The University Of Sheffield.

Access to Electronic Thesis

Author:	Elizabeth Cross
Thesis title:	On Structural Health Monitoring in Changing Environmental and Operational Conditions
Qualification:	PhD

On Structural Health Monitoring in Changing Environmental and Operational Conditions

A Thesis submitted to the University of Sheffield

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

by

E. J. Cross

Department of Mechanical Engineering

University of Sheffield

May 2012

Acknowledgments

First and foremost sincere thanks must go to my supervisor Professor Keith Worden, without whom this work wouldn't have been possible. I'd like to thank him for his guidance, help and continuing support. On this front thanks must also go to Dr Graeme Manson, Dr Charles Farrar and Professor James Brownjohn for their support and guidance.

A large part of this work has focused on data from the Tamar Bridge, none of the analysis in this thesis would have been possible without the hard work and dedication of Dr Ki Young Koo and other members of the Vibration Engineering Section. I'd like to thank Ki especially for his patience and kind help in accessing the data from the monitoring campaign. Other thanks must also go to the EU DAMASCOS consortium, in particular to Dr Gareth Pierce who is responsible for the experimental study used in Chapter 8. I'd also like to thank Villa Lämsä for providing me with data from the benchmark study on the Z24 highway bridge.

I would like to thank everyone in the Dynamics Research Group, it has been a privilege to work with such a lovely group of people, many who have helped me both directly and indirectly with my studies. Lastly, but by no means least, I would like to thank my family and friends for their support throughout my studies. Particular thanks to Oliver for his support and for putting up with me working evenings and weekends over the last six months!

Abstract

Structural Health Monitoring (SHM) is the monitoring of any type of structure for the express purpose of determining its condition and future lifespan and if, when and where any reparative action is needed. A focus of the work in this thesis is SHM for long-span bridges and particularly the effects of environmental and operational conditions on a monitoring campaign. There is currently a trend for heavily instrumenting civil structures with large sensor networks that continually collect terabytes of data. However, these large data sets are often redundantly stored and not used for anything. One of the principal aims in the thesis is to exploit such monitoring data for the development of diagnostic tools for structural condition assessment.

The first part of the thesis concerns formulating a baseline for the Tamar Bridge that represents the normal undamaged condition of the structure. To do this a large amount of analysis was needed in order to understand how different structural measurements are interrelated and how the bridge responds to normal environmental and operational conditions. Particular attention was paid to measurements that can be sensitive to structural degradation (such as modal properties). Often simple causal relationships were found between monitored variables, and response surface models were formulated that could predict selected variables with good accuracy given measurement of operational and environmental conditions, such as air temperature, traffic loading and wind profile. The predictive models developed are intended to be used as diagnostic tools, for example, a departure from the normal condition of the bridge will bring about a significant increase in prediction error, which may be monitored as a system alarm.

The second part of the thesis directly concerns how the influence of environmental and operational variation on features sensitive to damage can be lessened or removed

without measurement of these conditions themselves. This is a very important issue in SHM, as often the effects of fluctuating environmental and operational conditions can mask any indication of damage to a structure that may be evident in structural response. In the thesis a solution to the problem based on the econometric theory of cointegration is introduced. Application of this theory is found to be ideally suited to remove unwanted environmental and operational trends from SHM data, and forms an exceedingly promising contribution to the development of SHM technology.

Publications

Author publications to date

Journal papers

K. Worden, E.J. Cross and A. Kyprianou, 2012, A multiresolution approach to cointegration for time-varying systems with application to Structural Health Monitoring. Submitted to Mechanical Systems and Signal Processing.

E.J. Cross, K. Worden, G. Manson and G. Pierce, 2012, Features for damage detection with insensitivity to environmental and operational variations. Submitted to the Proceedings of the Royal Society A.

E.J. Cross, K.Y. Koo, J.M.W. Brownjohn and K. Worden, 2012, Long-term monitoring and analysis of the Tamar Bridge. Submitted to Mechanical Systems and Signal Processing.

E.J. Cross, K. Worden and Q. Chen, 2011, Cointegration; A novel approach for the removal of environmental trends in structural health monitoring data. Proceedings of the Royal Society A 467 2712-2732 doi:10.1098/rspa.2011.0023

E.J. Cross and K. Worden, 2011, Approximation of the Duffing oscillator frequency response function using the FPK equation Journal of Sound and Vibration 330 (2011) 743756.

Book chapters

E.J. Cross and K. Worden, 2011. An introduction to Structural Health Monitoring in the context of civil infrastructure. To appear in Health Assessment of Engineering Structures: Bridges and Other Infrastructure. Ed. A. Haldar. World Scientific.

Reviewed conference papers

E.J. Cross, K.Y. Koo, J.M.W. Brownjohn and K. Worden, 2012, Filtering environmental loading effects to enhance novelty detection on cable-supported bridge performance, 6th International conference on Bridge Maintenance, Safety and Management, Lake Como, Italy. In press.

E.J. Cross and K. Worden, 2011, Approaches to Nonlinear Cointegration with a View Towards Applications in SHM. Proceedings of the 9th International Conference on Damage Assessment of Structures, Oxford.

K. Worden, E.J. Cross and A. Kyprianou, 2011, Cointegration and Nonstationarity in the Context of Multi-resolution Analysis. Proceedings of the 9th International Conference on Damage Assessment of Structures, Oxford.

E.J. Cross, K. Worden, R.J. Sturgeon and Q. Chen, 2010, A tutorial on Cointegration for engineers - a tool for non-stationary time series analysis. Proceedings of the 10th International Conference on Recent Advances in Structural Dynamics.

E.J. Cross and K. Worden, 2009, Approximation of the Duffing oscillator frequency response function using the FPK equation, 7th International Conference on Modern Practice in Stress and Vibration Analysis.

Conference papers

E.J. Cross, K. Worden, P. Sartor, P. Southern, 2012, Prediction of landing gear loads using machine learning techniques, Proceedings of the 5th European Workshop on Structural Health Monitoring. In press.

E.J. Cross, K. Worden, E. Barton, 2012, Damage detection of the NPL footbridge under changing environmental conditions, Proceedings of the 5th European Work-

shop on Structural Health Monitoring. In press.

E. Figueiredo, R. Westgate, E.J. Cross, J.MW. Brownjohn, K. Worden, 2012, Applicability of Markov Chain Monte Carlo for damage detection on data from the Tamar suspension bridge and Z24, Proceedings of the 5th European Workshop on Structural Health Monitoring. In press.

E.J. Cross, K. Worden, P. Sartor, P. Southern, 2012, (Extended Abstract) Prediction of landing gear loads using machine learning techniques, SAFRAN Pole de Competence du Monitoring (Safran internal conference on health monitoring), Paris. March 2012. In press.

R.J. Barthorpe, E.J. Cross, E. Papatheou, K. Worden, 2011, Some recent developments in SHM, Proceedings of 2nd International Workshop on Smart Diagnostics of structures.

E.J. Cross, K.Y. Koo, J.M.W. Brownjohn and K. Worden, 2011, Cointegration and SHM of bridges, Proceedings of 8th International Workshop on Structural Health Monitoring, Stanford.

E.J. Cross, and K. Worden, 2011, Nonlinear cointegration as a combinatorial optimisation problem for damage detection, Proceedings of 8th International Workshop on Structural Health Monitoring, Stanford.

E.J. Cross, K. Worden and G. Manson, 2011, Cointegration for the removal of environmental trends in SHM data, Proceedings of the 2011 International Conference on Structural Engineering Dynamics, Tavira, Portugal.

E.J. Cross, K.Y. Koo, J.M.W. Brownjohn and K. Worden, 2010, Long-term monitoring and data analysis of the Tamar Bridge, Proceedings of ISMA, International Conference on Noise and Vibration Engineering.

E.J. Cross and K. Worden, 2010, An approach to nonlinear cointegration with application to structural health monitoring data. Proceedings of the 5th European Workshop on Structural Health Monitoring.

K.Y. Koo, E.J. Cross, K. Worden, J.M.W. Brownjohn, D. List, R. Cole and T. Wood, 2010 Long-term structural health monitoring for Tamar suspension bridge. Proceedings of the 5th International Conference on Bridge Maintenance, Safety and Management.

E.J. Cross, K. Worden, K. Y. Koo and J.M.W. Brownjohn, 2010, Modelling environmental effects on the dynamic characteristics of the Tamar suspension bridge Proceedings of IMAC XXVIII.

D. Harvey, E.J. Cross, R. Silva, C. Farrar and M. Bement, 2010 Input Estimation from Measured Structural Response Proceedings of IMAC XXVIII

R. Sturgeon, E.J. Cross, K. Worden, J.M.W. Brownjohn, K.Y. Koo and Q. Chen, 2009, Elimination of long-term trends in SHM data using cointegration, Proceedings of 7th International Workshop on Structural Health Monitoring, Stanford.

J.M.W. Brownjohn, K. Worden, E.J. Cross, D. List, R. Cole and T. Wood, 2009, Thermal effects on performance on Tamar Bridge 4th International Conference on Structural Health Monitoring on Intelligent Infrastructure (SHMII-4).

Table of Contents

Chapter 1

An Introduction to Structural Health Monitoring

Structural Health Monitoring (SHM) is, in short, any automated monitoring practice that seeks to assess the condition or health of a structure. Its beginnings as an area of interest to engineers can be traced back as far as the time when tap-testing for fault detection became common, although the field didn't really become established in research communities until the 1980s, when much interest was generated in the structural condition of oil rigs, and later in aerospace structures and their health [1]. Nowadays, SHM is a popular and still growing research field, which is more and more becoming a focus of the civil infrastructure community.

This chapter aims to provide a general overview of SHM. The potential benefits of a comprehensive SHM campaign will be outlined, before discussing the common issues arising when attempting to create/implement such a system.

1.1 The aims of SHM

The ideal that SHM strives towards is to be able to monitor a structure in such a way that any damage introduced, or any growth of inherent faults, would be immediately detectable. Further to this, the aim is that, after detection, any fault could be located and its severity inferred so that decisions can be easily made as to actions necessary (e.g. immediate halt to use of the structure, immediate repair, etc.).

These global objectives for SHM have been well formalised in Rytter's hierarchy [2], which classifies these aims into 'levels' of increasing difficulty. These levels can be summarised as follows:

- Level One - Detection: automatic detection of damage to the system/structure
- Level Two - Localisation: automatic determination of where damage has occurred in the system/structure
- Level Three - Quantification: automatic assessment of damage type and severity
- Level Four - Prognosis: prediction of the remaining useful life in the structure or specific component.

Although a number of changes and additions to this hierarchy have been suggested in the literature [3], these levels continue to provide a good basic summary of the fundamental aims of any SHM system, although, of course, without reference to how one might go about them.

1.1.1 Potential benefits of SHM

With regards to the advantages of any system able to fulfil the global objectives of SHM, the first and most obvious benefit is increased human safety, indeed, unsurprisingly much of the research in this field has been motivated by disasters such as bridge collapses and aeroplane crashes, where many lives have been lost (see [4], for example, for more details). Even at the lowest level of SHM - a detection of damage or degradation of structural condition could be hugely beneficial if used to provide an early warning that a structure may be unsafe. Additionally, with an automated detection system using non-visual assessment, any areas of a structure that are difficult or impossible to access, that otherwise may have been neglected in a visual inspection, can be assessed.

Other arising benefits will come from the regime change that comprehensive SHM could bring about. Currently civil and aerospace structures undergo routine inspection and maintenance at specific time intervals, for example bridge inspections in the USA are scheduled every two years [4], and commercial aircraft undergo a thorough inspection after a given number of flight hours/cycles. A time based approach to

management of structural assets such as this, firstly, has the implication that any unexpected faults occurring in between scheduled inspections may go un-noted and cause danger to life, or cause unnecessary stress on other structural components. Inversely, the set time scales for inspections may be overly conservative; if a structure continues to be in good health, the costs of thorough inspections could essentially have been saved. In the case of routine maintenance, where structural components may be replaced even if they are in excellent condition, the economic impact may be even greater. SHM has the ability to address both sides of this issue, as monitoring has the potential to become continual, and maintenance and repair could become condition-based. A switch to condition-based maintenance could also reduce the downtime a structure may undergo for routine and emergency maintenance, which, in turn, would be of economic and environmental benefit.

1.1.2 Disambiguation; SHM and similar areas of research

There are a number of research fields very closely related to SHM, which can be seen, depending on one's point of view, as either overlapping with, or perhaps even encompassed within SHM. For disambiguation, it seems sensible to discuss them shortly here, a more detailed discussion can be found in [3]. Although tap-testing was mentioned above as marking the beginnings of SHM, the example really belongs to the field of *Non-Destructive Testing* (NDT) or *Non-Destructive Evaluation* (NDE). NDE or NDT concerns the assessment of a structure or component's health through (offline) non-damaging procedures. Examples of tools commonly used for NDE are X-ray, electron microscopy, measurement of acoustic emissions and full scale vibration tests. Although all of these techniques may be used for SHM purposes, NDEs currently most commonly occur as one-off planned events, often applied to a small area of a structure where damage is suspected to have occurred. This is a different approach to SHM where monitoring aims to be continuous and global. In the future, it is likely that NDE inspection will form the basis for distinguishing between health and performance anomalies for civil infrastructure where this cannot be accomplished automatically. It is therefore true to say that NDE may be incorporated as part of an SHM system, but not vice-versa.

Another field related to SHM is *Condition Monitoring*. Condition monitoring largely concerns the health of rotating machinery; it has seen many successes and has been, in some areas, accepted as part of every day practice by industry [5]. Its commercial

success relative to SHM can be attributed to a number of factors that simplify the monitoring process: the machinery operate in a controlled environment, it is usually easy to access and is typically on a small scale. Importantly, rotating machinery have been found to exhibit well defined dynamic responses for particular fault categories, which makes fault detection and identification a more easily attainable goal than it perhaps is for SHM [6].

It should be noted here that one thing SHM is not is *monitoring*. Simple collection of data does not constitute SHM, however complex and comprehensive the sensor network is. It is true to say that the current trend for civil infrastructure is to configure as many sensors as possible on a single structure, indeed state of the art monitoring campaigns nowadays plan to employ in the number of thousands of sensors [7]. If such sensor networks stream data continuously, the amount of data stored is huge; however, this is of no value unless performance or health knowledge is extracted, and this is, unfortunately, far from the norm.

1.2 SHM in practice

The many benefits of SHM come hand in hand with many challenges that must be overcome. The fundamental problem at the heart of SHM is that of how a measure of structural condition can be gained from an automated process. No sensor can measure damage directly [8] and so this fundamental problem breaks down into a number of separate issues; what can be measured that correlates to damage, how to measure it and, importantly, how to use the raw measurements to make inferences and decisions about structural condition. Finally, before an SHM system can be relied upon, it must be proven to work, and issues such as how to cope with sensor failures, for example, must be addressed. In the following, each of these issues is addressed separately. As an introduction to SHM, this chapter certainly does not aim be exhaustive, instead for a comprehensive review of practices in SHM readers are referred to [1, 9, 10].

1.2.1 Instrumentation for SHM

The first questions asked of any planned implementation of SHM, as alluded to in the paragraph above, are what is the most useful thing that can be measured for

structural assessment purposes, and how can one best measure it? This is part of the *operational evaluation stage* in the four-stage implementation of SHM discussed in [8].

The most common measurements sought by far in SHM are of the dynamic response of a structure. Dynamic responses contain information about the mass, stiffness and damping of a structure, all of which could feasibly change with the onset or progression of damage, hence the interest in these measurements. Measurement of acceleration is perhaps the most common in SHM and is used for structures and components of all sizes. Measurements of strain are also very common, but are currently most regularly used for small scale structures/components and for composite materials. Strain measurements also find good employment in usage monitoring, where load cycles are counted [11, 12].

A dominating issue when considering the dynamic response of a structure is that many common practices developed for SHM, such as modal analysis, for example, rely on knowledge of the excitation source. Structures in the real world experience excitation from operational conditions which, in practice, cannot be measured, such as the excitation experienced by a bridge from traffic passing over it. In these circumstances, an assumption as to the properties of an excitation source must be made [13, 14]. In other circumstances, artificial excitation that can be measured is introduced, with a hammer, a shaker or an electrical pulse, for example [15]. As knowledge of an excitation source is desirable, much research effort is currently focused on sensor systems that can provide their own measurable excitation source [16]. One particular area of interest where this is relevant is the use of higher frequency guided waves for damage assessment [17]. Guided waves have mainly been used for the detection of damage in plates and pipes, and are therefore arguably of most interest to the aerospace and process industries, although recently a growing interest in SHM for wind turbines has led research into guided waves in that direction as well [18].

