and experimental testing, i.e. NDT and DT. Furthermore, it is very common in industrial world to avoid the dynamic evaluation of structures and focus on the analysis of structural mechanics and static magnitudes (axial, shear loads and moments).

In the last decade SHM started evolving and its role is expected to be vital in the near future. Many societies are promoting the so-called *Digital Agenda*, which mainly focuses on the exploitation of bridge *big data* analysis, which in other words is the SHM data. The main reason of this promotion is the necessity of society to understand in detail the behaviour of structures on real-time scale (not small-scale experiments), evaluate it continuously and perform a more educated decision making, since factors such as cost and structural significance are of great importance for major structures. This is also the direction that the current thesis has taken.

Here it is important to say that although this thesis is mainly focused on bridges, the methodologies discussed and demonstrated could be applied to many different types of SHM data, provided by continuous monitoring activities.

9.2 Summary based on Layout

The layout of the thesis consists of eight main Chapters. The first Chapter discusses the role of SHM in today's society, its origins, fields of application, multi-disciplinary nature and generalised objectives. In addition, it talks about the subsystems comprising a typical SHM system, the available sensor technologies and categories of bridge SHM, while a review of the conventional practice (visual inspection and condition evaluation) and bridge management systems is made, underlining some basic issues faced by todays developed societies. Moreover, an introduction to damage detection is attempted, emphasising the measures provided by SHM, techniques for DSFs extraction, with the main focus placed on data-based approaches and unsupervised SPR.

In the second Chapter, the main interest of the thesis is placed on novelty detection, focusing on the role of EOVs on the variability of SHM signals, discussing methodologies that have been used inside SHM literature in the past to project out their effect, such as cointegration and PCA. In addition, in the second Chapter the scope and aims of the thesis are discussed.

In the third Chapter, the experimental setup and vibration analysis performed on a laboratory truss bridge model manufactured and tested inside a Sheffield University laboratory are described. The natural frequencies of the truss model are measured between -10 to 25^oC, while three damage cases are simulated by the removal of connecting bolts in two different location of the bridge. The obtained results are analysed in order to evaluate the impact of temperature both in short and longer-term. From this analysis it was observed that the effect of temperature is very significant on natural frequencies, even for relatively small temperature changes between 0.1 to 0.5^oC.

In Chapter 4 the SHM data provided by two real-time SHM campaigns, Tamar Bridge and Z24 overpass, where analysed employing regression modelling, aiming to evaluate the effect of EOVs on them. The impact of temperature in both cases, Tamar Bridge cable tensions and natural frequencies of Z24 was pointed, as the main reason why the series are sharing common trends over time and are correlated between them.

Next, Chapter 5 introduces the concept of *Johansen's approach to cointegration*,

discusses each step of the procedure, concentrates on the way in which the common trends are manipulated inside cointegration and finally demonstrates its use on the data of three bridge cases; the cable tensions of Tamar Bridge, $Z24$ and laboratory truss bridge natural frequencies. The main aim of these is to study the impact that EOVs and damage introduce on real-time and experimental data, focusing on the interrelationships between SHM series. More specifically, from the cable data from the Tamar Bridge, the dominant effect of temperature on cable tension signals is discussed and its effect on series' nonstationarity. The *Johansen's approach to cointegration* employed using multiple cable tension series providing a stationary residual. It is important to add that it was shown that signals sharing inverse proportional trends can be used together inside cointegration and a provide a stationary residual, capable also for novelty detection. Then, in the case of $Z24$, the effect of temperature on the relationships between natural frequencies is studied, as well as the effect that damage have on these. The effect of temperature led to the introduction of a specific pattern on the data, showing an increase of natural frequencies as temperature decreases.

On the other hand, the introduction of damage led to the generation of a data cluster in greater distance from the main data population. Z24 natural frequencies provided a useful example where large artefacts observed on the natural frequency variation, over a year of monitoring, can affect the novelty detection capability of the cointegration residual. In particular, it was observed that the inclusion of all large artefacts during the training of cointegration can lead to incapability detecting damage in the residual signal. Therefore, it is important to select the most appropriate training set, that can both provide a stationary residual and is still sufficiently damage sensitive. However, from further investigation on non-linear cointegration [82], the damage detection sensitivity of cointegration residual can be further increased. Finally, from the laboratory data, the effect of temperature on small-scale structures is discussed. In particular, temperature is observed to introduce short-term nonstationarity in the signals, affecting the relationships between natural frequencies and consequently the cointegration residual. However, cointegration was able to detect damage in all damage cases simulated.

The sixth Chapter speculates on the nature of stationarity and how this can be quantified using nonstationarity measures, such as the ADF test. As a matter of fact, it is difficult to judge a finite sample, such as an SHM signal, as a judgement on stationary or nonstationarity only based on the mathematical definition of weak

stationarity. Therefore, more complex measures, such as ADF test can be used for this purpose. However, in order to use ADF a main hypothesis is made. It is assumed that a good way to judge if a series is stationary or not is to count the number of dominant periods of the signal. This Chapter proposed that the ADF is a function of this number, This idea was investigated through three groups of simulations, which are based on dimension analysis and basic concepts of signal processing. It was found that the main factors on which the ADF t-statistic depends on are the number of observation multiplied by time interval over the dominant signal's period ($N\Delta t/T$), which basically counts the number of signals dominant periods, the normalised frequency (f_n), which describes the ratio between the dominant frequency (f) and the sampling one (f_s) and finally, the signal to noise ratio ($\sigma_y^2/\sigma_\epsilon^2$; SNR). Another important conclusion made was that given the sampling frequency (f_s) of any signal and setting the significance level (critical t-statistic), a critical frequency (F_c) in Hz can be found, which depends on the normalised critical frequency (f_c), which can be used to judge if a signal is stationary or not. For this analysis the significance level set at 99% confidence intervals.