Some other measurements that have been found useful for SHM are measurement of acoustic emissions which occur with damage initiation and progression (for example, this has been investigated for monitoring individual cable snapping in the main cables of suspension bridges [19]), and measurement of electrical impedance, which has been found to correspond to the mass, stiffness and damping properties of a structure [20].

One of the largest concerns in SHM is *how* to obtain the measurements that will be most useful; research into suitable instrumentation for SHM probably attracts more interest in the community than any other single topic. A complete SHM monitoring system will more than likely require a large number of sensors of differing types in order to monitor all components and be able to identify different damage scenarios. If suitable sensors are available, a number of important questions must be addressed for a useful (or ideally optimal) monitoring system, these include; where sensors are best placed, how many are needed, how they can be powered, where an excitation source will come from and how data will be transmitted.

Most monitoring systems in place now, especially on civil infrastructure, use wired sensors, both for a power source and for data transferral. For numerous reasons, however, using wired sensors for monitoring structures outside the laboratory proves to be difficult; the amount of wiring necessary to instrument whole structures quickly becomes infeasible for large scale structures, further to this, the addition of a large amount of wires is often very unappealing to operators (due to, for example, the extra weight or the increased lightning conductivity that a network of wires may introduce). For these reasons, wireless sensing has become a popular topic of research over the last few years, and is seen by some within the community as the future for SHM [4].

Sensing wirelessly naturally introduces a new set of problems to address, the most pressing of which are how to power the sensor and data telemetry. To overcome some of the powering and telemetry issues, it is thought that some on-board processing of data at the sensor before transmission would be of great benefit [21]. Self-powering sensors (energy harvesting) are also an emerging field of interest [22], as well as other novel techniques for power and data transferral, such as the use of remotely controlled vehicles [23].

An additional monitoring issue that has more recently arisen, concerns the management of large amounts of data collected by a monitoring system. Nowadays, many monitoring campaigns, especially for civil infrastructure, involve dense sensor arrays from which terabytes of data are collected [7, 24]. Consequently the development of systems for storage of, and importantly, access to large amounts of data has become important. Aside from all the necessary signal processing inherent in monitoring campaigns, efficient management of data is essential for a successful SHM system.

1.2.2 Assessment of structural condition from measurements

The question of how to infer structural condition from different measurements is at the heart of SHM. Once measurements with some correlation with damage have been obtained, the process for arriving at a judgement on structural condition can be divided between two major tasks. The manipulation of measurements from a structure in order to create a usable variable that can give an indication of structural condition is often named *feature extraction*, this forms the first major task. The use of extracted features to make decisions (such as damaged or not damaged) is the second major challenge that must be faced, and one which has been identified by many as a problem in statistical pattern recognition [25].

Feature Extraction

As previously stated, no sensor is available that can measure any type of damage directly, instead measurements can at best be correlated with the damage type one is interested in. A raw measurement is also unlikely to be directly useful for damage detection and assessment; on a practical level this is simply often because a raw measurement provides too much data/information than is feasible to work with (high dimensionality), and can also often be difficult to interpret. Often the pattern recognition techniques that are used to infer structural condition from data can only work well in a low dimension. For these reasons, feature extraction is used to create useful metrics from raw measurements which are often of a lower dimension than the raw data. Simple features that can be extracted from raw measurements include, for example, statistics from a signal such as the mean and variance. Where the aim of feature extraction is purely to reduce the dimension of measurements, approaches such as principal component analysis can be used, which acts to transform data in such a way that redundancy is simple to identify and remove. Many other feature extraction methods rely on working in the frequency domain of a signal provided through use of a Fourier transform [26]. In the frequency domain, the spectrum of a signal is particularly useful as it is often of low dimension and can be easily interpreted.

Damage sensitive features in the form of modal properties extracted from acceleration measurements are probably the most frequently occurring features used in SHM. These include, but are not limited to, natural frequencies, mode shapes, and

mode shape curvatures. The modal approach is attractive due to the interpretability of the features and additionally the fact that typically only a low number of sensors is required in order to extract these features. An important point to consider, however, is that modal features are global indicators for structural condition, meaning that any inference on condition applies to the whole structure. Because of this fact, and the fact that modal analysis can be carried out with a small number of sensors and therefore with relatively little trouble, these approaches have been found useful for the provision of a good one-off general assessment of a structure [27]. The well known disadvantage, however, associated with modal properties is that they have been found to be insensitive to structural degradation on a local scale [1].

Other approaches to feature extraction seek to fit measurements to mathematical models or functions (other than the Fourier transform) and use parameters from these models as features. A common example of this is fitting an ARMA type model to measurement data and using the model coefficients as features [28]. Another way in which fitting models to raw data can be used as a feature extraction methodology, is to use the residual error of a predictive model as a feature [28]. For more details on feature extraction and selection see [26].

It is useful to note that where any measurement data are fitted to a model in this way, (including for modal analysis), this may be referred to as *system identification*, a topic on which a wealth of research has been conducted (see for example [29]). Where output-only modal analysis is concerned, a growing area of interest is in the use of stochastic subspace identification [30]. System identification for nonstationary random vibration is also a growing area of interest (see [31, 32]).

Pattern Recognition for inference on structural condition from features

Once a particular feature has been extracted from raw measurements, a decision process is needed to infer structural condition from this feature. As previously mentioned, this is essentially a problem in pattern recognition, where a feature will be classified according to whether it has arisen from a damaged or undamaged structure. At higher levels of SHM, a feature will be classified as to the location, type and severity of the damage, if present. Pattern recognition in this form relies on one of two different approaches, the first is a supervised learning approach, the second relies on novelty detection [3].

Supervised learning for SHM is any procedure for the classification of a feature which is informed with data from all states of interest. In terms of the lowest level of Rytter's hierarchy this simply means that data must be available from the damaged and undamaged state of a structure. Techniques that use supervised learning can include all algorithms capable of classification, such as neural networks, support vector machines and Gaussian processes [26]. A supervised learning approach is considered to be necessary where identification of different damage types and locations is required [8]. Unfortunately, data collected from the damaged state of a structure is rarely available (let alone data from multiple damage scenarios), as, naturally, introducing damage to a structure to inform an SHM decision process is not an acceptable option. In cases where supervised learning is necessary, future advancements in the field may rely on physics based, or high fidelity models that can accurately simulate the response of a structure in a damaged state. It is possible that physical proxies for damage can also be found [33].

Indeed, research into the use of high fidelity models (i.e. finite element) for SHM is popular. Model updating is used as a tool for inference on structural condition which, in simple terms, is the use of measurements from a real structure to update a physics based model, which can then be used, via an inverse problem, to predict the current state of the structure. Research into model updating approaches is reviewed comprehensively in [34]. Such approaches are philosophically different to the data-based approaches that will be employed in this thesis.

Novelty detection algorithms are required when data from the damaged condition of a structure are not available (which is most often the case). A decision process reliant on novelty detection will aim to define a baseline, with data from the undamaged condition of a structure, that represents the normal response of the structure in its undamaged condition. An abnormal response is then detected by any significant departure from this baseline. Common examples of techniques for novelty detection include outlier analysis and the use of statistical process control charts, both of which rely on the selection of a threshold which, if crossed, signifies that a structure is not responding in a normal way [35]. A disadvantage of novelty detection approaches is the unavoidable fact that an indication that a structure has departed from its normal condition is uninformative as to what may have caused this departure. In this case, further investigation after a novel response has been detected will be necessary in order to eliminate the possibility that change has occurred for a benign reason e.g. because of a temperature change. However, in the author's opinion,

novelty detection must be considered for successful SHM, at least at the lowest level of SHM, simply due to the fact that a supervised learning approach is limited to scenarios that can be anticipated, or have occurred before and for which data are available. Novelty detection must be incorporated into any comprehensive SHM system to safe guard against unforeseen circumstances (Black Swan events [36]), an indication that a structure is responding in a abnormal way can then be investigated further.

The pattern recognition problem as a whole is further complicated by the fact that many features undergo variability caused by operational and environmental conditions (as alluded to above), which must also be accounted for by the inference procedure [37]. The influence of environmental and operational variation can make classification problems in supervised learning very complex, as the data may become separable in a large number of ways according to different operating conditions. Novelty detection is also compromised if external conditions produce a novel structural response from an undamaged structure. How this problem may be overcome is the focus of the remainder of this thesis, and will be introduced more comprehensively in the next chapter.

1.3 Conclusions

The ideals and aims of SHM have been discussed in this introductory chapter along with some of the common practice of those attempting to implement it. Although SHM as a field of research could now be considered mature, the fact that the developed technology has seen little success to date for real world applications, is indicative of the huge challenges that are faced. Of these challenges, the largest appear to be the development of sensing technology capable of detecting critical fault types and how technology developed in the laboratory can be applicable to structures in operation. This thesis generally concerns the latter, and undertakes what is considered to be a major problem when attempting to apply SHM technology outside of laboratory conditions, which is how to deal with the often confounding influence of changing environmental and operational conditions.

Chapter 2

SHM in changing environmental and operational conditions

This chapter introduces how the influence of changing environmental and operational conditions can be problematic when attempting to infer structural condition from monitoring data. Firstly, the motivation for this research is set out, the chapter also gives an outline of others' approaches for dealing with the effects of a changing environment on SHM features. Finally, the layout of this thesis is summarised.

2.1 Motivation

Over the last few years, there has been an increasing trend in the civil and structural engineering community of instrumenting dense sensor networks on structures for SHM purposes. Such sensor networks commonly stream measurement data continuously, collecting and storing a huge amount of information on a daily basis [7]. Although this is an encouraging trend for SHM, very often it seems that little is done with the data collected. In some ways, this is indicative of the early stage that research into comprehensive monitoring for SHM is at. For many monitoring campaigns the means to carry out reliable system identification are still being sought, as an example, GPS technology is being trialled in many bridge monitoring campaigns (see [38] for a good example of typical investigations currently under way).

The work covered in this thesis is funded by EPSRC, and a principal aim of the

funded project was to make the first steps towards utilising data from comprehensive monitoring campaigns of bridges, specifically cable-stayed bridges, for SHM. The research detailed in this thesis was carried out in collaboration with the Vibration Engineering Section (VES) in the Department of Civil and Structural Engineering at the University of Sheffield, who currently monitor a wide range of structures including three cable-stayed bridges. The overall aim of the collaborative research was to develop sensible ways to access the large amounts of data measured in these campaigns and discover the most helpful ways this data could be used. Within the remit of this project, the aim of the author's work is to address the effects of changing environmental and operational conditions on the measured responses of these structures.

One of the most comprehensive of the VES monitoring campaigns to date is undoubtedly that of the Tamar Bridge in Southwest England [39]. A milestone of this research will be to gain a greater understanding of the effects of changing operational and environmental conditions on the measured response of this structure. The second challenge is then to develop ways to account for the effects of the varying environment so that reliable information on structural condition can be inferred from the measured response variables. In the following, a short overview of the issue of environmental and operational variations in SHM will be given, before an outline of the remainder of this thesis is made.

2.2 The problem of changing environmental and operational conditions for SHM

As previously alluded to, the effect of changing environmental and operational conditions on a structure is an important issue in SHM, and has been identified a key concern to the research community [40]. This interest arises from the inconvenient fact that measured responses from a structure that demonstrate sensitivity to damage or structural degredation, will, in general, also exhibit sensitivity to any change in operational and environmental conditions [8]. This is especially relevant in the context of civil monitoring campaigns, where typically, for a structure in operation, all measured structural responses are subject to daily and seasonal variations induced by (amongst others) temperature, wind loading and operational loading (such as traffic loading for bridges). Such structures will often exhibit inherently

nonstationary dynamic and quasi-static responses which can mask any changes in structural response that would be indicative of the occurrence or progression of damage, or of a change that could signify a performance anomaly. In these cases the effects of the environmental and operational variation must be accounted for in some way before a reliable measure of structural condition can be inferred. This confounding influence of environmental and operational conditions is, in fact, considered as one of the main inhibiting factors slowing the uptake of SHM by industry. The problem is often referred to as the data normalisation problem [37].

In the SHM literature, undoubtedly the most commonly occurring discussion of the confounding influence of environmental and operational conditions on damage sensitive features arises from the sensitivity of structural response to temperature. For bridges, temperature is generally considered to be a dominant environmental factor affecting the normal dynamic response, due to its effect on the stiffness of structural components, and also its potential effect on the boundary conditions of a structure (for instance from the freezing of foundations etc.). Historically, many previous studies have found fluctuations in modal frequency to be correlated with ambient temperature, although different mechanisms have been used to explain this, see for example [15, 41–43]. Cornwell et al. [42] suggested that the thermal gradient across the deck of the Alamosa canyon bridge drives the observed fluctuations in modal frequency. In colder climates significant shifts in frequency between above and below freezing temperatures have been attributed to an increase in stiffness explained by the Young's modulus of the asphalt on the deck at colder temperatures [15]. In this case, the modal frequencies of a bridge deck were observed to have a bi-linear relationship with temperature. A similar behaviour has also been observed in a steel truss foot bridge in the US [44].

Besides temperature, the importance of other operational conditions have also been considered for bridge structures. A dominating topic of research that must be mentioned here is on unstable bridge response to wind conditions, which has been a major concern for long span bridges since the collapse of the Tacoma Narrows bridge in 1940. A large body of research has been undertaken to better understand the interaction between wind and phenomena such as buffeting and flutter of bridge structures (see for example [45, 46]). Generally the aim of work in this field is to ensure that new bridge designs are safe and to monitor structures during construction and in early life to ensure that design criteria to avoid self excited and wind induced oscillation have been met.

More in the context of this thesis, where problems with unstable responses of such structures are not addressed, are the monitorable correlations between damage sensitive features and environmental/operational conditions. In this context the response of a long span bridge to high and low wind speeds was investigated in [47], where it was concluded that the modal frequencies of the structure decreased with increased response amplitude levels directly caused by increased wind speed. The effect of humidity alongside temperature has also been studied. In [48], the effect of humidity and temperature on the modal parameters of a reinforced concrete slab are investigated, where it is reported that increased humidity effectively adds mass to a structure, and has a strong negative correlation with modal frequency. The effect of traffic loading has also been addressed in [49], where, for long span bridges, the influence of traffic loading on the structure's frequency was considered negligible due to the fact that the mass of a single vehicle is very small in comparison to the mass of the 'superstructure'. In a separate study, however, the modal frequencies of a cable stayed bridge were found to vary up to one percent a day due to traffic loading [50].

A review of the relevant literature reveals a number of potential options already explored for dealing with the problem of operational or environmentally induced variations in structural response. Perhaps the most common approach has been to attempt to model the monitored parameters or damage sensitive features in question with respect to those environmental/operational factors considered to be driving its/their variation [15, 44, 51–58]. If a model can predict the value of a damage sensitive feature given the conditions affecting it, the error of the model could be suitable as a robust indicator of structural condition. Often these approaches have employed a simple regression of the damage sensitive feature (normally natural frequencies) onto measured structural temperature [15, 44, 52–54, 58]. More complex approaches for regression have also been explored [55–57], where modal parameters of the Ting Kau bridge, Hong Kong, have been regressed onto measured temperature using support vector machines, principal component analysis and neural networks. In a very similar vein, tracking the correlation between the measured strain of a harbour wall and temperature has been explored in [51].

With such approaches, the main limiting factor is that the changing environmental and operational conditions have to be identified and accurately measured. While this may be feasible where only one or two environmental or operational factors are important, such as temperature, where multiple factors affect the features of interest

a substantial monitoring campaign will then become necessary.

A very different approach to others in the literature, suggested in [59], is to incorporate temperature compensation directly into the structural identification step. In this case an analytical expression for the effect of temperature on a structure is incorporated into an algorithm for the identification of modal parameters. So far, only the effect of temperature has been considered and the methodology only trialled in simulation.

On the topic of addressing this problem during a system identification procedure, as mentioned in the introductory chapter, research into the modelling of nonstationary random vibration is a growing area of interest. Models with time varying parameters have been employed that can account for signal nonstationarity (although not necessarily induced by environmental and operational variations) [32, 60]. Where employed such models could provide a more straightforward means of inference on the condition of structures in operation.

Other approaches, where perhaps measurements of the environmental/operational conditions are not available have also been explored. A simple potential solution to the problem is to use a long span of response data to define the normal condition of a system, an idea explored in [61], this could be, for example, data collected over a whole year where all ranges of environmental/operational conditions have occurred. New measurements may then be compared in some way with the defined normal condition. Evidently this approach requires storage of a large amount of data, and a further drawback is that using a large normal condition set may reduce feature sensitivity to damage [58].

A number of other studies employ what may be described as latent variable models [62], which, without measurement of the changing environment, attempt to capture the variation in the feature data caused by it. Principal component analysis (PCA) has been used in a number of studies to re-express multivariate SHM feature data with a new set of orthogonal coordinate axes [63, 64]. These axes (called principal components) are linear combinations of the original coordinate axes ordered according to the amount of variance in the data each axis accounts for (see Chapter 4 for more details). The assumption employed in these studies is that the high variance signatures of changes induced by environmental and operational conditions in SHM features will be trapped in the higher principal components. In [63], this assumption is exploited by discarding the higher variance principal components and projecting

temperature dependent data onto the minor components which constitute a temperature independent feature set. In [64], only the higher variance principal components are retained and used as a model to predict/reconstruct the feature data. In a similar way as described above for regression, the 'model' error is then used to indicate an abnormal response. This idea of linear projection and trapping of environmental variation was also independently proposed through the use of Factor Analysis (FA), which is a very similar algorithm to PCA [65, 66].

Although such approaches appear to be promising solutions to the data normalisation problem, it has been shown that nonlinearity can hamper the effectiveness of employing PCA and similar algorithms. In [67], a remedy to problems introduced by nonlinearity was to cluster feature data into several (linear) regions and then employ PCA separately to each region. An auto-associative neural network, which may be said to be equivalent to nonlinear PCA, is used in [68] for data normalisation of features extracted from an auto-regressive type model. An auto-associative neural network (nonlinearly) maps its inputs onto themselves. The premise of using them for data normalisation is that if the network is trained on data from an undamaged structural condition, it will learn the effect of latent variation on the features inputted to the network. It is then expected that the network error will increase if damage occurs. In [69, 70] nonlinear PCA, achieved through a kernel based algorithm, is used in a different way. The nonlinear principal components are calculated for a set of data, the individual mapping for each data point is then considered. Data points with similar mappings (measured by Euclidean distance), which presumably come from similar environmental/operational conditions, are compared with each other, abnormal response is then detected if a single measurement can be considered as an outlier to this cluster.

Along similar lines, a new approach for data normalisation has recently emerged, where, for multiple sensor arrays, Gaussian process (GP) regression is used to predict the measurement of each single sensor from the measurements of all other sensors in the network [62]. Given suitable training data from different environmental and operational conditions, the GP should be able to accurately predict structural response at each sensor if the structure continues to operate in similar way as in the period where the training data was recorded. In a similar way to the regression techniques, the GP regression model error is used as an indicator of abnormal structural response.

Singular value decomposition (SVD) has also been explored as a potential way to

detect damage from structural responses under the influence of these confounding trends [71, 72]. When using an SVD for this purpose, the general assumption that is made is that environmental and operational conditions will produce a global change in the structure (to the stiffness or mass for example) and that damage will only produce a local change. When studying natural frequencies, this assumption implies that environmental and operationally induced changes in structural response will be linear in nature, whilst changes induced by damage will be nonlinear [72]. If this assumption holds true, an SVD can be used to determine the effective rank of a matrix of response measurements from an undamaged structure. When new measurements are added to this matrix, if the effective rank increases, a nonlinear change of the damage sensitive feature can be inferred which then implies that damage has occurred. Under a similar assumption, in very recent work [73, 74], a state space representation of response data is considered. A state space reconstruction is used for the prediction of the state of a structure in a healthy condition, which as in other approaches, will fail to predict new states well if an abnormality occurs (in this case if the structure responds in a nonlinear manner).

Despite the sophistication of some of the solutions to the data normalisation problem summarised here, this area of research can still be considered as exploratory. The vast majority of the studies cited here only consider (and attempt to account for) the effect of temperature on structural response. Furthermore, although the use of latent variable models seems like a very promising solution, none have been comprehensively trialled, and can still be considered as under development. In this thesis, the effect of multiple environmental and operational conditions will be considered, and a new approach for data normalisation suggested.

2.3 Scope of this thesis

This thesis will address the data normalisation problem in the context of data collected, mainly, from bridge monitoring campaigns. A large part of the work will be based on data collected by the Vibration Engineering Section at the University of Sheffield from the Tamar suspension bridge located in South West England. A major aim of this work is to begin to develop diagnostic tools that could be used to indicate the condition or performance of this structure, which is, of course, subject to changing environmental and operational conditions. In order to achieve this, the first

task is to understand how the varying environment affects the measured responses of the bridge, in other words to understand what constitutes the '*normal condition*' of the structure. With a sound understanding of this gained, ways to account for environmentally induced trends that may obscure any indication of degradation in a structure's conditions can be attempted. The second half of this thesis is devoted to the development of a new way to remove trends induced by environmental and operational conditions from damage sensitive data. This new approach is based on the idea of *cointegration* which originates in the field of econometrics.

2.3.1 Brief outline of thesis

- Chapter 3 introduces the Tamar bridge and VES monitoring campaign. Details are given of the data available for analysis in this work.
- Chapter 4 explores how one can define what constitutes the 'normal condition' of a structure. The main drivers of the variation in modal frequency of the Tamar bridge are investigated.
- Chapter 5 builds on the findings of Chapter 4 in order to create features for monitoring the structural condition of the Tamar bridge. Regression models are used to predict the modal frequency variation and displacement of the deck.
- Chapter 6 introduces the concept of cointegration, theory from the field of econometrics, and how it can be used to remove environmental and operational trends from damage sensitive features.
- Chapter 7 demonstrates the application of cointegration for the data normalisation problem using data from the Tamar monitoring campaign. The implications of using econometric theory for engineering applications is also discussed.
- Chapter 8 applies the developed cointegration theory to a benchmark study involving Lamb-wave propagation for detecting damage in a composite plate. A comparison is made with the PCA approach applied in [63].
- Chapter 9 begins investigations into nonlinear cointegration, for problems where damage sensitive features are nonlinearly related.

- Chapter 10 concludes the thesis. Future work is discussed.

Chapter 3

The Tamar Bridge

A large part of the work carried out in this thesis will be based on data collected from the Tamar Suspension Bridge which is located in South West England. The bridge is monitored by the Vibration Engineering Section (VES), in the Department of Civil and Structural Engineering at the University of Sheffield. The EPSRC grant that funds this research brings the author and colleagues together with VES to work with the data collected from this monitoring campaign. Some background information about the bridge will be given in this chapter along with details of the monitoring campaign before any analysis of the available data is attempted. Although the author has been given access to all data from the Tamar Bridge, it must be noted that she took no part in the measurement and processing of the available data, all credit for this must be given to the members of VES.

3.1 The Tamar Bridge and some of its history

The Tamar Bridge (Figure 3.1) has been a vital transport link over the River Tamar carrying the A38 trunk road from Saltash in Cornwall to the city of Plymouth in Devon since its construction in 1961. The original bridge was designed as a conventional suspension bridge with symmetrical geometry, having a main span of 335 metres and side spans of 114 metres. With anchorage and approach spans, the overall length of the bridge reaches 643 metres. The towers are constructed from reinforced concrete, and have a height of 73 metres with the deck suspended

Figure 3.1: The Tamar Suspension Bridge

at half this height. The towers sit on caisson foundations founded on rock. The main suspension cables are 350mm in diameter, each consist of 31 locked-coil wire ropes and carry vertical locked-coil hangers at 9.1m intervals. The main cables are splayed at anchorages and anchored some 17 metres into rock. The truss is 5.5 metres deep and composed of welded hollow steel boxes. The original three-lane deck, spanning between cross trusses, was of composite construction with a 150mm deep reinforced concrete slab on five longitudinal universal beams and surfaced with 40mm of hand-laid mastic asphalt.

3.1.1 Upgrade

When opened in 1961 Tamar Bridge was, for a short time, the longest suspension bridge in the UK and was also the first to be built after the end of World War II. In the late 1990s, after nearly four decades of use, it was found that the bridge would not be able to meet a new European Union directive that bridges should be capable of carrying lorries up to 40 tonnes in weight. Since restricting use by such vehicles would damage the local economy, the bridge was strengthened and widened. After considering a number of options, the appointed consultant (Hyder) proposed replacement of the main deck with a lightweight orthotropic steel deck, with construction of temporary relief lanes cantilevered either side of the bridge truss. These lanes were originally intended to act as a supplementary diversion route while the main deck was being replaced but were finally adopted as part of

(a) Below deck view (b) Birds-eye view

Figure 3.2: Cantilever lanes added in the upgrade to the Tamar Bridge

the permanent solution.

Pairs of prefabricated orthotropic panels, each typically 15m long and 3m wide were welded longitudinally to form the 6m wide cantilever sections, also surfaced with hand-laid mastic asphalt. As proposed, a new light-weight orthotropic steel deck replaced the original three lane composite deck slab. Eighteen new locked-coil cables were installed and stressed to supplement the original suspension system, primarily to help carry the additional dead load of the new cantilever lanes and associated temporary works (Figure 3.2).

In summary, approximately 2,800 tonnes of structural steel was added together with 125 tonnes of cables; however, when offset by the removal of the old main deck, the final weight of the suspended structure rose by just 25 tonnes to 7,925 tonnes. The deck replacement process was completed in December 2001 and the bridge now carries about 50,000 vehicles per day. This upgrade gave rise to interest in the bridge performance, and various sensor systems have been installed to measure parameters such as tensions on the additional stays, wind velocity and structural temperature. As the bridge displacement information is essential for assessing performance, surveys of the bridge deflection profile have been carried out periodically and a hydraulic levelling system has also been installed to monitor vertical deflections of the main span.

3.2 Monitoring the Tamar Bridge

Currently three monitoring systems are in place and running at the Tamar Bridge. The first is a Structural Monitoring System (SMS) installed by Fugro Structural Monitoring, which is used to monitor cable loads, structural and environmental temperatures and wind speed and profile. This system was installed during the upgrade to provide information on the performance and condition of the bridge during and after the strengthening and widening. The sensors used in the SMS include:

- anemometers to measure wind speed and profile,
- a fluid pressure-based level sensing system to measure deck vertical displacement,
- temperature sensors for the main cable, deck steelwork and air temperature,
- extensometers and resistance strain gauges to measure loads in additional cables.

An additional set of sensors was installed by the University of Sheffield (VES) in 2006 to monitor dynamic behaviour of the bridge deck and selected cables. Four stay cables are instrumented, each with a pair of accelerometers: one oriented horizontally and one in the vertical plane (an example of which is shown in Figure 3.3a). As well as the eight cable accelerometers, three accelerometers are installed to measure acceleration of the deck, two of which are shown in Figure 3.3b. The Sheffield system records 64Hz-sampled time series in files at 10-minute intervals. From the accelerometers an automated system implemented by VES identifies modal parameters every ten minutes using an output-only modal analysis [14] that relies on stochastic subspace identification (SSI).

Stochastic subspace identification is commonly used when attempting output-only modal analysis on structures in operation [75], it is thought to be one of the most powerful identification techniques currently available for modal parameters [76]. Estimates of modal parameters are obtained via the identification of a discrete state-space model where the (ambient) inputs to the system are assumed to be white noise. The VES implementation of SSI is classified as a data-driven automated approach, identification of the state space model is achieved by construction of a block Han-

(a) Cable accelerometer

(b) Vertical and horizontal deck accelerometers

Figure 3.3: Accelerometers installed on the Tamar Bridge

kel matrix of measurement data and employment of numerical techniques such as singular value decomposition and QR factorisation. For more details readers should consult [30, 75–77], particularly [77], where examples of code for the implementation of SSI routines are available. In this thesis, the natural frequency data extracted by the SSI routine will be utilised, in particular, only the lowest five frequencies will be studied. This limitation is due to the fact that the properties of the higher modes cannot be estimated with as much fidelity.

The newest monitoring system introduced by VES is a total positioning system (TPS) which uses a robotic total station (RTS) for precise three dimensional displacement monitoring of the deck and towers, accurate to within 2 or 3mm. The RTS, shown in Figure 3.4, was installed in September 2009 on the roof of the Tamar Bridge office which sits close to the bridge on the Plymouth side bank of the river. Fifteen reflectors (see Figure 3.5) have been installed around the bridge including on the deck, main towers and side towers. Single static displacement measurements from each of the 15 reflectors are repeated every 30 minutes, with each measurement cycle taking about 10 minutes to cover all 15 reflectors.

A traffic count is also available from the toll gates that monitor the flow of traffic in one direction (east-bound) across the bridge. The information available is the number of vehicles passing the toll every hour, automatically classified into ten different categories (from motorcycles up to 4-axle heavy goods vehicles with trailers,

Figure 3.4: Robotic Total Station installed on the roof of the Tamar Bridge office

Figure 3.5: Robotic Total Station reflectors installed on the bridge deck

as listed in Table 3.1). For the work in this thesis, an estimate of the traffic loading on the bridge at any one time is calculated from this traffic count in the following way. Firstly, the average weight of vehicles in each of ten categories has been roughly estimated, as recorded in Table 3.1 (informed by a government transport statistics report [78]). Each vehicle count is multiplied by its respective estimated weight and summed, this is then is multiplied by two to account for the fact that the traffic is only monitored in one direction. Finally, to obtain an estimate of the instantaneous load on the bridge, the summed load is divided by 45 which reflects the estimated time it takes for a vehicle to cross the bridge.

Vehicle Category	**Estimated average mass (Kg)**
Vehicles not picked up by Automatic Vehicle Classification	0
Motorcycles	150
Cars, Trikes and PLG[1] under 3.5 tonnes (gross vehicle weight)	1500
2 axle HGVs	18000
3 axle HGVs	26000
4+ axle HGVs	32000
Cars with trailers	8250
2 axle HGVs with trailer	21000
3 axle HGVs with trailer	30000
4+ axle HGVs with trailer	36000

Table 3.1: Estimated vehicle weights used in this thesis.

3.2.1 Data Reliability

As described above the Tamar monitoring campaign provides a large amount of data in various forms. The data made available for analysis in this thesis spans over three years from 2007 to 2011 and provides ample opportunity for in-depth analysis. Before measurements from any monitoring campaign can be analysed, however, it is inevitable that some action will be needed to obtain a consistent and reliable data set with which to work. Firstly, in any bank of historic data there will likely be periods (short and long) where the monitoring system has failed, this may be relating to power supply, telemetry or memory issues, amongst others. Further to gaps in data caused by failure of the monitoring systems, missing individual data points are also common for campaigns of this nature. Two examples of monitored response that are prone to missing data points in the Tamar monitoring campaign

are the extracted modal frequencies of the bridge deck and the deflections of the bridge deck and towers as measured by the TPS. Occasional missing data points in the modal frequency time histories originate from a mode misclassification by the SSI routine. Missing data points in the TPS system are often caused by weather conditions (cloud or fog) that obscure the reflectors that are used to make the displacement measure. Aside from missing data, faulty sensors are also a common problem that must be tackled.

The first task at hand when provided with a large data set is to assess the reliability of each individual data channel, and to identify any long or short gaps. In this work, all suspect data channels (those with anomalous drifts, dubious scale, etc.) were excluded from the analysis that follows in this thesis. Any gaps in the data set have also been identified, occasional missing data points have been replaced through linear interpolation. Again it must be pointed out that the majority of the necessary data/signal processing had been carried out by VES before the author accessed it.

3.2.2 Summary of data available from the Tamar Bridge

Table 3.2 provides a summary of the data available for analysis from the Tamar monitoring campaign. The sampling rates and intervals vary for different measurements, however, synchronised half hourly measurements made available by VES will be used in any analysis in this thesis. Figure 3.6 also shows a schematic of the bridge, provided by VES, with the location of each sensor marked.

As an example of some of the data available from this monitoring campaign, samples of the measured cable tensions (of the stay cables added in the upgrade at six locations) and structural temperature are plotted in Figures 3.7 and 3.8. Each cable tension plotted is an average of the measured tension in the stay cables on either side of the bridge at the locations listed in the legend (which correspond to the sensor labels in Figure 3.6). Figure 3.9, then shows the correlation between cable tension and structural temperature. From this figure, it is evident that cable tension is correlated with temperature, as is to be expected. However, it is interesting to note that most of the cable tensions are negatively correlated with temperature apart from at two locations where the correlation is positive. These positive correlations are with the stay cables attached at the Plymouth tower towards the centre of the bridge. A simple explanation for these observations is that as the bridge deck

expands with increased temperature, the expansion joint at the Saltash tower allows the deck to move in westerly direction (towards Saltash), this increases the tension in the stay cables attached to the Plymouth tower, and slackens the others.

Measurement description	**Sensor description**	**Number of measurements available**	**Units**	**Sampling Frequency (Hz)**
Cable tensions	Extensometer & resistive strain gauges (Fugro)	34	kN	1
Deck level	Fluid manometer system (Fugro)	20	mm	1
Displacement of deck and towers	Total positioning system (Leica)	45	m	Half-hourly (static)
Cable and deck acceleration	Accelerometers (Honeywell QA750)	7	g	64
Traffic count	Toll count	10	-	Hourly cumulative
Temperature (ambient and structural)	Resistance thermometers (Fugro)	22	$^{\circ}C$	1
Humidity	Hygrometer (Fugro)	4	%	1
Wind speed and direction	Anemometer and vane (Vector Instruments)	7	mph	1

Table 3.2: Summary of available measurements from the Tamar Bridge.