Next, the seventh Chapter discussed the ways that have been used in the past inside SHM literature, in order to monitor the cointegration residual. The use of different SPC charts and nonstationary testing are demonstrated and compared using natural frequency data generated by a simple mass-spring system. Two main conclusions were drawn from this analysis; (i) stationarity is needed in order to monitor efficiently the cointegration residual using SPC charts and (ii) nonstationaity testing, i.e. ADF, can be used for damage detection, but it is a weak measure of novelty, because the ADF provides the mean between the stationary and nonstationary part of the signal and in cases that the noise amount is relative high, its novelty detection capabilities are low.

Furthermore, Chapter eight discussed and demonstrated the use of the discrete orthogonal wavelet transform (multiresolution analysis; MRA), ADF and ACFs for the decomposition of SHM signals. In particular, a way to perform MRA in a more principal manner is described, as the signal under consideration can be analysed into multiple frequency components and using ADF, these can be distinguished as nonstationary and stationary. In addition, ACFs can be used to find those stationary MRA levels that can be considered as *delta-correlated* and used to form the noisy part of the signal. In this way, the original signal can now be decomposed into a mean, a nonstationary, a stationary and a noisy part. This type of analysis can

be useful not only inside SHM but also for those interested in time-series analysis. The low-frequency levels can be associated with nonstationarity, while the high-frequency ones can be associated with stationarity. In case of damage, it is expected an introduction of nonstationarity in the SHM signals, which shows that the low-frequency components are more damage sensitive that the higher ones.

9.3 Novelty of Thesis

The main aim of the author was to try and present novel results in each chapter of the thesis. Besides the first two mainly theoretical chapters of the thesis, some novel ideas discussed and important results for future research generated. To begin with, in the third chapter a completely new data set generated by the experimental laboratory truss bridge test. These results are used inside the thesis as well as can be shared for future work on them. Going now on the fourth chapter, it was the first time that the cable tension measures of Tamar Bridge was analysed providing interesting results and explanation on the effect that EOVs impose on them. This particular part of the fourth chapter is planned to be published as an interesting show case.

Chapter 5, besides talking about the Johansen's approach to cointegration, showed and discussed the results of its application in three different real-time bridge SHM data. This was very useful to understand what are the current limitations of the cointegration analysis. Furthermore, during the Z24 natural frequencies analysis, a brief description is given about the current development on nonlinear cointegration and which is main problems that the linear cointegration is facing at the moment.

Going to chapter 6, a completely new interpretation about the ADF test is provided. This is one of the most commonly employed nonstationarity tests and based on dimension analysis and basic signal processing properties, it was shown that the ADF is mainly counting the number of cycles of the fundamental periodic component of the signal. The latter was shown through a series of simulations using a simple sinewave funtion. However, the most important point of this analysis was that assuming that one knows the sampling frequency of a signal, it is possible to determine a critical frequency in Hz after which the signal from nonstationary becomes stationary. This is important not only inside SHM activities, but also to anyone interested in signal processing, time series analysis, because provides an interesting

perspective on how to judge if a signal is stationary or not.

The seventh chapter, through two examples, mass spring system and laboratory truss bridge test, gave a review of the most useful ways to track novelty on the cointegration residual. From this it was shown that the fundamental moments of stationarity, such as the mean and variance can be tracked down to provide informative inference about the existence of nonstationarity and consequently abnormality. It is also important here to say that the ADF was tested as a measure of novelty and damage detection. ADF test has been used in the past as an detector for SHM. However, after the analysis and explanation provided it was shown that it can be characterised as a weak one, because the t-statistic provides the average between the stationary and nonstationary component of the signal, and performs better when novelty dominates the signal.

Finally, chapter 8 provided a principled way to perform multiresolution analysis, in order to decompose a given signal into each mean, stationary, nonstationary and noisy components. In particular, the ADF test was able to identify a critical level L* after which the signal becomes stationary, whilst the use of autocorrelation functions was able to provide the noisy part inside the signal. This type of decomposition is important inside SHM because different parts of the signals can be used inside regression with available EOVs or inside cointegration to enhance the damage detection capability of Johansen's approach to cointegration. In addition, this concept can be used inside other scientific fields such as time series analysis, signal processing and from anyone with interest in signal decomposition.

9.4 Limitations of Cointegration and Suggestions for Further Research

Although the ADF and MRA methods can be used as standard-alone procedures, the main method and direction followed in this thesis is to investigate and, if it is possible, to improve Johansen cointegration. As it has already been mentioned throughout the thesis, the two main advantages of cointegration inside the context of SHM are its capability to quantify and project out the impact of EOVs and the fact that it provides a feature (cointegration residual) which is damage sensitive and stationary, thus it can be used in combination with SPC charts for novelty detection.

Based on the two above advantages, one can investigate further the ways in which the method can be improved in the part of quantification and elimination of EOVs impact and secondly at the part of monitoring the cointegration residual. The current thesis investigates more the first main issue, making an attempt to understand and simply interpret one of the most commonly employed nonstationarity measures, that of ADF t-statistic. This is incorporated inside the Johansen's approach to cointegration. Secondly, the thesis tries to discuss the fact that the impact of EOVs occurs on a widely disparate time-scales and therefore its elimination is not a straight-forward procedure. For the latter, is proposed a more principled way to perform MRA in order to decompose a given signal, aiming to obtain those decomposition levels that are more damage sensitive and can be used inside Johansen cointegration. It should be mentioned that there is a great amount of work done on the aforementioned issue, employing methods, such as MRA [105], PCA [63] and non-linear cointegration [2, 82].