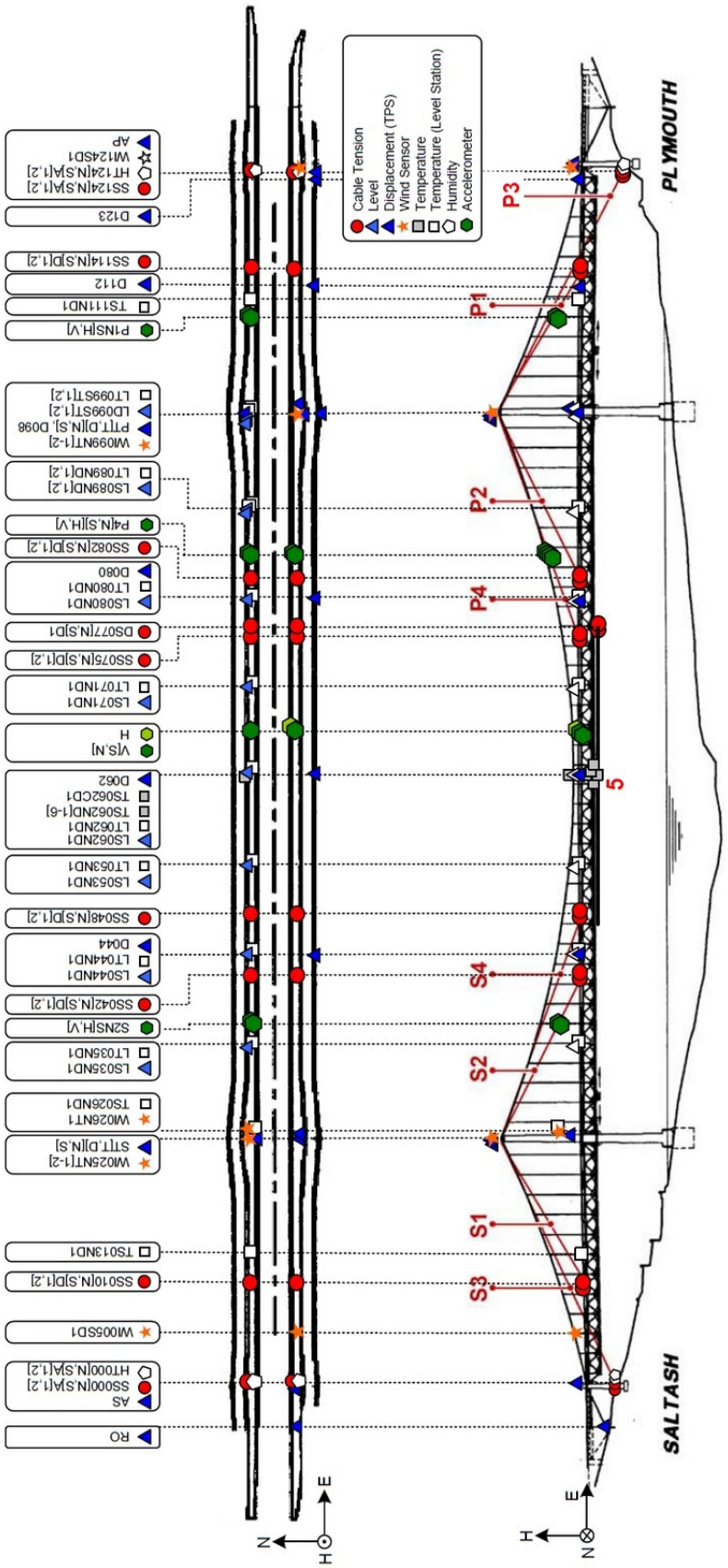


Figure 3.6: Sensor locations for the Tamar Bridge

3.3 Conclusions

This chapter has introduced the comprehensive monitoring campaign of the Tamar Bridge, which is led by the Vibration Engineering Section (VES) at the University of Sheffield. The monitoring campaign provides a wealth of data that can now be used to make the first steps towards using real monitoring data for SHM. This bridge forms the main case study in this thesis. In the next chapter the data from this campaign will be investigated in more depth and attempts will be made to understand what constitutes a normal structural response in the face of changing operational and environmental conditions.

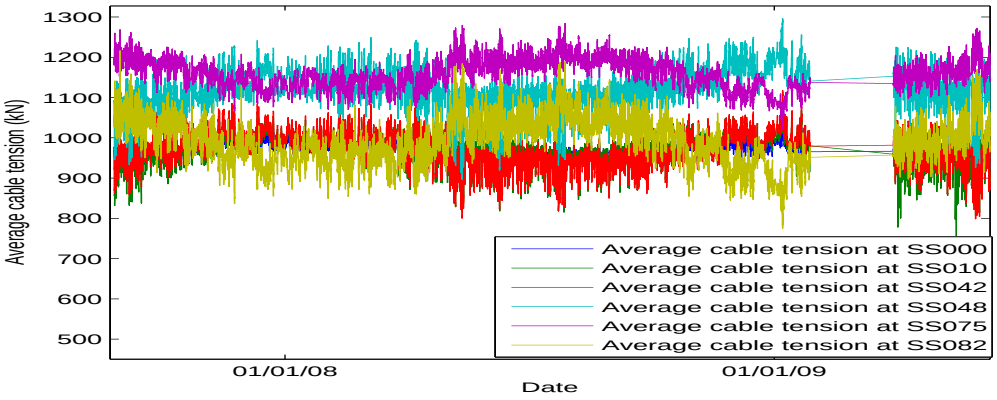


Figure 3.7: Average cable tensions at different locations

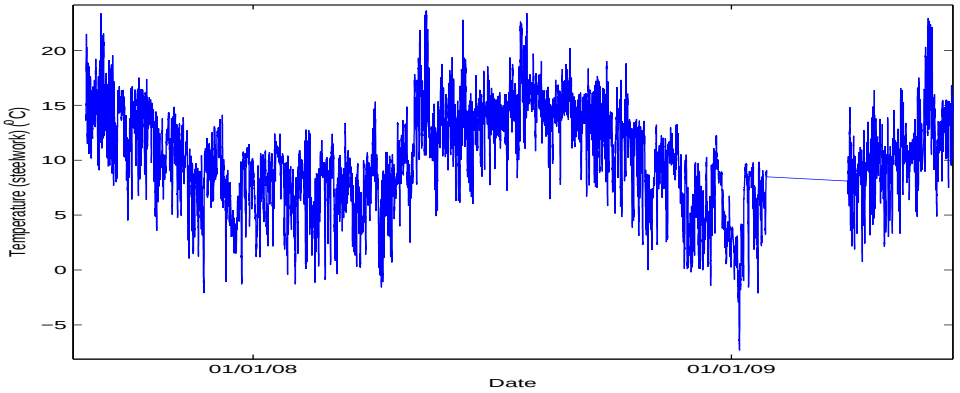


Figure 3.8: Measured structural temperature

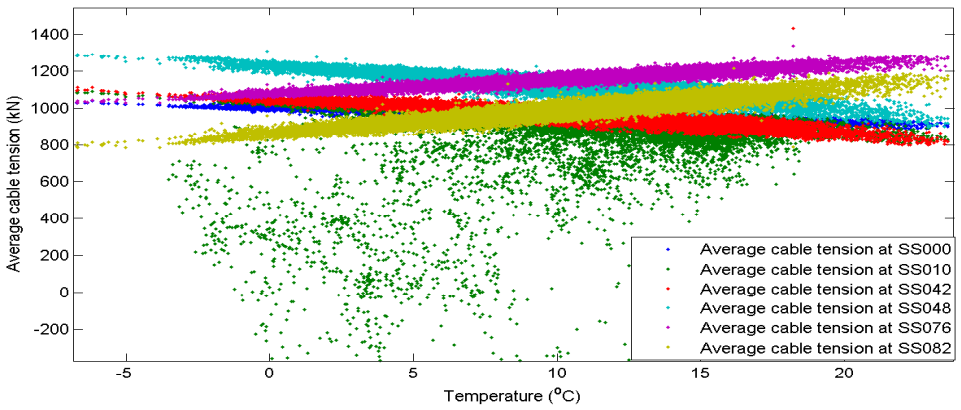


Figure 3.9: Correlation between stay-cable tension and structural temperature

Chapter 4

Defining the normal condition for the Tamar Bridge

The current trend of collecting large amounts of data in the name of SHM from civil structures cannot be constructive unless this data is utilised. As one of the first projects where a concerted effort is being made to use monitoring data from a large scale in-service structure for the development of helpful SHM strategies, the first priority in this work is to understand a structure's response in the context of changing environmental and operational conditions.

At least three years of reliable monitoring data is now available from the Tamar monitoring system, which provides a most unique opportunity for development of a reliable SHM system. From this large database, the first task on the way to developing an SHM system is to understand the structure's normal condition, which is to say, how the structure responds to normal variations in environmental and operational conditions. For a successful SHM routine a sound understanding of all mechanisms affecting a structure's response is extremely beneficial.

This chapter specifically aims to study the effect of multiple environmental and operational conditions on the dynamic response of the Tamar Bridge. If the dynamic response of a structure is used for structural condition assessment, all variations due to anything other than changes in the structure itself must be understood and accounted for. If the environmental and loading effects can be accounted for and filtered out to normalise response data, any changes in dynamic characteristics must

signal some structural change, which could be a slow degradation (e.g. loss of cable tension or reduction in member stiffness through corrosion) or some sudden change such as a seized bearing or failure of a structural member. In the literature of SHM, particularly for aerospace applications, these gradual or sudden changes are collectively called 'damage', and the technology used to identify them from response data is termed 'damage detection'. Since proposing 'damage detection' technology to a bridge operator or a structural engineering research proposal review is not a winning approach, the term 'structural performance anomaly' is more appropriate than 'damage' in this context.

For the purposes of data normalisation, the effects of temperature, traffic loading, wind loading (and consequently deck acceleration) will be considered in the following analysis. The study of traffic loading here is especially rare due to the fact that many previous bridge monitoring campaigns have had to be conducted while the structure was out of action.

Despite the fact that data spanning three years from the Tamar monitoring system are available, there are, understandably, periods within these three years when the monitoring system failed. In the following analysis data collected in 2007 and 2008 will be considered; this is principally because of a suspicion that the casings around one of the sensors may have become waterlogged early in 2009, which may have affected the response recording.

4.1 Dynamic response

As previously described, dynamic data for the Tamar Bridge is extracted from accelerometer data using a data-driven SSI technique [30]. In this section the variation of the first five modal frequencies of the deck with respect to temperature, wind speed and traffic loading are investigated. Discussion of the potentially complex effects of deck acceleration will follow in its own subsection, as it can itself be affected by traffic and wind speed.

The simplest approach in determining an environmental or operational variable's impact or importance on the fluctuations of the modal frequencies of the bridge is to plot each frequency with respect to those variables. Figure 4.1 shows the first five modal frequencies of the deck plotted with respect to temperature, wind speed,

and traffic loading, along with linear best fit lines which have been added as a visualisation aid. In Figure 4.1 the lowest frequency trend corresponds to the first vertical symmetric mode (denoted VS1), the second lowest corresponds to the first lateral symmetric mode (LS1), the next to the first vertical anti-symmetric mode (VA1), the second highest to the first lateral anti-symmetric mode and finally the highest corresponds to the first anti-symmetric torsional mode (TA1).

For further visual clarification of the influence of each variable, a principal component analysis (PCA) of the frequency data is carried out. Principal component analysis takes a multivariate data set and projects it on to a new set of variables, or 'principal components', which are linear combinations of the old variables. Of these new variables, the first principal component will account for the biggest proportion of the variance in the data set that can be described by a single axis, the second principal component will account for the second biggest proportion of the variance in the data set independent of the first, and so on. If the original number of variables is some number p, up to p new variables may be formed. Now, if the first n of these principal components represent a significant amount of the variance, it is fair to say that the data can be suitably represented solely by these n principal components without loss of any real information. Principal component analysis, therefore, works to reduce the dimensionality of the dataset, which can considerably ease analysis of datasets of high dimensionality. PCA is commonly used for a wide variety of tasks, here, the reduction of dimensionality of the data greatly aids visualisation of any possible structure within the data. In this work, the data are transformed to have a zero mean and unit standard deviation prior to analysis, throughout the thesis this action will be referred to as *normalisation*. Specifically, for the uses of this work the first two principal components of the frequency data are plotted, for the data set considered (which spans the period of 16 months), the first two principal components account for 78.6% of the variance in the data. Figure 4.2 shows the first two principal components of the data plotted against each other with each of the data points coloured according to whether they occurred at high, medium or low temperatures, wind speeds and traffic loadings. For further information on PCA readers are referred to any text book on multivariate analysis (a good example being reference [79]).

Figure 4.1 demonstrates that each of the first five modal frequencies of the deck have a tendency to decrease with increased temperature and increased traffic loading. For the frequency that appears most sensitive to temperature (the second,

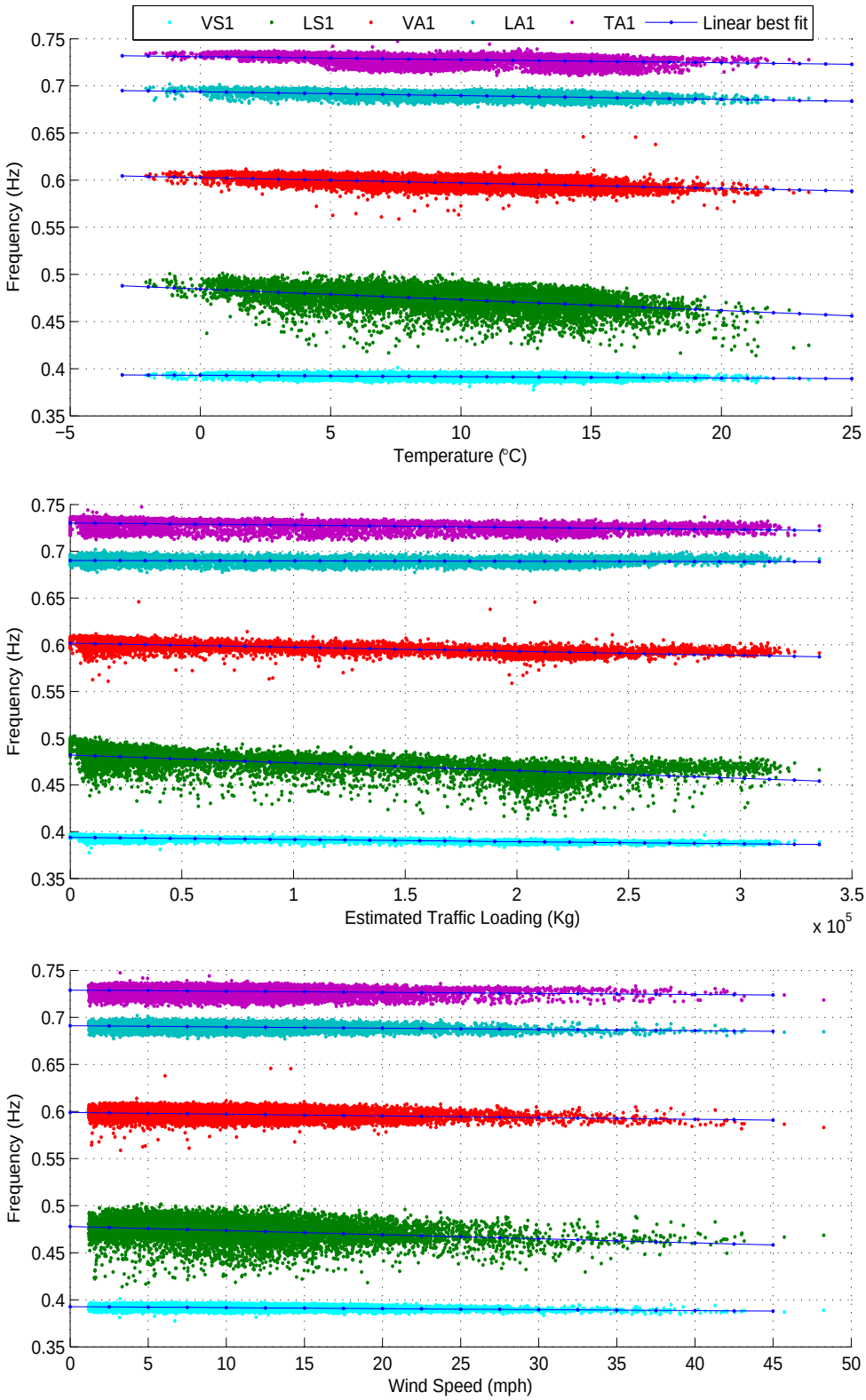


Figure 4.1: Deck Modal Frequencies plotted with respect to traffic loading, wind speed and temperature