On the other hand, there is the part of monitoring the cointegration residual. Chapter 6 focused on this issue, presenting some of the basic methods which can be employed. In particular, from the analysis it was concluded that SPC charts such as X-bar and S-bar seems to be the best way to perform cointegration residual monitoring, at least for now. The analysis performed here on both synthetic data generated using a simple mass-spring system, as well as the natural frequencies of laboratory truss bridge. In addition, the last part of Chapter 6 explored the use of ADF t-statistic for residual monitoring. Based on the discussion of Chapter 5 about the interpretation of ADF t-statistic and the results of Chapter 6, it was concluded that ADF can be used for damage detection purposes when a high amount of nonstationarity is introduced in a signal, however, theoretically and pragmatically it is a weak measure of novelty, because it depends on the relative sizes between the periodic and noisy components of a signal, providing the mean value of these two. In other words, if the stationary part of a signal which is assessed for damage, has a high high-frequency content (more stationary signal), then ADF t-statistic is not a strong indicator of novelty.

The aforementioned, in general sense, are the main points where the investigation for Johansen's approach to cointegration focused. In the two paragraphs that follow, a discussion is made on two practical issues that can be considered as limitations for the use of cointegration inside SHM, while two suggestions about future work on Johansen's approach are also made, with the emphasis placed in application to

bridges.

9.4.1 Practical Limitations

Two practical limitations of Johansen cointegration are discussed in this section and some thought is given in ways these can be dealt with.

To begin with, the first limitation of Johansen cointegration is the fact that its use is for data coming from continuous monitoring of structures. This issue is general (does not concerns only cointegration) and is probably the main reason why SHM is not applied extensively in civil infrastructure so far. Considering the case of bridge applications, the relative cost of instrumentation, maintenance and generating data-based results, is not cost-effective and can be time consuming. This happens, more specifically, in the case that the efficient management of the bridge stock of a whole nation is needed. The budget for the assessment and repair for each bridge is significantly low. Hence, due to the relatively high cost for setting an SHM system, such a solution may not be considered. Furthermore, the second limitation that is discussed here is that cointegration is based on the concepts of data machine learning and therefore data are needed for training purposes. The existence of a newly installed SHM system in a bridge does not mean that there are enough data to work with. As a matter of fact, time is needed in order to collect and process data of structural response in order to draw conclusions based on that. In addition, there is a lack of data coming from the damage state of bridges and supervised methods cannot been employed, because NDT is not permitted due to bridges' significant structural importance to the society.

There is no immediate measure that can be followed in order to solve the previously mentioned practical limitations. These issues will be solved during time. In particular, it is a fact that the majority of the newly designed bridges and many major bridges around the world nowadays, have already monitoring systems installed on them. The latter means that the construction budget planning for a bridge includes also funding for SHM systems. This shows that many existent bridges, constructed the past two decades, while also any new that will be constructed or rehabilitated in the upcoming years will need the expertise of SHM. Therefore, it is of great importance for the scientific community and industry to prepare the ground for the role of SHM for civil infrastructure.

In other words, although the two aforementioned issues have clearly a practical character explaining the reasons for which SHM is not extensively used in industrial context, however, it is explained that in the upcoming decades SHM will be become necessary for the assessment and maintenance planning of civil infrastructure.

9.4.2 Suggestions for Further Research using Cointegration

Johansen's approach to cointegration is already an established methodology inside SHM and unsupervised learning. Its main concept is based on its capability to remove the impact of EOVs from the damage sensitive features and provide stationary signals (cointegration residual) that can be used for novelty detection. Since cointegration can provide successful measures for damage detection, further research can be orientated in the direction of damage localisation. For example, in cases where multiple SHM series are available, multiple cointegration residuals can be generated and can be assessed based on probabilistic statistical measures, to provide indications (i.e. alarms) for potentially damaged locations in a structure.

Furthermore, one interesting field in which cointegration and more specifically the error correction model (ECM) can be explored, is the future response prognosis/forecasting. Some basic information about the use of ECM in such a way are provided in [136]. Multiple forecasts of the model can be generated, depicting multiple future scenarios for expected response. These can be employed in combination with existent or newly formed deterioration and decision models in order to perform life-cycle analysis (LCEA) of structures for a more efficient management and decision making. Some interesting ideas for LCEA can be found in [137],[138].

In particular, in the field of bridge engineering, there is a specific area focusing on bridge management systems (BMS). This area is separated into two types of management; the project-level bridge management (P-LBM), which includes in-situ inspection, testing, maintenance requirements, prevention measures, remedial methods and monitoring strategies, while there is also the network-level bridge management (N-LBM), which includes estimation rates of deterioration, prediction of future condition (i.e. using Markov chain models), planning optimal maintenance programs, prioritising maintenance, assess effectiveness of different strategies, estimate costs and budgets [19].

It seems that in the case that Johansen's approach to cointegration test, explored

for its prognosis/forecasting capabilities could be incorporated successfully and/or in combination with other methods for life-cycle estimation for N-LBM.

Appendix A

Cable Tensions and VAR model

A.1 Example of Tamar Bridge's Cable Measuring Creep and Restoring Deformation

Such cables (referred to as cable Type 5 in Tamar Bridge references) are commonly installed in stay-cable and suspension bridges, because cable forces require control and adjustment during construction stages [56]. The latter can be found in Appendix A. Figure A.1 depicts the configuration of the stay-cable system of Tamar Bridge and the location of the aforementioned cables (red dashed line). It should be mentioned that there are two, in total, such cables one in each side of the deck (DS077SD1 and DS077ND1). The tension signal of cable DS077SD1 is depicted in Figure A.2

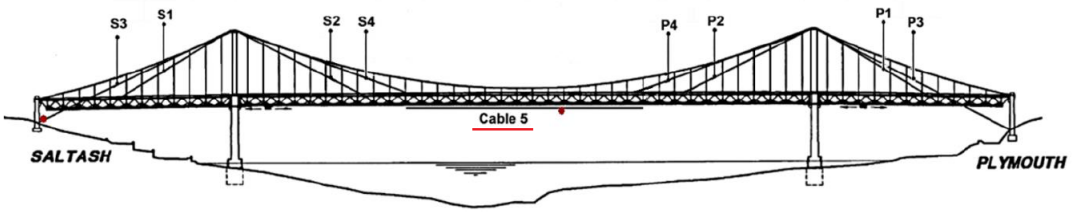


Figure A.1: Stay-cable system configuration and cable type 5.

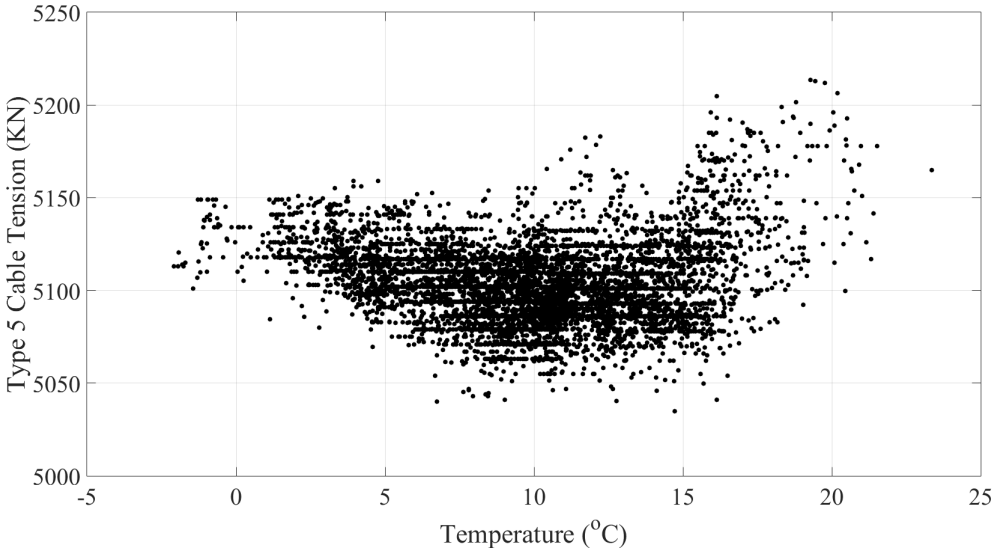


Figure A.2: Tension signal of cable SS077SD1 with respect to temperature.

A.2 VAR Model Demonstration

From this Figure an interesting observation is that cable tension increases both when temperature increases and decreases. Assuming a typical cable, this has an initial curvature in its configuration. During low temperature it should be in compression, while during higher temperatures in tension. In other words, during longitudinal contraction and expansion of the bridge due to thermal strains developed in the truss. However, it seems that the cable has a specific function, which is to restore the deformations and hence this is the reason why the tension increases.

chapterVAR Model Demonstration For demonstration purposes, a *VAR* model with one lag (VAR(1)) including two (Y_{1t} and Y_{2t}) variables ($k = 2$) is presented. The VAR model can be written in terms of a variable vector (Y_{it}) (vector of $I(1)$ series), a coefficient matrix $[a_{ij}]$ (or alpha matrix), a vector of lagged terms $Y_{i,t-i}$ and a vector of error terms e_{it} (equation (A.1)). In equation (A.1), more lag terms and variables can be introduced. The equation (A.1) can be written also as given in equation (A.2), for $i = 1, .., n$, where n is the number of series and j defines the model order (lags).

$$\begin{Bmatrix} Y_{1t} \\ Y_{2t} \end{Bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{Bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{Bmatrix} + \begin{Bmatrix} e_{1t} \\ e_{2t} \end{Bmatrix} \tag{A.1}$$

$$\{Y_i\} = \sum_{j=1}^{p} [A_i] \{Y_{t-p}\} + \{e_i\} \tag{A.2}$$

Bibliography

[1] LMS Instruments. Lms scadas iii data acquisition front-end, breda, the netherlands. Technical report, Siemens, 2009.

[2] E.J. Cross. *On Structural Health Monitoring in Changing Environmental and Operational Conditions*. PhD thesis, University of Sheffield, 2012.

[3] K.Y. Koo, J.M.W. Brownjohn, D.I. List, and R. Cole. Structural health monitoring of the tamar suspension bridge. *Structural Control and Health Monitoring*, 20(4):609–625, 2013.

[4] C. Kramer, C.A.M. de Smet, and G. de Roeck. Z24 Bridge damage detection tests. In *Proceedings of IMAC 17, the International Modal Analysis Conference, Kissimmee, FL, USA*, pages 1023–1029, 1999.

[5] V.M. Karbhari and F. Ansari. *Structural Health Monitoring of Civil Infrastructure Systems*. Elsevier, 2009.

[6] H. Sohn, C.R. Farrar, F.M. Hemez, and J.J. Czarnecki. A review of structural health review of structural health monitoring literature 1996-2001. Technical report, Los Alamos National Laboratory, 2002.

[7] C.R. Farrar and K. Worden. *Structural Health Monitoring: A Machine Learning Perspective*. John Wiley & Sons, 2012.

[8] A. Rytter. *Vibrational Based Inspection of Civil Engineering Structures*. PhD thesis, Dept. of Building Technology and Structural Engineering, Aalborg University, 1993.

[9] B. Bakht and A. Mufti. *Bridges: Analysis, Design, Structural Health Moni-*

toring and Rehabilitation. Springer, 2015.

[10] L. Duan and W.F. Chen. *Bridge Engineering Handbook*. CRC Press, 1999.