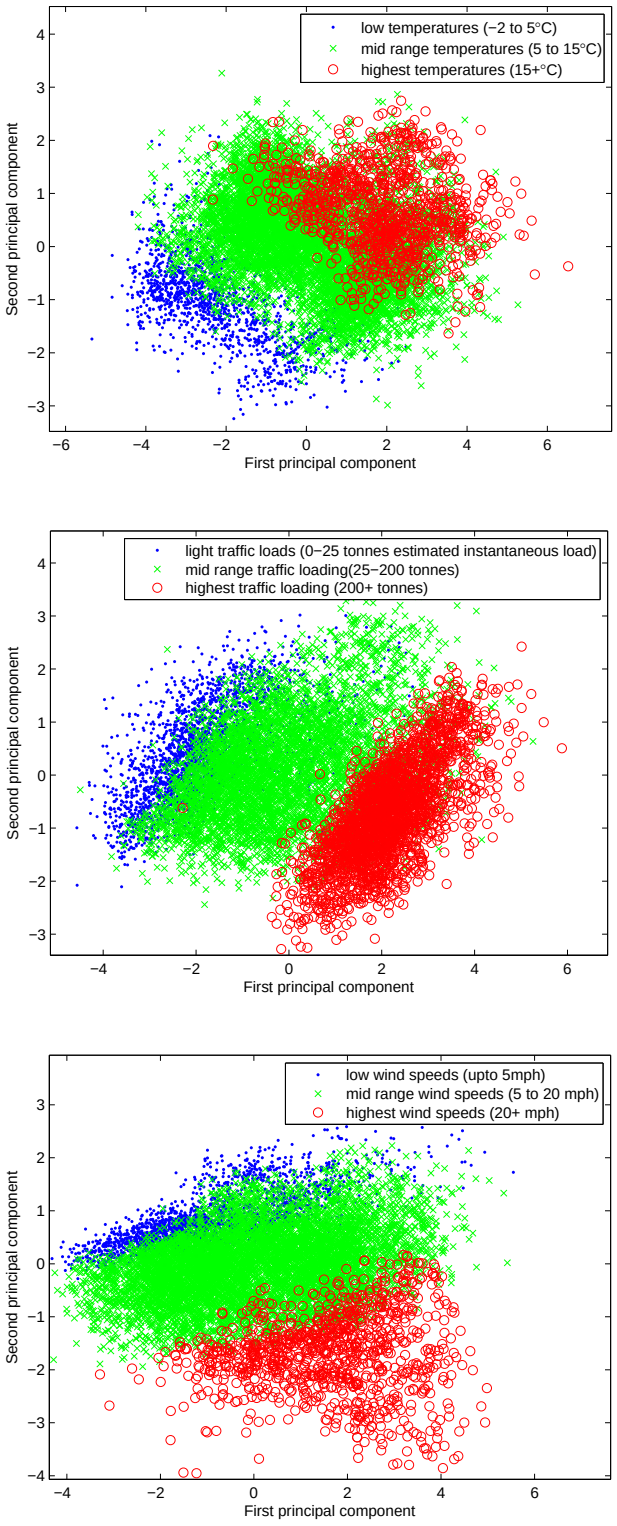


Figure 4.2: First two principal component scores of frequency data sorted according to temperature, wind speed and traffic loading

which corresponds to the first lateral symmetric mode), the frequency decreases by approximately 4.5% over a 20°C change in temperature. Similarly, the frequency for the second mode changes by around 3.5% between periods of low and high traffic. Figure 4.2 shows that the frequency data are uniquely distinguishable by both temperature and traffic loading. This indicates that both temperature and traffic loading have a separate but significant influence on how the frequencies vary.

The dependency of any modal frequency on wind speed is unclear from Figure 4.1. Although it appears that generally frequencies are lower at very high wind speeds, no clear conclusion can be drawn as the majority of the data occurs at low to moderate wind speeds. Figure 4.2, however, demonstrates that the frequency data can be sorted according to wind speed. High wind speeds have previously been found to have an influence on the stable dynamic characteristics of long span bridges [47], which is supported by the pattern shown in Figure 4.2. The effect of wind loading on the bridge will be considered later on in the analysis in more detail, when the effect of the deck acceleration is addressed.

In summary, Figures 4.1 and 4.2 indicate that temperature and traffic loading are the dominant environmental and operational factors affecting the modal frequencies of the deck, the PCA plots in Figure 4.2 also suggest that wind speeds may have some influence. Despite the obvious mass increase that must arise from heavy traffic, previously, little attention has been given to the effect of traffic loading on modal parameters in this context. From toll counts and web cam images, the instantaneous traffic loading on the bridge is estimated to increase by between 100 to 200 tonnes during very busy periods, which occur around 8am on weekdays. For a fixed stiffness, this change in mass would account for a 1.5-3% reduction of the modal frequencies, which is consistent with the variation encountered in Figure 4.1. This is in direct contrast to the conclusions drawn by Kim *et al.* [49], where traffic loads were not found to influence the modal frequencies of a long-span bridge. It should, however, be noted that the Tamar Bridge has to endure much larger traffic loads than those considered in [49].

Having determined that traffic loading should indeed be important it remains to separate out the effect of temperature from that of traffic loading. Figure 4.3[1] shows how a simple linear model with estimated traffic loading as its only input can predict the frequency change of the first mode (VS1).

[1]'Normalised' here indicates that the time series has been standardised to have a zero mean and unit standard deviation.

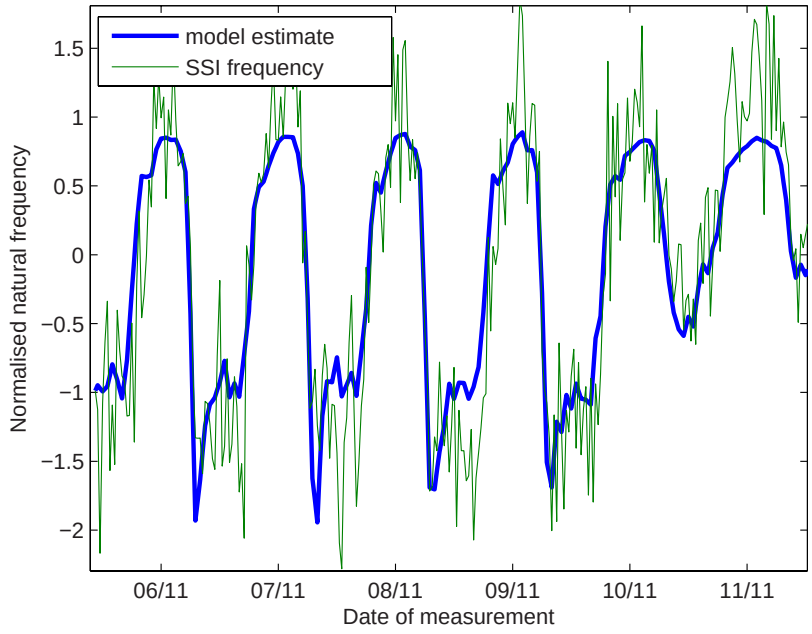


Figure 4.3: Linear model of first deck modal frequency with traffic loading input only

The model takes the form

$$\omega_1 = 0.099 - 0.79 \times (traffic\ load)\,. \tag{4.1}$$

This type of model is called a response surface model [80], the idea being originally developed by Box and Wilson [81] for modelling chemical processes. Such models are learned from data rather than established by using the underlying physics or chemistry; they are essentially regression models of varying degrees of sophistication. Response surface models are often used to learn the input-output relations from large computer models in order to produce fast-running approximations for Monte Carlo analysis; in this context they are called meta-models, surrogate models, fast-running models or emulators. Response surface models will be utilised in this work as a tool to better understand the interaction between normal structural response and the varying environmental and operational conditions. Where employed, they will all be low-order polynomials with the parameters established using simple least-squares analysis. The usefulness of response surface models in this context arises from their simplicity; in a simple regression model one can easily infer the importance of an independent variable on a feature of interest. To aid the interpretation of the

response surface models used in this thesis a measurement of model error will be utilised, as well as classic t- and F- tests, which will be employed to clarify the statistical significance of multiple parameters in the models. A short description of these statistical tests will be given below, however, as they are common statistical constructs, the descriptions will be brief. For more detail see [82] or any other good text book on applied statistics.

A normalised mean squared error (MSE) is introduced here as a measure of model fitness. Specifically, the normalised MSE used here is defined as:

$$MSE = \frac{100 \sum (model\ errors)^2}{n\ (\sigma[predictions])^2} \tag{4.2}$$

where n is the number of data points predicted and σ denotes standard deviation. This MSE has the property that, if the mean of the data is used as the model, the MSE will be 100%. With this normalisation, values of MSE below 100% are indicative of captured correlation. Some readers may be more familiar with the R^2 correlation coefficient as a means of studying model fitness; the MSE is related to the R^2 coefficient as follows:

$$R^2 = 1 - \frac{n.MSE}{100} \tag{4.3}$$

so R^2 increases as the MSE decreases.

A common t-test for regression can be used to assess the significance of individual parameters in a response surface model. To infer the significance of any parameter in a model, a test statistic is calculated that is dependent on the estimated parameter for each variable.

To conduct a t-test, the test statistic t for an estimated parameter/coefficient $\hat{\beta}$ that describes the contribution of an independent variable, X say, should be calculated and compared with the tabulated t-value $t_{(\alpha/2,n-p-1)}$, where α is the significance level required, n the number of observations used to establish the model and p the number of parameters in the model. The test statistic t to be calculated is:

$$t = \frac{\hat{\beta}}{se(\hat{\beta})} \tag{4.4}$$

where $se(\hat{\beta})$ is the standard error of the coefficient $\hat{\beta}$.

To test the significance of the estimated parameter/coefficient, $\hat{\beta}$, a null hypothesis of $\beta = 0$ is adopted; in other words the null hypothesis states that parameter X with estimated coefficient $\hat{\beta}$ is not important to the model under consideration. This hypothesis can be rejected at a significance level α if the calculated test statistic t is larger in magnitude than the tabulated t-value for that significance with the relevant degrees of freedom.

An F-test can be used to gauge the significance of a complete regression model by assessing the amount of variance in the dependent variable, y, say, that is accounted for by the model. Of interest here is the partial F-test, which assesses the significance of adding a single, or group of variables, to an established regression model and addresses the question of *'is the model significantly improved by adding certain input variables?.'* Suppose the established model has inputs $X_j, \{j = 1, 2, \cdots, p\}$ and corresponding coefficients $\beta_j, \{j = 1, 2, \cdots, p\}$. The usefulness of adding extra model terms $X_k^*, \{k = 1, 2, \cdots q\}$, with corresponding coefficients $\beta_k^*, \{j = 1, 2, \cdots, q\}$, to the established model is assessed by calculating the F statistic, which is a ratio of the amount of variance added to the model prediction when including the extra parameters, to a measure of the mean squared error of the augmented model. More formally:

$$F = \frac{SSR_{augmented} - SSR_{established}}{\sum_1^n \left(y_i - \hat{y}_{(augmented)i}\right)^2/(n - p - q - 1)} \tag{4.5}$$

where SSR_{model} refers to sum of squares of each regression (either of the already established model or the new augmented one); $SSR_{model} = \sum_1^n \left(\hat{y}_{(model)i} - \bar{y}\right)^2$, y_i are the observed targets of the model, $\bar{y}$ their mean, and finally $\hat{y}_{(model)i}$ are the model predictions.

The null hypothesis of the partial F-test is that the new variables X_k^* do not significantly improve the prediction capabilities of the model, given that the variables X_j are already included in it. This hypothesis is rejected at a significance level α if the F test statistic is larger than the tabulated F value, $F_{(q,n-p-q-1,\alpha)}$ (from the F distribution with q and $n - p - q - 1$ degrees of freedom).

Returning now to the analysis, the very simple (linear, univariate) model (4.1) does a surprisingly good job of modelling the fluctuation of the lowest modal frequency of

the bridge deck and suggests that the traffic loading is a major driving force at work for this mode. In this case, the model was trained with 2500 samples, which were half-hourly measurements from 27th October to 18th December 2007, and tested on 2500 samples of measurements from 14th May to 5th July 2008. The MSE values were 32.52 and 29.42 for the training and testing set respectively.

Interestingly, the model's prediction capability is seemingly not improved by adding a temperature dependent variable; for the same training and testing period adding a temperature dependent variable reduces the training set MSE very slightly to 29.84, the test MSE is, however, slightly higher at 30.83. Furthermore, if data over a longer period of time are considered in the training set, where one might expect to encounter seasonal effects, temperature still appears to be unimportant to the regression; for a training data set that spans a whole year the MSE of the model only decreases from 31.07 to 30.06 when a temperature dependent parameter is added, suggesting that temperature is not a dominant driving factor. Despite this, when conducting a t-test as described above for the model where a temperature dependent variable is included, the null hypothesis that temperature is an insignificant parameter in the model is rejected at a 99.9% confidence level, suggesting that although temperature may not have a dominant effect on the lowest modal frequency, it still plays some part.

Similarly for the next two frequencies (LS1 and VA1), a linear model of the form (4.1) can predict the main trend of frequency change reasonably (and again perhaps surprisingly) well (see Figures 4.4 and 4.5).

As before, it is interesting to consider how the inclusion of temperature may affect the prediction capability of the models. Unlike the first mode, adding a temperature dependent variable to the model does seem to improve the prediction capability over a longer time period for the second and third modal frequencies. For example, for a training data set spanning several seasons, the model MSE of the third frequency reduces from 60.58 to 50.93 with the addition of a temperature dependent variable to the model. Figures 4.6 and 4.7 demonstrate the difference in model prediction when including an additional temperature dependent variable in the model of the third modal frequency; Figure 4.6 reveals the model prediction of the frequency change when only a traffic load dependent variable is included in the model, Figure 4.7 shows how this prediction changes when a temperature dependent parameter is added. When studying the model predictions it seems that over short time periods the addition of a temperature dependent variable has no visible effect, however,

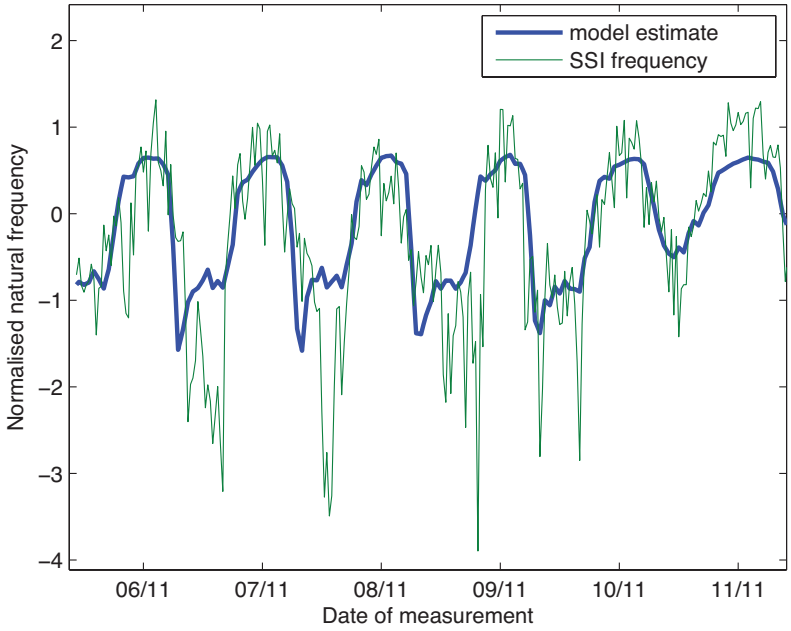


Figure 4.4: Linear model of second deck modal frequency with traffic loading input only

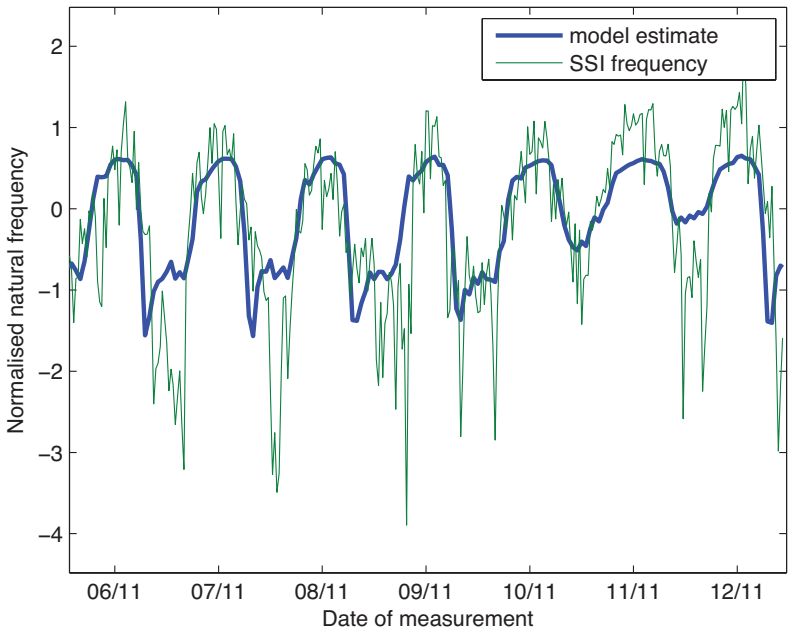


Figure 4.5: Linear model of third deck modal frequency with traffic loading input only

the general fit to longer periods of data appears to be improved, as one can see in Figure 4.7. This suggests that the temperature has more of a seasonal influence than daily, for this mode at least. For both the models of the second and third natural frequency, when including temperature parameters, a t-test can reject a null hypothesis that temperature is unimportant to the regression model with 99.9% confidence.

It should be noted that although a simple model form including temperature and traffic loading inputs can recreate the general trend of the second mode (LS1), there are large daily drops recorded in the second modal frequency which cannot be recreated. These large drops generally occur at times of high traffic, and are rare at weekends. The current hypothesis is that these large drops are caused by short term large traffic loadings, such as would result from a traffic jam. Unfortunately the traffic loading estimates have to be interpolated from hourly traffic counts, and as such cannot predict short term traffic loads. It is expected that a more sophisticated traffic loading estimate would improve model fidelity for the second mode considerably.

Similar model fits can predict the general trends in the fluctuations of the fourth and fifth frequencies (LA1 and TA1), however, the prediction errors are comparably large. A more complex model structure will be needed to accurately predict and therefore better understand the changes in the fourth and fifth modal frequencies.

4.1.1 Deck Acceleration

Closely linked to the wind profile, and also to the dynamics of traffic loading is the acceleration of the deck. Figure 4.8 shows plots of the first five modal frequencies with respect to the root-mean-squared (RMS) values of vertical and horizontal deck acceleration, linear best line fits have, again, been added purely to aid visualisation of the trends. These plots show a clear tendency for decreased frequencies at higher amplitudes of deck acceleration, for both horizontal and vertical accelerations. This amplitude dependency indicates that the system is nonlinear, which is not unexpected for such a complex structure. Indeed, Zhang *et al.* report that the dynamic behaviour of all cable-supported bridges is amplitude dependent [50].

This discovered nonlinearity does not in fact overly increase the complexity of the response surface analysis carried out here, it rather just increases the number of pa-

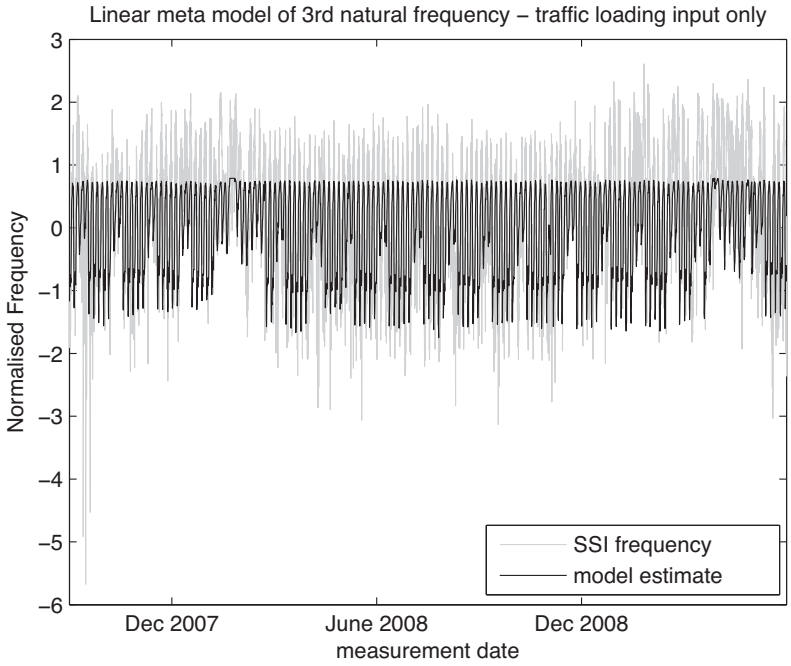


Figure 4.6: Linear model of the third deck modal frequency with traffic loading input only

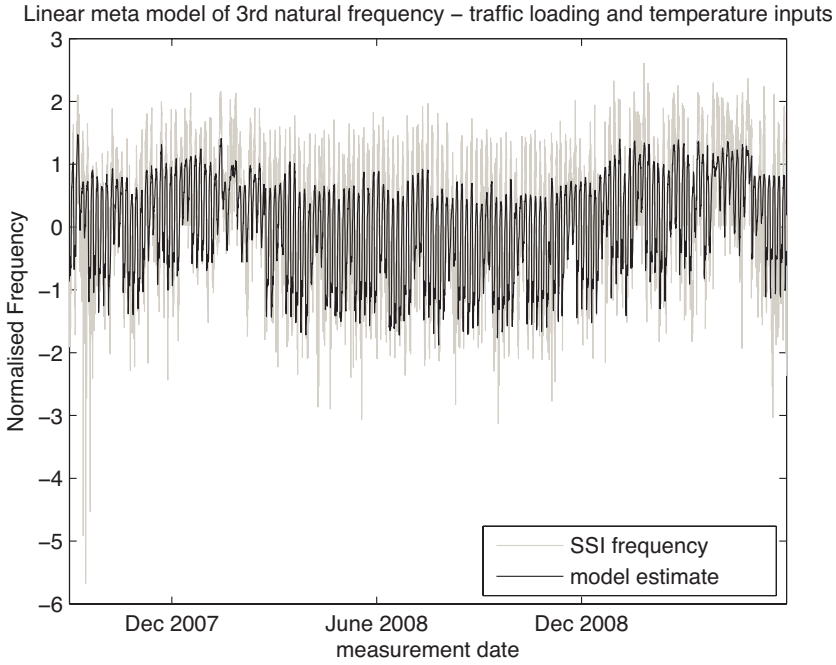


Figure 4.7: Linear model of third deck modal frequency with traffic loading and temperature inputs

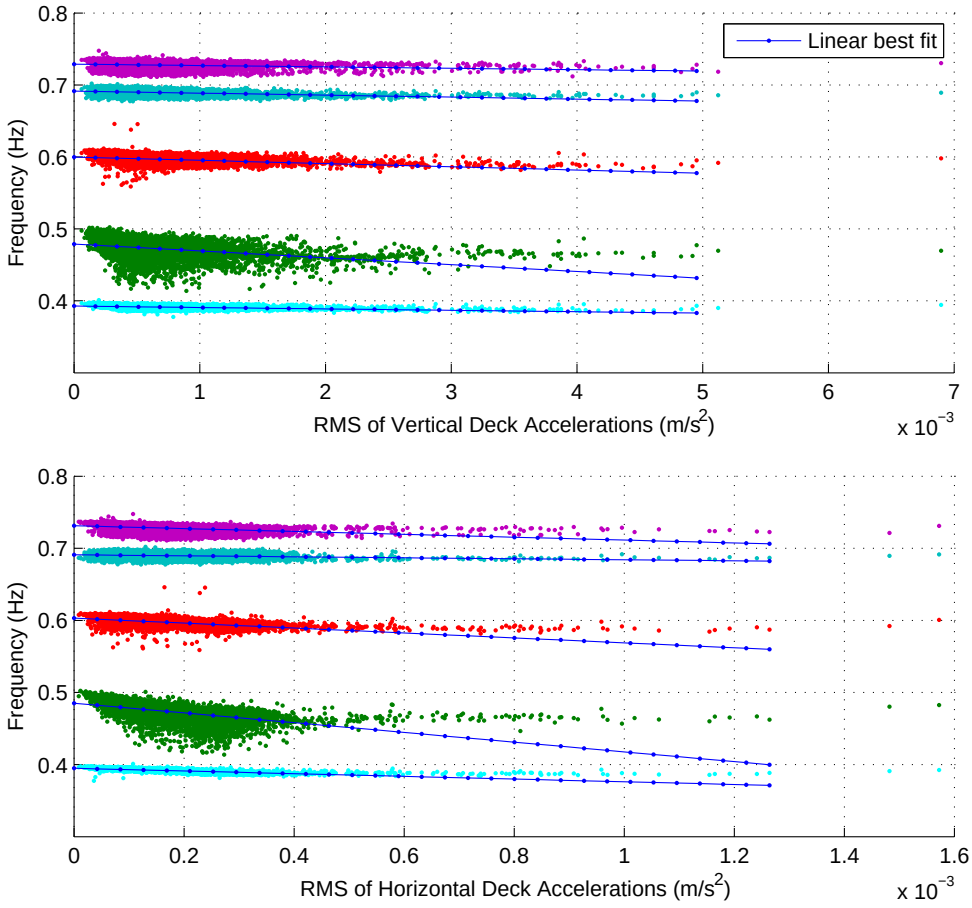


Figure 4.8: Modal frequency plotted according to RMS of vertical deck accelerations (above) and horizontal deck accelerations (below)

rameters that must be considered when attempting to understand, or even predict, the fluctuations in the modal frequencies. This having been said, on close inspection of Figure 4.8, the correlations between modal frequency and deck acceleration appear non-trivial, especially for the second modal frequency; this will require further investigation if the relationship between the two is to be well understood.

Figure 4.8 demonstrated that increases in the horizontal and vertical accelerations of the deck correspond to a decrease in the modal frequencies of the first five modes. Returning to the simple response surface models for predicting frequency change, the effect of adding a variable dependent on the vertical RMS deck acceleration is now investigated. Figure 4.9 demonstrates the improvement of the model prediction when adding an acceleration dependent variable for the first frequency over a time when high wind speeds were recorded. To understand the role of the wind speed, and therefore hopefully better understand the relationship between deck acceleration

and modal frequencies, the first sensible step is to study the deck accelerations themselves.

Figure 4.10 is composed of plots of wind speed against vertical and horizontal deck acceleration (RMS), the plot points are also sorted according to the direction of the wind at the time. From this figure, two clear response mechanisms can be seen; when the wind is from the east or west, there is no increase in response with increased wind speed, conversely when the wind is from the north or south, i.e. normal to the bridge span, above 25 mph the deck acceleration response increases (from inspection - nonlinearly) with increased wind speed. As the bridge is orientated east-west, the increasing response with increased wind speed occurs, not surprisingly, when the wind hits the bridge side on. This bi-functional relationship must be considered with any attempt to model the bridge's behaviour with respect to deck acceleration. The effect of traffic on the deck acceleration should also be considered (see Figure 4.11); here, the RMS of vertical and horizontal deck acceleration increases linearly with increased traffic load.

Having now a better understanding of the deck acceleration, one can return to the relationship between deck acceleration and modal frequency. Although the acceleration response of the deck acts in two different regimes according to wind direction, it does not necessarily follow that this should be reflected in the relationship between deck acceleration and modal frequency; it is possible that one regime could define the acceleration-frequency relation. Figures 4.12 and 4.13, however, show a closer view of selected plots from Figure 4.8, sorted according to wind speed (and coloured according to wind direction), which clearly demonstrate a more complex relation between the two variables than perhaps expected. On inspection of Figures 4.12 and 4.13, there generally appears to be two different trends roughly separable by wind speed and direction, namely, the frequencies appear to act under a different regime when high wind speeds from a southerly direction are recorded.

4.2 Mathematical Models of Modal Frequencies

Based on the above analysis, more complex models to predict modal frequency change can now be contrived. The primary reason for the development of such models here is for a better understanding of the mechanisms at work as the structure responds to normal environmental and operational conditions. However, a

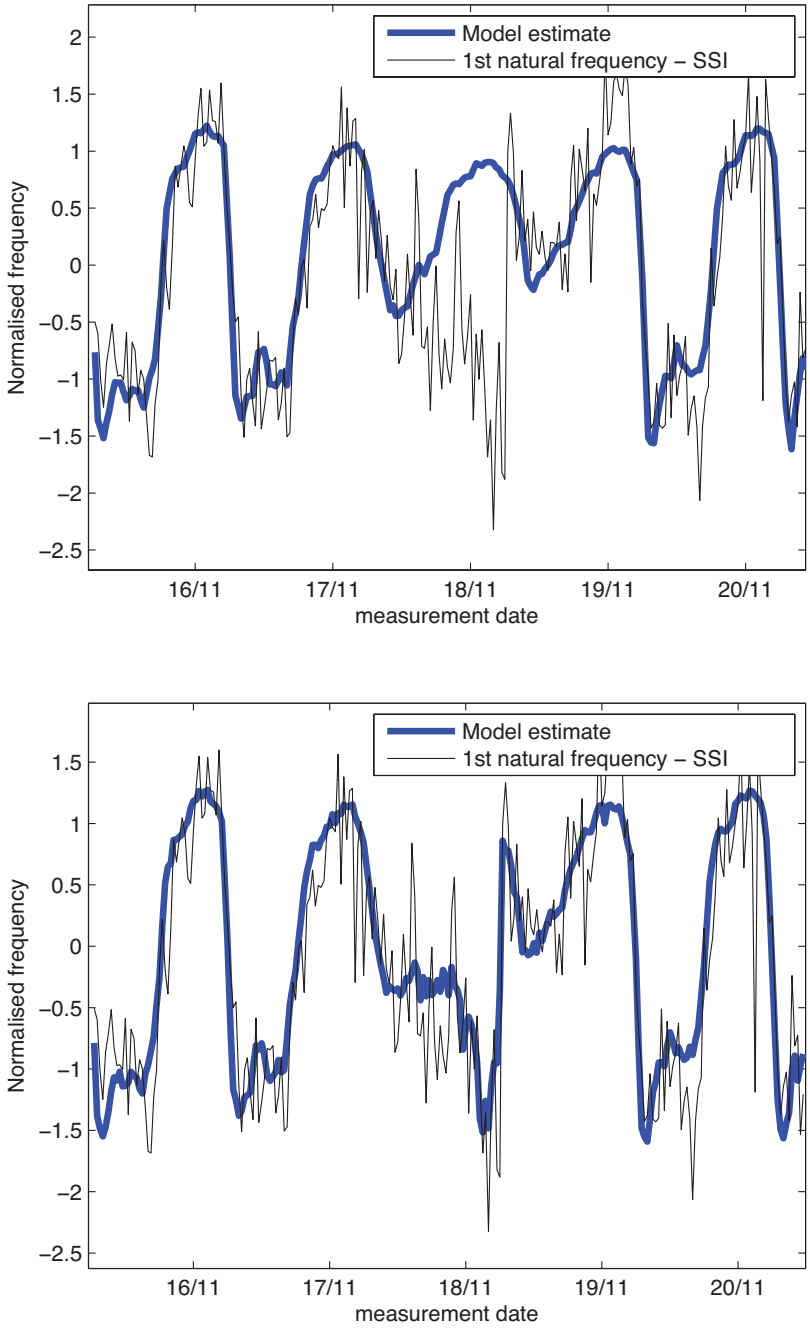


Figure 4.9: Above: linear model of the third deck modal frequency with traffic loading input. Below: linear model of third deck modal frequency with traffic loading and temperature inputs.