[11] L.A. Bisby and M.B. Briglio. An introduction to structural health monitoring. Technical report, ISIS Educational Module, 2005.

[12] T.H. Yi, H.N. Li, and M. Gu. Optimal sensor placement for structural health monitoring based on multiple optimization strategies. *The Structural Design of Tall and Special Buildings*, 20(7):881–900, 2011.

[13] S. Lovejoy. Applications of structural health monitoring to highway bridges. Technical report, The Oregon Department of Transportation, 2010.

[14] H. Wenzel. *Health Monitoring of Bridges*. John Wiley & Sons, 2008.

[15] Highways England. *The Design Manual for Roads and Bridges*. Highways England, 2015.

[16] ARMY TM. Bridge inspection, maintenance and repair. Technical report, ARMY, 1994.

[17] F.K. Chang. *Structural Health Monitoring 2000*. CRC Press, 1999.

[18] American Association of State Highway and Transportation Officials (AASHTO): Load Rating Factor Design (LRFD). Specifications for highway bridges. Technical report, American Association of State Highway and Transportation Officials (AASHTO), 1997.

[19] G.AR. Parke and N. Hewson. Manual of bridge engineering. Technical report, Instutition of Civil Engineers (ICE), 2008.

[20] S.K.U. Rehman, Z. Ibrahim, S.A.M̃emon Zainah, and M. Jameel. Nondestructive test methods for concrete bridges: A review. *Construction and Building Materials*, 107:58–86, 2016.

[21] A.W. Herrmann. Asce 2013 report card for america's infrastructure. *IABSE Symposium Report*, 99(33):9–10, 2013.

[22] A.C. Estes. *Bridge Maintenance, Safety, Management and Life-cycle Optimization*. Taylor & Francis, 2011.

[23] European Commission Community Research and Development Services (CORDIS). Project pantura: Flexible processes and improved technologies for urban infrastructure construction sites. Technical report, Project EU 261172, 2011-2013.

[24] R. Brincker and C. Ventura. *Introduction to Operational Modal Analysis*. John Wiley & Sons, 2015.

[25] R.J. Barthorpe. *On Model-and Data-based Approaches to Structural Health Monitoring*. PhD thesis, University of Sheffield, 2010.

[26] M.I. Friswell. Damage identification using inverse methods. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1851):393–410, 2007.

[27] S.W. Doebling, C.R. Farrar, M.B. Prime, and D.W. Shevitz. Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: A literature review. Technical report, Los Alamos National Lab., NM (United States), 1996.

[28] P. Cawley and R.D. Adams. The location of defects in structures from measurements of natural frequencies. *The Journal of Strain Analysis for Engineering Design*, 14(2):49–57, 1979.

[29] K. Christodoulou and C. Papadimitriou. Structural identification based on optimally weighted modal residuals. *Mechanical Systems and Signal Processing*, 21(1):4–23, 2007.

[30] A. Bagchi, J. Humar, H. Xu, and A.S. Noman. Model-based damage identification in a continuous bridge using vibration data. *Journal of Performance of Constructed Facilities*, 24(2):148–158, 2010.

[31] M.M.A. Wahab and G.D̃e Roeck. Damage detection in bridges using modal curvatures: Application to a real damage scenario. *Journal of Sound and vibration*, 226(2):217–235, 1999.

[32] T. Toksoy and A.E. Aktan. Bridge-condition assessment by modal flexibility. *Experimental Mechanics*, 34(3):271–278, 1994.

[33] R. Plude, J. Christenson, J. DeWolf, and A. Jamalipour. Quantifying damage measures for a composite steel girder bridge using finite element analysis.

Structural Health Monitoring, 2011.

[34] T. Guo and D.M. Frangopol Y. Chen. Fatigue reliability assessment of steel bridge details integrating weigh-in-motion data and probabilistic finite element analysis. *Computers & Structures*, 112:245–257, 2012.

[35] D.V. Val, M.G. Stewart, and R.E. Melchers. Life-cycle performance of rc bridges: Probabilistic approach. *Computer-Aided Civil and Infrastructure Engineering*, 15(1):14–25, 2000.

[36] Z. Zong, X. Lin, and J. Niu. Finite element model validation of bridge based on structural health monitoringpart i: Response surface-based finite element model updating. *Journal of Traffic and Transportation Engineering (English Edition)*, 2(4):258–278, 2015.

[37] M.L. Fugate, H. Sohn, and C.R. Farrar. Vibration-based damage detection using statistical process control. *Mechanical Systems and Signal Processing*, 15(4):707–721, 2001.

[38] E.P. Carden and J.M. Brownjohn. Arma modelled time-series classification for structural health monitoring of civil infrastructure. *Mechanical systems and signal processing*, 22(2):295–314, 2008.

[39] F.P. Kopsaftopoulos and S.D. Fassois. Vibration based health monitoring for a lightweight truss structure: Experimental assessment of several statistical time series methods. *Mechanical Systems and Signal Processing*, 24(7):1977–1997, 2010.

[40] P. Moser and B. Moaveni. Environmental effects on the identified natural frequencies of the dowling hall footbridge. *Mechanical Systems and Signal Processing*, 25(7):2336–2357, 2011.

[41] P. Omenzetter and J.M.W. Brownjohn. Application of time series analysis for bridge monitoring. *Smart Materials and Structures*, 15(1):129, 2006.

[42] J. Kullaa. Damage detection of the z24 bridge using control charts. *Mechanical Systems and Signal Processing*, 17(1):163–170, 2003.

[43] A.A. Mosavi, D. Dickey, R. Seracino, and S. Rizkalla. Identifying damage locations under ambient vibrations utilizing vector autoregressive models and mahalanobis distances. *Mechanical Systems and Signal Processing*, 26:254–

267, 2012.

[44] B. Peeters, J. Maeck, and G.D̃e Roeck. Dynamic monitoring of the z24-bridge: Separating temperature effects from damage. In *Proceedings of the European COST F3 Conference on System Identification and Structural Health Monitoring, Madrid, Spain*, pages 377–386, 2000.

[45] R. Brincker, P. Andersen, and R. Cantieni. Identification and level i damage detection of the z24 highway bridge. *Experimental Techniques*, 25(6):51–57, 2001.

[46] B. Peeters and G.D̃e Roeck. One-year monitoring of the z24-bridge: Environmental effects versus damage events. *Earthquake Engineering & Structural Dynamics*, 30(2):149–171, 2001.

[47] R.G. Rohrmann, M. Baessler, S. Said, W. Schmid, and W.F. Ruecker. Structural causes of temperature affected modal data of civil structures obtained by long time monitoring. In *SPIE Proceedings Series*, pages 1–7. Society of Photo-Optical Instrumentation Engineers, 2000.

[48] C.R. Farrar, S. Doebling, P.J. Cornwell, and E.G. Straser. Variability of modal parameters measured on the alamosa canyon bridge. Technical report, Los Alamos National Lab, NM (United States), 1996.

[49] C.Y. Kim, D.S. Jung, N.S. Kim, S.D. Kwon, and M.Q. Feng. Effect of vehicle weight on natural frequencies of bridges measured from traffic-induced vibration. *Earthquake Engineering and Engineering Vibration*, 2(1):109–115, 2003.

[50] E. Caetano, A. Cunha, V. Gattulli, and M. Lepidi. Cable–deck dynamic interactions at the international guadiana bridge: On-site measurements and finite element modelling. *Structural Control and Health Monitoring*, 15(3): 237–264, 2008.

[51] D. Siringoringo and Y. Fujino. Observed along-wind wibration of a suspension bridge tower. *Journal of Wind Engineering and Industrial Aerodynamics*, 103: 107–121, 2012.

[52] W.F. Chen and L. Duan. *Bridge Engineering Handbook: Construction and Maintenance*. CRC Press, 2014.

[53] C.C. Chen, W.H. Wu, and C.Y. Liu. Effects of temperature variation on cable forces of an extradosed bridge. In *6th European Workshop on Structural Health Monitoring*, pages 1–8, 2012.

[54] X.X. Li, W.X. Ren, and K.M. Bi. Fbg force-testing ring for bridge cable force monitoring and temperature compensation. *Sensors and Actuators A: Physical*, 223:105–113, 2015.

[55] C. Machelski and M. Hildebrand. Efficiency of monitoring system of a cable-stayed bridge for investigation of live loads and pier settlements. *Journal of Civil Structural Health Monitoring*, 5(1):1–9, 2015.

[56] P. Krishna, A.S. Arya, and T.P. Agrawal. Effect of cable stiffness on cable-stayed bridges. *Journal of Structural Engineering*, 111(9):2008–2020, 1985.

[57] A. Bosdogianni and D. Olivari. Wind-and rain-induced oscillations of cables of stayed bridges. *Journal of Wind Engineering and Industrial Aerodynamics*, 64(2-3):171–185, 1996.

[58] M. Matsumoto, Y. Daito, T. Kanamura, Y. Shigemura, S. Sakuma, and H. Ishizaki. Wind-induced vibration of cables of cable-stayed bridges. *Journal of Wind Engineering and Industrial Aerodynamics*, 74:1015–1027, 1998.

[59] K. Worden and G. Manson. The application of machine learning to structural health monitoring. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1851):515–537, 2007.

[60] H. Sohn, K. Worden, and C.R. Farrar. Statistical damage classification under changing environmental and operational conditions. *Journal of Intelligent Material Systems and Structures*, 13(9):561–574, 2002.

[61] N. Dervilis, E.J. Cross, R.J. Barthorpe, and K. Worden. Robust methods of inclusive outlier analysis for structural health monitoring. *Journal of Sound and Vibration*, 333(20):5181–5195, 2014.

[62] H. Sohn and C.R. Hoon. A statistical pattern recognition paradigm for vibration-based structural health monitoring. Technical report, Los Alamos National Lababoratory, NM (US), 2000.

[63] E.J. Cross, G. Manson, K. Worden, and S.G. Pierce. Features for damage

detection with insensitivity to environmental and operational variations. *Proc. R. Soc. A*, pages 1–31, 2012.

[64] A. Deraemaeker, A. Preumont, and J. Kullaa. Modeling and removal of environmental effects for vibration based shm using spatial filtering and factor analysis. In *Proceedings of IMAC XXIV*, 2006.

[65] R. Ruotolo and C. Surace. Using svd to detect damage in structures with different operational conditions. *Journal of Sound and Vibration*, 226(3):425–439, 1999.

[66] C. Liu, J.B. Harley, Y. Ying, I.J. Oppenheim, M. Bergés, and J.H.G̃arrett D.W.G̃reve. Singular value decomposition for novelty detection in ultrasonic pipe monitoring. In *Sensors and Smart Structures Technologies for Civil, Mechanical and Aerospace Systems 2013*, volume 8692, page 86921R. International Society for Optics and Photonics, 2013.

[67] K. Worden, G. Manson, and D. Allman. Experimental validation of a structural health monitoring methodology: Part i. novelty detection on a laboratory structure. *Journal of Sound and Vibration*, 259(2):323–343, 2003.

[68] G. Manson, K. Worden, and D. Allman. Experimental validation of a structural health monitoring methodology: Part ii. novelty detection on a gnat aircraft. *Journal of Sound and Vibration*, 259(2):345–363, 2003.

[69] G. Manson. Identifying damage sensitive, environment insensitive features for damage detection. In *Proceedings of the Third International Conference on Identification in Engineering Systems*, pages 187–197, 2002.