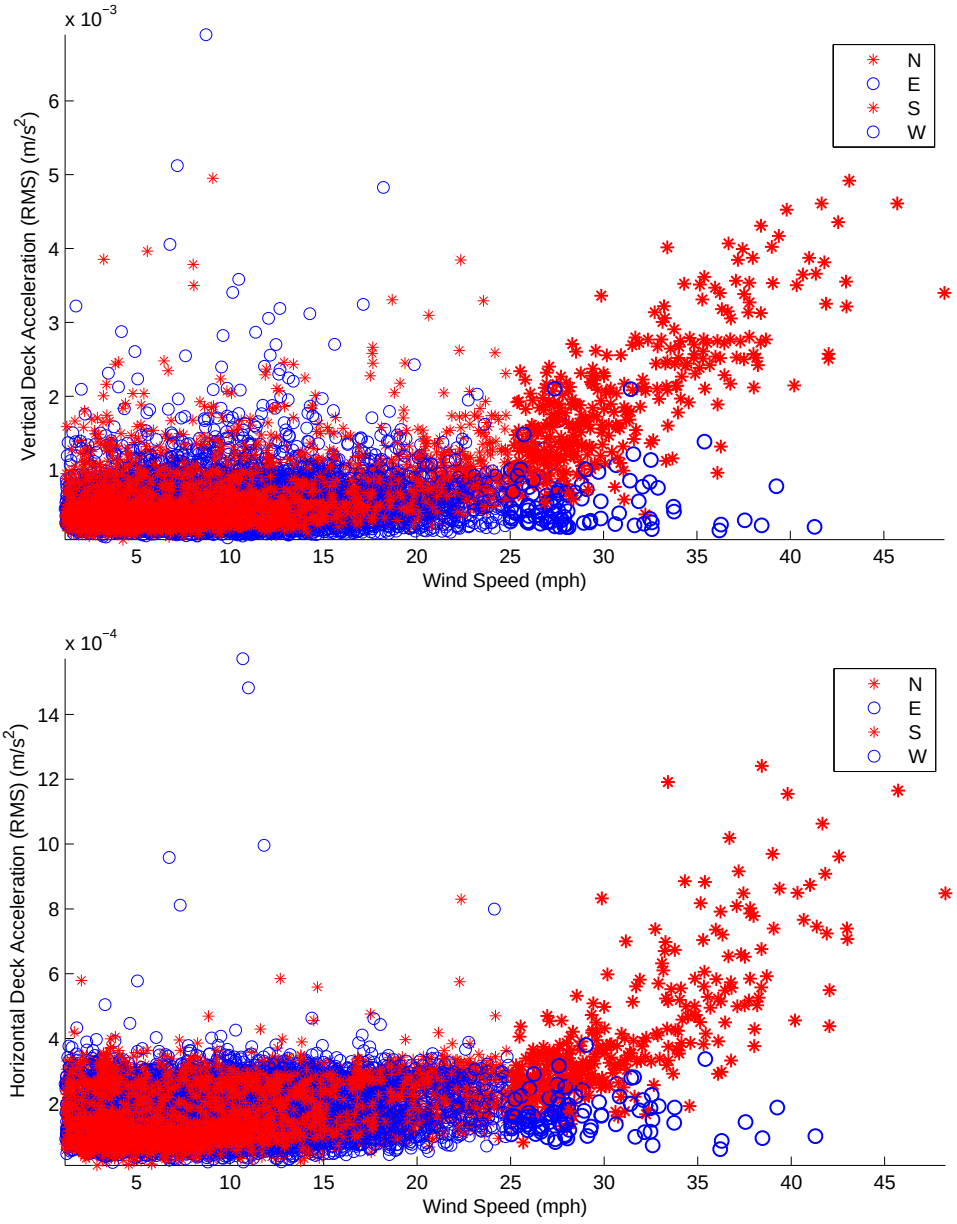


Figure 4.10: Vertical and horizontal deck acceleration plotted with respect to wind speed and sorted by wind direction

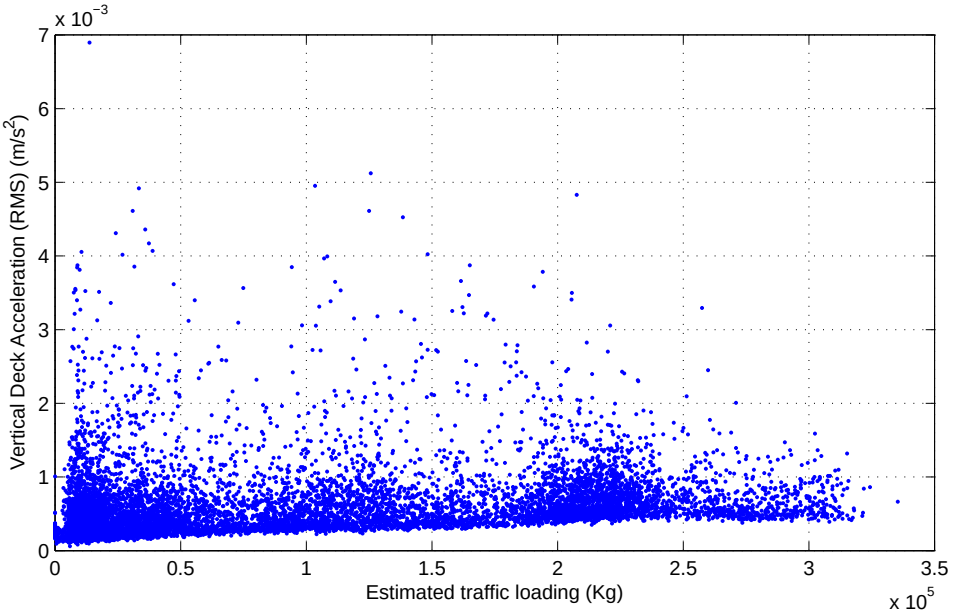


Figure 4.11: RMS of vertical deck acceleration plotted according to estimated traffic loading

secondary motive, which is also an important one, is that any reliable predictive models will, understandably, be of great use as a condition indicator. From the knowledge gained in the previous sections, useful inputs to any model of the deck modal frequencies would include parameters dependent on traffic loading, temperature and also horizontal and vertical deck acceleration. Indeed, inputs based on deck acceleration should take into account the two possible response regimes discovered, which occur most likely because of differing wind patterns. Furthermore, more complex additional parameters may be considered to reflect any nonlinearity in the response. Time-lagged parameters may also be added to account for any dynamic relations between variables, which, for example, would come into play if the modal frequency at any one time depended, say, on the temperature at that same time and also the time(s) preceding it. Response surface models (see (4.1), for example) will be used here again. One of the main advantages of polynomial response surface models is their simplicity; they are easily fitted using least-squares methods, and they are very easy to interpret, as coefficient values can indicate the significance of a parameter (as long as input variables are normalised prior to use). Alternative predictive modelling approaches such as neural networks and support-vector machines have previously been explored in the literature for monitoring data from other long-span bridges, [55, 57], although in a slightly different context. In these papers, sophisticated modelling techniques have been used with the aim of

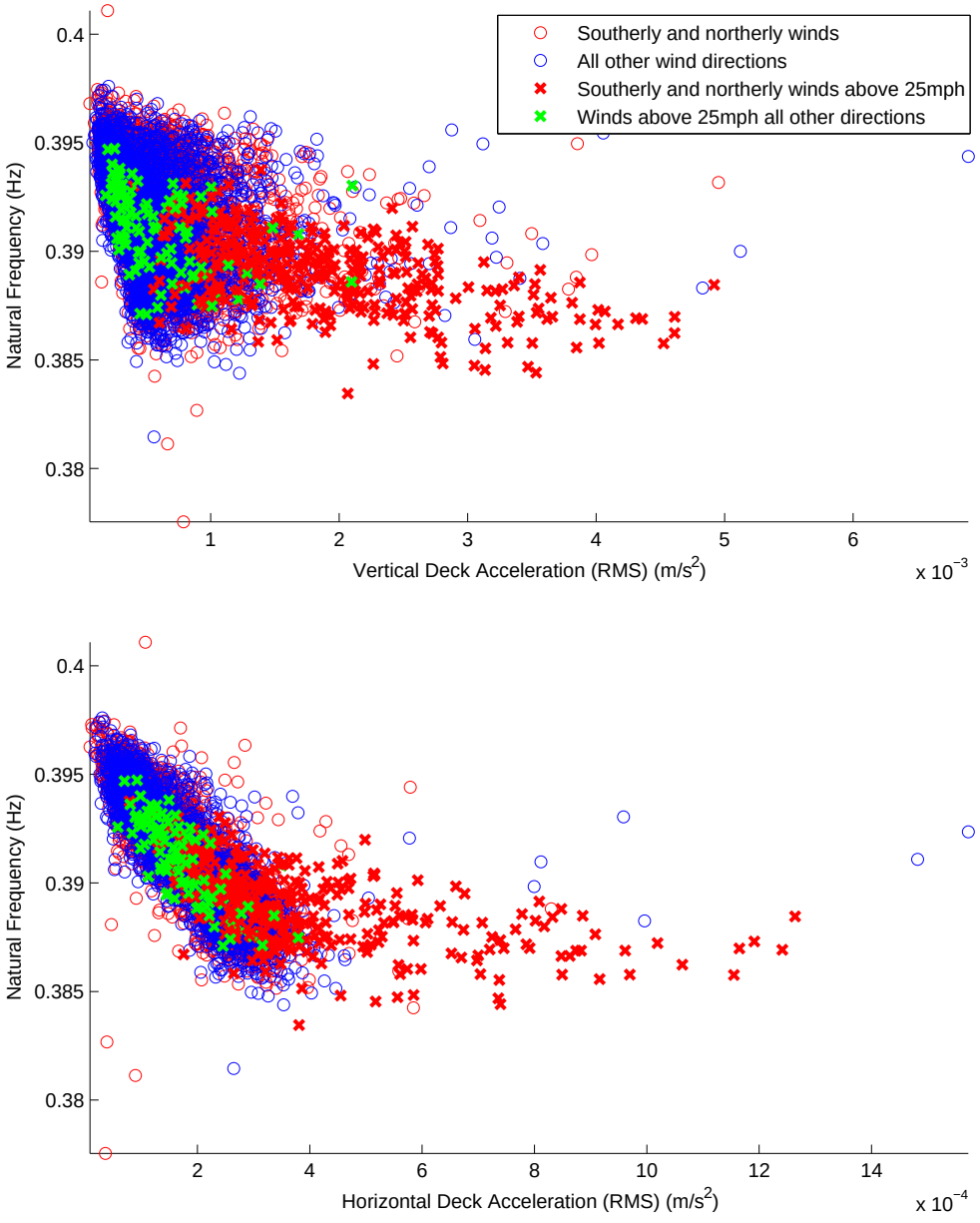


Figure 4.12: Variations of first modal frequency plotted according to RMS of deck acceleration

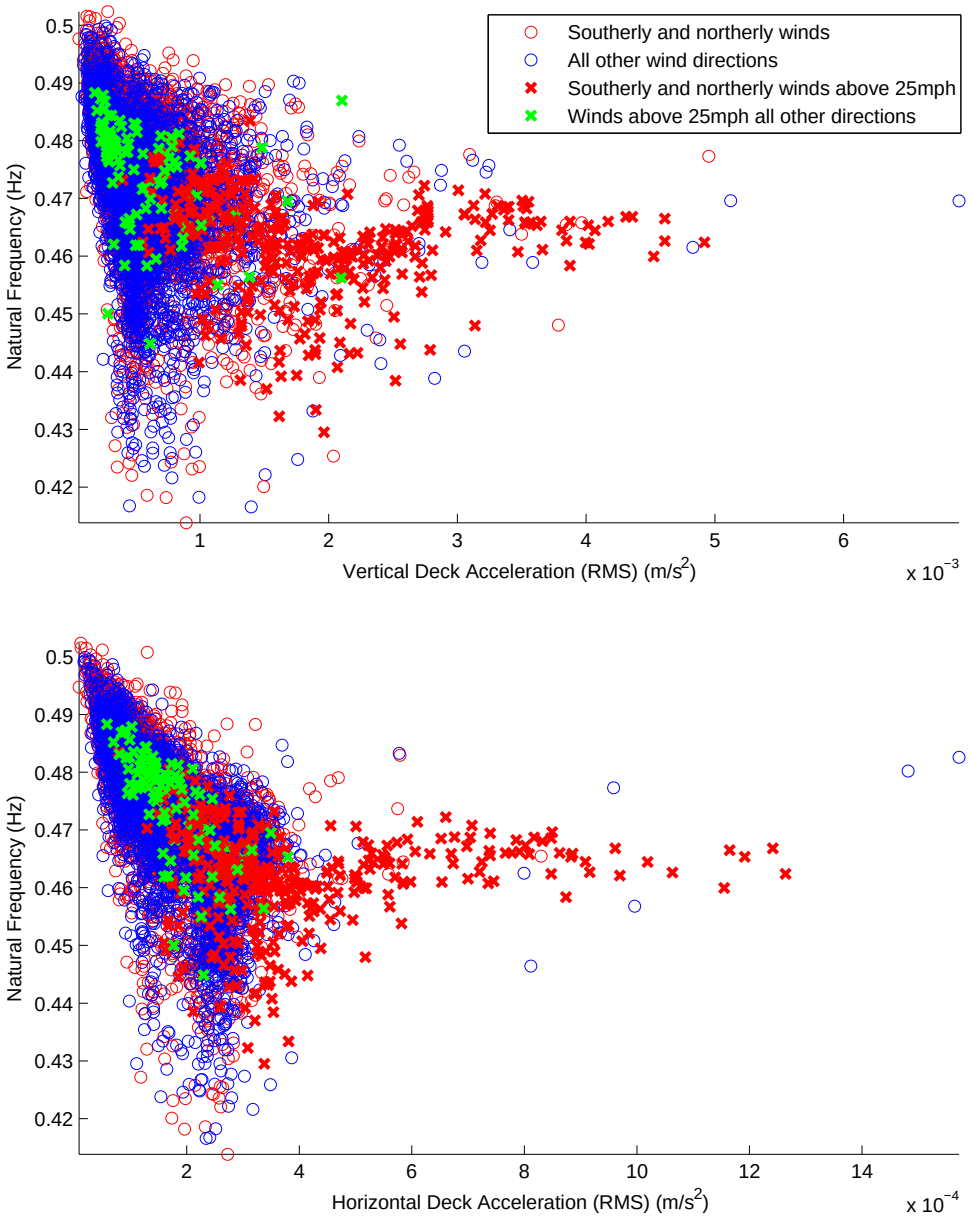


Figure 4.13: Variations of the second modal frequency plotted according to RMS of deck acceleration

developing diagnostics that can separate changes in modal parameters caused by environmental variations from those caused by damage. These methods are known to have powerful prediction capabilities, however, no knowledge of the physical system can be gained directly from these non-parametric approaches. As the primary focus of this work is to better understand the interaction between environmental and operational conditions and measured structural response, the simpler, more interpretable response surface models are used here. Furthermore these more complex techniques will often require much larger quantities of training data and are much more computationally expensive.

Here, to begin, response surface models will be fitted for the first and second modal frequencies of the deck, which correspond to a symmetrical vertical mode and a symmetrical lateral mode respectively. As previously, the normalised mean squared error will be used as a performance indicator for the goodness of fit of the response surface models and a t-test will be used to assess the significance of each parameter coefficient. In the following, the response surface models are trained on 5922 data points (data collected over the period of a year) and tested on an independent data set of 3735 data points. Initially six input parameters for the response surface models will be considered. Reflecting the above analysis, these will include variables dependent on:

- **T:** temperature
- **TR:** traffic loading
- $\mathbf{VA}_W$**:** vertical deck acceleration occurring at times when high wind speeds hitting the deck side on (i.e. from the north or the south) are recorded (zero at all other times)
- $\mathbf{HA}_W$**:** horizontal deck acceleration occurring at times when high wind speeds hitting the deck side on (i.e. from the north or the south) are recorded (zero at all other times)
- $\mathbf{VA}_N$**:** vertical deck acceleration in all other wind conditions
- $\mathbf{HA}_N$**:** horizontal deck acceleration in all other wind conditions.

Before attempting to assess the importance of each of these parameters to a model of modal frequency, however, it is important to see how these input variables interrelate. If any are highly correlated then inference on their importance in a regression

model may be misleading, this is referred to as the problem of multicollinearity (again see any good book on applied statistics or regression for more details). To investigate correlations in the input parameters listed above their correlation matrix is calculated for data in the training period and shown in Table 4.1.

	T	**TR**	$\mathbf{VA}_W$	$\mathbf{HA}_W$	$\mathbf{VA}_N$	$\mathbf{HA}_N$
T	1.0	0.24	0.12	0.10	0.17	0.20
TR	0.24	1.0	-0.05	0.01	0.31	0.82
$\mathbf{V}A_W$	0.12	-0.05	1.0	0.91	0.03	0.02
$\mathbf{HA}_W$	0.10	0.01	0.91	1.0	0.03	0.02
$\mathbf{VA}_N$	0.17	0.31	0.03	0.03	1.0	0.35
$\mathbf{HA}_N$	0.20	0.82	0.02	0.02	0.35	1.0

Table 4.1: Correlation matrix of potential six input variables for response surface models

Studying this table, it is evident that horizontal and vertical deck acceleration at times of high wind speeds are highly correlated. Interestingly, traffic loading and horizontal deck acceleration under normal wind conditions are also highly correlated. To avoid issues with multicollinearity, horizontal deck accelerations will not be included as input parameters to the response surface models. Any variation due to horizontal deck acceleration will be equally well accounted for by the traffic loading and vertical deck acceleration parameters in the implemented models.

When fitting a response surface model with the remaining four inputs, the MSE of the model for the first mode was 22.6 for the training data set, and 24.8 for the test set. When conducting a t-test to check the significance of each of the parameters included in the basic model, all parameter coefficients were found to be statistically significant at a 0.05 significance level. A model with these input parameters of the second modal frequency could not perform as well and had a training MSE of 49.7 and a MSE of 51.1 on the test data. This was due, as explained previously, to the fact that large drops occur in the time history of the second modal frequency that the model cannot recreate, which are thought to be caused by traffic patterns. For the second frequency, when using a t-test, again all four parameters considered were found to be statistically significant at a 0.05 significance level.