[70] G. Lederman, Z. Wang, J. Bielak, H. Noh, J.H. Garrett, S. Chen, J. Kovacevic, F. Cerda, and P. Rizzo. Damage quantification and localization algorithms for indirect shm of bridges. In *Proc. Int. Conf. Bridge Maint., Safety Manag., Shanghai, China*, 2014.

[71] A.M. Yan, G. Kerschen, P.D̃e Boe, and J.C. Golinval. Structural damage diagnosis under varying environmental conditionspart i: a linear analysis. *Mechanical Systems and Signal Processing*, 19(4):847–864, 2005.

[72] A.M. Yan, G. Kerschen, P.D̃e Boe, and J.C. Golinval. Structural damage diagnosis under varying environmental conditionspart ii: Local pca for non-

linear cases. *Mechanical Systems and Signal Processing*, 19(4):865–880, 2005.

[73] A.D.F. Santos, M.F.M. Silva, C.S. Sales, J.C.W.A. Costa, and E. Figueiredo. Applicability of linear and nonlinear principal component analysis for damage detection. In *Instrumentation and Measurement Technology Conference (I2MTC), 2015 IEEE International*, pages 869–874. IEEE, 2015.

[74] S. Johansen. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3):231–254, 1988.

[75] S. Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: Journal of the Econometric Society*, pages 1551–1580, 1991.

[76] E.J. Cross, K. Worden, and Q. Chen. Cointegration: A novel approach for the removal of environmental trends in structural health monitoring data. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 467:2133:2712–2732, 2011.

[77] P.B. Dao and W.J. Staszewski. Lamb wave based structural damage detection using cointegration and fractal signal processing. *Mechanical Systems and Signal Processing*, 49(1):285–301, 2014.

[78] P.B. Dao and W.J. Staszewski. Data normalisation for lamb wave-based damage detection using cointegration: A case study with single-and multiple-temperature trends. *Journal of intelligent Material Systems and Structures*, 25(7):845–857, 2014.

[79] F. Li, Z. Wang, and G. Liu. Towards an error correction model for dam monitoring data analysis based on cointegration theory. *Structural Safety*, 43: 12–20, 2013.

[80] C. Bao, H. Hao, and Z. Li. Vibration-based structural health monitoring of offshore pipelines: Numerical and experimental study. *Structural Control and Health Monitoring*, 20(5):769–788, 2013.

[81] E.J. Cross and K. Worden. Approaches to nonlinear cointegration with a view towards applications in shm. In *Journal of Physics: Conference Series*, volume 305, pages 12–69. IOP Publishing, 2011.

[82] H. Shi, K. Worden, and E.J. Cross. A nonlinear cointegration approach with

applications to structural health monitoring. In *Journal of Physics: Conference Series*, volume 744:1, pages 12–25. IOP Publishing, 2016.

[83] DIN Standard. Nen1601: Mechanical properties of fasteners 7984. Technical report, Technical Committee of the European Standards Organization CEN, 2000.

[84] LMS Instruments. Lms scadas iii data acquisition front-end. Technical report, Siemens, 2007.

[85] D.J. Ewins. *Modal Testing: Theory and Practice*. Research Studies Press Letchworth, 1984.

[86] J.K. Anderson. Tamar bridge. *Proceedings of the Institution of Civil Engineers*, 31(4):337–360, 1965.

[87] R.F.J. Gill. Tamar suspension bridge-strengthening and capacity enhancement. In *Bridge Modification 2: Stronger and Safer Bridges: Proceedings of the International Conference Organized by the Institution of Civil Engineers and Held in London on 7 November 1996*, volume 2, pages 93–104. Thomas Telford, 1997.

[88] J.M.W. Brownjohn and P. Carden. Real-time operation modal analysis of tamar bridge. In *26th International Modal Analysis Conference (IMAC XXVI)*, 2008.

[89] G.E.P. Box and N. Draper. *Empirical Model-Building and Response Surface*. John Wiley and Sons Inc., 1986.

[90] H. Tabatabai. *Inspection and Maintenance of Bridge Stay Cable Systems: a Synthesis of Highway Practice*. Transportation Research Board, 2005.

[91] Y.L. Xu. *Wind and Structural Health Monitoring*. Wiley Online Library, 2013.

[92] A. Bosdogianni and D. Olivari. Wind-and rain-induced oscillations of cables of stayed bridges. *Journal of Wind Engineering and Industrial Aerodynamics*, 64(2-3):171–185, 1996.

[93] M. Matsumoto, Y. Daito, T. Kanamura, Y. Shigemura, S. Sakuma, and H. Ishizaki. Wind-induced vibration of cables of cable-stayed bridges. *Journal of Wind Engineering and Industrial Aerodynamics*, 74:1015–1027, 1998.

[94] D.A. Dickey and W.A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–431, 1979.

[95] D.A. Dickey and W.A. Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 49:1057–1072, 1981.

[96] P.C. Phillips and P. Perron. Testing for a unit root in time series regression. *Biometrika*, 72:335–346, 1988.

[97] D. Kwiatkowski, P.C. Phillips, and P.C. Schmidt. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54:159–178, 1992.

[98] J.G. MacKinnon. *Critical Values for Cointegration Tests*. Oxford University Press, 1991.

[99] C.R. Nelson and C.R. Prosser. Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, 10:139–162, 1982.

[100] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.

[101] G.W. Schwert. Margin requirements and stock volatility. *Regulatory Reform of Stock and Futures Markets*, pages 55–66, 1989.

[102] R.A. Yaffee and M. McGee. *An Introduction to Time Series Analysis and Forecasting: with Applications of SAS and SPSS*. Academic Press, 2000.

[103] S.G. Mattson and S.M. Pandit. Statistical moments of autoregressive model residuals for damage localisation. *Mechanical Systems and Signal Processing*, 20(3):627–645, 2006.

[104] C.S. Li, W.J. Ko, H.T. Lin, and R.J. Shyu. Vector autoregressive modal analysis with application to ship structures. *Journal of Sound and Vibration*, 167(1):1–15, 1993.

[105] K. Worden, E.J. Cross, I. Antoniadou, and A. Kyprianou. A multiresolu-

tion approach to cointegration for enhanced shm of structures under varying conditions–an exploratory study. *Mechanical Systems and Signal Processing*, 47(1):243–262, 2014.

[106] K. Worden and G.R. Tomlinson. *Nonlinearity in Structural Dynamics: Detection, Identification and Modelling*. CRC Press, 2000.

[107] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. CRC Press, 1986.

[108] D. Scott. *Multivariate Density Estimation: Theory, Practice and Visualization*. John Wiley and Sons, 1992.

[109] E.J. Cross, G. Manson, K. Worden, and S.G. Pierce. Features for damage detection with insensitivity to environmental and operational variations. *Proceedings of the Royal Society - Series A*, 468:4098–4122, 2012.

[110] D.C. Montgomery. *Introduction to Statistical Quality Control*. John Wiley and Sons, 2009.

[111] H.L. Hurd and A. Miamee. *Periodically Correlated Random Sequences: Spectral Theory and Practice*. John Wiley and Sons, 2007.

[112] D. DeJong, J. Nankervis, N. Savin, and C. Whiteman. Integration versus trend-stationarity in macroeconomic time series. Technical Report Working Paper, University of Iowa. Department of Economics, 1988.

[113] C.S. Hakkio and M. Rush. Cointegration: How short is the long run? *Journal of International Money and Finance*, 10:571–581, 1991.

[114] J. Otero and J. Smith. Testing for cointegration: Power versus frequency of observation-further Monte Carlo results. *Economics Letters*, 67:5–9, 2000.

[115] R.J. Shiller and P. Perron. Testing the random walk hypothesis: Power versus frequency of observation. Technical report, National Bureau of Economic Research, Cambridge, 1985.

[116] H. van den Berg and S.C. Jayaretti. A novel test of the monetary approach using black market exchange rates and the Johansen-Juselius cointegration method. *Economics Letters*, 41:413–418, 1993.

[117] M.A. Hooker. Testing for cointegration: Power versus frequency of observa-

tion. *Economics Letters*, 41:359–362, 1993.

[118] K. Lahiri and N. Mamingi. Testing for cointegration: Power versus frequency of observation – another view. *Economics Letters*, 49:121–124, 1995.

[119] R. Perman. Cointegration: An introduction to the literature. *Journal of Economic Studies*, 18:3–30, 1993.

[120] A.C. Palmer. *Dimensional Analysis and Intelligent Experimentation*. World Scientific Publishing Co., 2008.

[121] H. Cramer. *Mathematical Methods of Statistics*. Princeton University Press, 1964.

[122] C.R. Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Breakthroughs in Statistics*, 1:235–247, 1992.

[123] M. Frigo and S.G. Johnson. Fftw: An adaptive software architecture for the fft. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 3, pages 1381–1384. IEEE, 1998.

[124] K. Worden, E.J. Cross, and E. Barton. Damage detection on the NPL footbridge under changing environmental conditions. In *Proceedings of* 6^{th} *European Workshop on Structural Health Monitoring*, pages 1–8, 2012.

[125] G. Liu, J. Zhang, X. Yao, and P. Qin. Damage detection from continuous long-term static response using cointegration and mewma control chart. *Internal Conference of Sustainable Development of Critical Infrastructure*, pages 421–429, 2014.

[126] E.J. Cross and K. Worden. Cointegration and why it works for shm. In *Journal of Physics: Conference Series*, volume 382, pages 12–46. IOP Publishing, 2012.

[127] G. Liu, J. Zhang, X. Yao, and P. Qin. Damage detection from continuous long-term static response using cointegration and mewma control chart. In *International Conference of Sustainable Development of Critical Infrastructure*, pages 421–429, 2014.

[128] P.B. Dao, W.J. Staszewski, and A. Klepka. Stationarity-based approach for the selection of lag length in cointegration analysis used for structural damage detection. *Computer-Aided Civil and Infrastructure Engineering*, 32(2):138–

153, 2017.

[129] K. Worden, I. Iakovidis, and E.J. Cross. On stationarity and the interpretation of the ADF statistic. In *In Proceedings of the 36th International Modal Analysis Conference, Orlando, FL*. IEEE, 2018.

[130] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic press, 1999.

[131] Y. Meyer. Wavelets-algorithms and applications. *Wavelets-Algorithms and Applications Society for Industrial and Applied Mathematics Translation., 142 p.*, 1993.

[132] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, and G.M. Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 2015.

[133] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. Numerical recipes in c++. *The Art of Scientific Computing*, 1992.

[134] C. Chatfield. *The Analysis of Time Series: An Introduction*. Chapman and Hall, fifth edition, 1996.

[135] Roadside Features London. Design manual for roads and bridges: Volume 3 section 4. Technical report, U.K Highway Agency, 2001.

[136] K. Juselius. *The Cointegrated VAR model: Methodology and Applications*. Oxford University Press, 2006.

[137] D.M. Frangopol, M.J. Kallen, and J.M. van Noortwijk. Probabilistic models for life-cycle performance of deteriorating structures: Review and future directions. *Progress in Structural Engineering and Materials*, 6(4):197–212, 2004.

[138] R. Schöbi and E.N Chatzi. Maintenance planning using continuous-state partially observable markov decision processes and non-linear action models. *Structure and Infrastructure Engineering*, 12(8):977–994, 2016.