As previously mentioned, the addition of dynamic and nonlinear elements to these models may prove beneficial. It may be, for example that a past temperature measurement has an influence on the current frequency. When studying Figure 4.12, one can see that a quadratic function may be pertinent for accounting for the frequency

change at times when high winds hit the side of the bridge.

To investigate the incorporation of dynamic parameters into the response surface models, time lagged versions of the four basic parameters identified above were considered. However, it was found that, understandably, the four baseline parameters are highly correlated with the lagged versions of themselves. In order to avoid issues with multicollinearity a first difference of each parameter is considered instead.

Quadratic terms for each of the variables are also considered. The baseline variable that describes vertical deck acceleration when strong winds hit the bridge side on is again highly correlated with the quadratic version of itself (most likely due to the fact that it is zero the majority of the time). For this reason a quadratic form of this variable will be considered separately.

A response surface model will therefore be fitted with the following input variables considered:

- **T:** temperature
- **TR:** traffic loading
- $\mathbf{VA}_W$**:** vertical deck acceleration occurring at times when high wind speeds hitting the deck side on (i.e. from the north or the south) are recorded (zero at all other times)
- $\mathbf{VA}_N$**:** vertical deck acceleration in all other wind conditions
- Δ**T:** first difference of temperature
- Δ**TR:** first difference of traffic loading (value at time t - value at time $t-1$)
- $\Delta\mathbf{VA}_W$**:** first difference of vertical deck acceleration during strong winds hitting the deck side on
- $\Delta\mathbf{VA}_N$**:** first difference of vertical deck acceleration in all other wind conditions
- $\mathbf{T}^2$**:** temperature squared
- $\mathbf{TR}^2$**:** traffic loading squared
- $\mathbf{VA}_N^2$**:** vertical deck acceleration in normal wind conditions squared

Each of the additional variables was added onto the basic four parameter model

	Independent/Input Variables	**Parameter Mode 1**	**Parameter Mode 2**
$\mathbf{T}$	Temperature	-0.03	-0.17
$\mathbf{TR}$	Traffic Loading	-0.84	-0.59
$\mathbf{VA}_W$	Vertical Acceleration (strong winds from North/South)	-0.39	-0.31
$\mathbf{VA}_N$	Vertical Acceleration (normal wind conditions)	-0.15	-0.21
$\mathbf{\Delta T}$	First difference temperature	-	0.21
$\mathbf{\Delta TR}$	First difference traffic loading	0.25	0.13
$\mathbf{\Delta VA}_W$	First difference of vertical acceleration (strong winds from North/South)	0.10	-
$\mathbf{\Delta VA}_N$	First difference of vertical acceleration (normal wind conditons)	0.04	0.06
$\mathbf{T}^2$	Squared Temperature	0.02	-
$\mathbf{TR}^2$	Traffic Loading Squared	0.07	0.08
$\mathbf{VA}_N^2$	Vertical Acceleration (normal wind conditions) squared	0.02	0.02

Table 4.2: Parameter coefficients for models predicting the first and second modal frequencies

separately and its individual contribution assessed using a partial F-test. When considering the first modal frequency, each of these additional parameters was found to make a statistically significant contribution to the response surface model when applying a partial F-test with the exception of the first difference of temperature (ΔT). For the model of the second modal frequency, of these additional parameters the first difference of the vertical deck acceleration at times of strong winds ($\Delta \mathrm{VA}_W$) and the variable describing temperature squared (T^2) were both found by a partial F-test not to provide a statistically significant contribution to the model.

The coefficients of the final models for the first and second modal frequencies which include all parameters judged to make a statistically significant contribution are presented in Table 4.2. As each input variable is normalised (actually standardised to zero mean and unit variance), the model parameters can be (roughly) interpreted as an importance value, the higher the value, the more influence that the term has on the modal frequency change.

Table 4.2 suggests that the traffic loading parameter is most influential to both the first and second modal frequencies, with vertical deck acceleration at times of strong northerly and southerly winds second most important. As expected from the anal-

ysis above (see Figure 4.7), temperature plays a much more dominant role in the prediction of the second modal frequency than the first. Interestingly, for the first mode the first difference of the traffic loading variable shows some significance, indicating that the change in mass helps to predict the frequency. For the second mode it is the first difference of the temperature variable that shows some significance, indicating that thermal gradients may influence the frequency change. Once again this confirms the increased dependence on temperature of the second mode.

It is also interesting to note which variables are not influential to the modal frequency. Quadratic parameters do not have large coefficient values and can, therefore, be considered not important. This is most likely because most input variables have a linear effect, but may also be because any true nonlinearity cannot be represented suitably by a quadratic form.

From the above analysis it seems most likely that the vertical acceleration of the deck at times of strong winds is nonlinearly related to modal frequency, however a quadratic term for this variable was not added to the response surface model to avoid issues with high correlation between the linear and nonlinear term. Instead, using a squared term of the acceleration is investigated separately here. When *replacing* VA_W with its square in the model fitted above (whose parameter coefficients are described in Table 4.2), the MSE of the model fit increases from 23.57 to 26.16. When the squared parameter is used its coefficient is -0.06 which is much lower than the coefficient attributed to the original variable (-0.39 from Table 4.2). Interestingly, when the squared variable is used $\Delta\mathrm{VA}_W$, the parameter describing the first difference of the vertical acceleration in strong northerly/southerly winds becomes statistically insignificant at a 0.05 significance level.

To study the different model fits at times of strong northerly/southerly winds, the model predictions are plotted in Figure 4.14, along with the prediction of a response surface model with just temperature and traffic input parameters, as plotted in Figure 4.9, for comparison. Studying this figure one can see that the model with the squared acceleration values can predict the trough of the frequency more accurately, but performs worse than the model with the linear variable before the large trough. It seems that overall replacing the deck acceleration variable, VA_W, with its square is not beneficial. In light of the relations highlighted in Figure 4.12, this is perhaps surprising. However, it is likely that model fit would improve with a more complex form to model the interaction between the deck acceleration and modal frequency at high wind speeds. It must also be noted that training data at times of higher

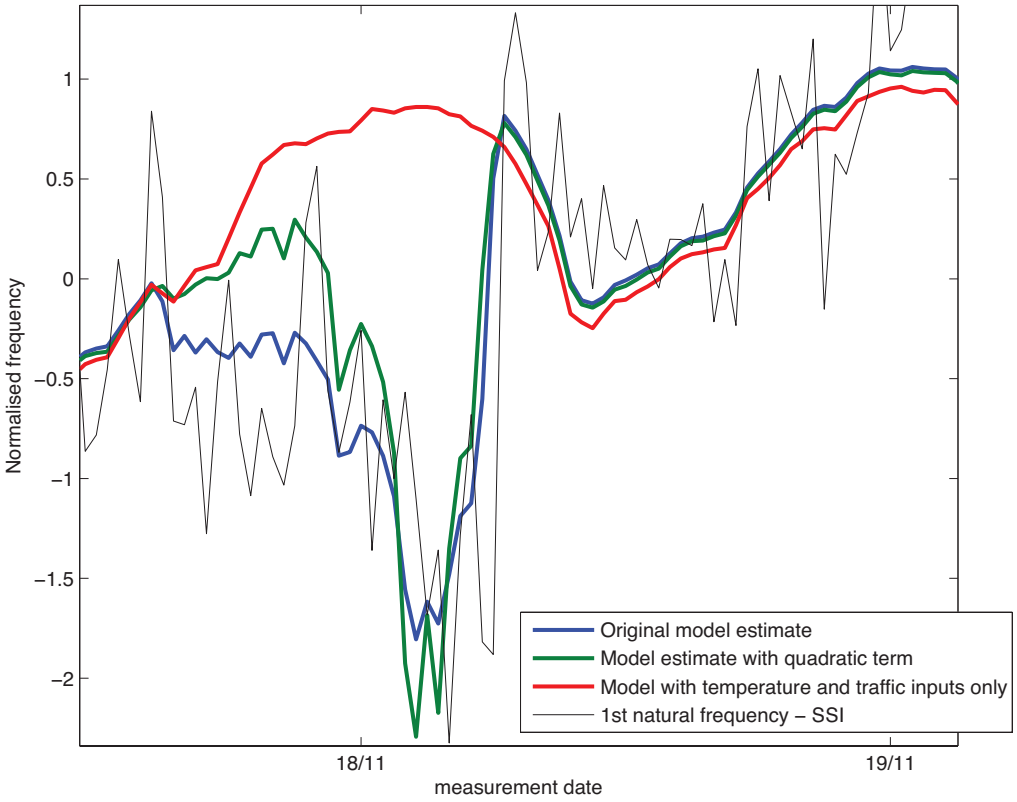


Figure 4.14: Model predictions of the first modal frequency when including a squared parameter of vertical deck acceleration at times of strong northerly/southerly winds

wind speeds are rare, it is anticipated that the model fit would improve with a larger training set incorporating more data occurring in these wind conditions.

4.3 Conclusions

The current chapter has introduced data from the Tamar monitoring campaign. The work covered has addressed which environmental/operational conditions drive the fluctuations observed in the modal frequencies of the deck obtained from acceleration data by a data-driven SSI routine. Traffic loading was found to be a dominant driver of daily frequency fluctuation, whilst temperature was found to have more of a seasonal effect than daily. Lastly, the acceleration of the deck was found to have a significant effect on the modal frequencies at times when the wind speed was higher

than 25mph and hitting the bridge side-on.

Finally, response surface models have been fitted in attempt to predict the modal frequency changes of the bridge deck given the measured environmental/operational conditions. It was found that a simple response surface model with input variables based on the estimated traffic loading, temperature and deck acceleration (in turn dependent on the wind speed and direction) can predict the change in the first modal frequency to a good degree of accuracy. The higher modal frequencies can also be predicted with similar models, although with less accuracy. In the next chapter, how these models may be utilised for SHM will be explored further.

Chapter 5

Towards SHM for Tamar

In the previous chapter simple regression models were employed in an attempt to better understand the normal condition of the Tamar Bridge. An important aim of the research project funded by EPSRC that this thesis forms part of is to be able to identify anomalies in the response data that may relate to performance or structural condition. This chapter addresses how this might be possible for the Tamar monitoring campaign and considers the natural frequency and displacement measurements of the deck.

5.1 Novelty detection for Tamar

As discussed in the introductory chapter, a major pitfall for practical SHM implementation, from a machine learning perspective, is the lack of data available from the 'damaged state' of structures, which often necessitates an unsupervised learning approach. Due to this, the wish to identify anomalous responses lends itself directly to the idea of *novelty detection*, where baseline data is used to define a 'normal' response, and new measurements are compared to this baseline in order to assess whether a structure continues to respond within this normal condition.

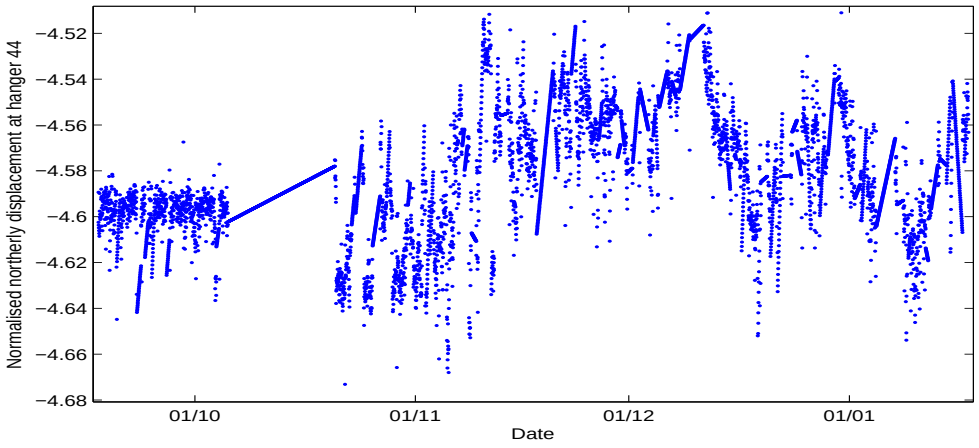
The use of novelty detection for performance and structural health monitoring is complicated by the fact that most measurements or structural responses of interest are influenced by changing environmental and operational conditions, as evidenced in the previous chapter. Practical application of novelty detection often works on

the premise that some monitored damage sensitive feature will remain stationary, or within some limits, all the time a structure continues to respond in its normal condition. The occurrence of damage is then inferred by any significant change occurring in the feature. If responses driven by, for example, a temperature differential, are not defined within the normal condition, then a novelty detection process will wrongly assign these responses as anomalies.

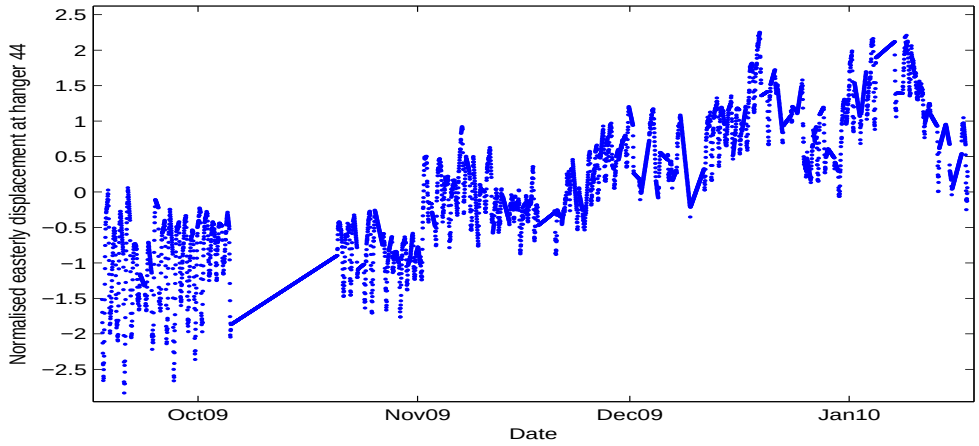
When undertaking any task that requires judgement as to whether a structural response is typical or anomalous, knowledge as to what constitutes a normal response will always be of benefit. In the previous chapter, the author investigated how the modal frequencies of the deck are influenced by environmental and operational conditions. The findings were that the modal frequencies of the deck are most influenced by the traffic loading and acceleration of the deck. Temperature and wind profile also influence frequency change, in particular, the wind has a strong effect when hitting the side of the bridge at high speeds. Along with modal frequencies, this chapter also considers deflection measurements of the deck and tower from the TPS system described previously.

The TPS system installed at the Tamar Bridge is the newest system in the monitoring campaign, the available data therefore spans a much shorter time than the data analysed in the previous chapter. The data available for analysis here are half hourly readings of displacements in a northerly, easterly and vertical direction from 17th September 2009 -17th January 2010. To aid visualisation of the available data, plots of the normalised northerly, easterly and vertical displacement measurements at hanger 44 (see Figure 3.6) over the four month period are shown in Figure 5.1. Studying Figure 5.1, one can firstly notice a large gap in the data, also visible is an increasing trend for the easterly and vertical displacements as time progresses. The northerly displacement also appear to have greater variability in the winter months.

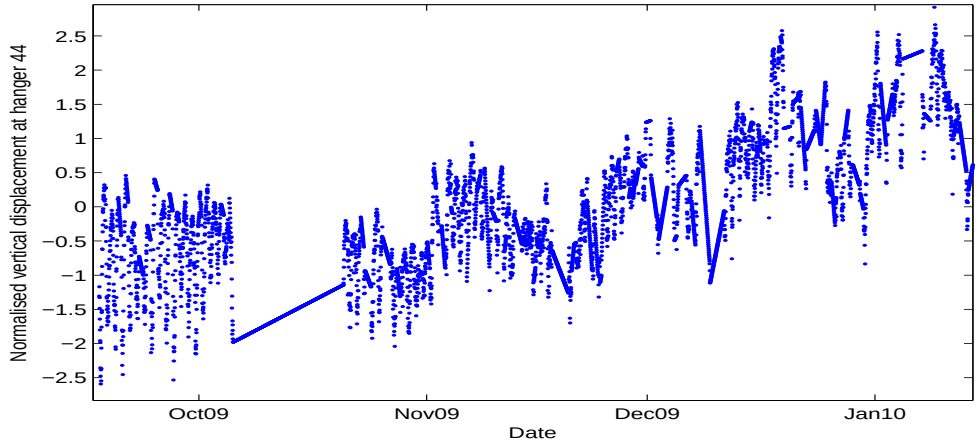
From similar analysis to that carried out in the previous chapter, it has been found that the measured easterly and vertical deflections correlate most highly with temperature (as shown in Figure 5.2, where displacements measured at hanger 44 are plotted against temperature). Vertical and easterly deck displacements also show some correlation with the estimated traffic loading, although with large variance.



(a)

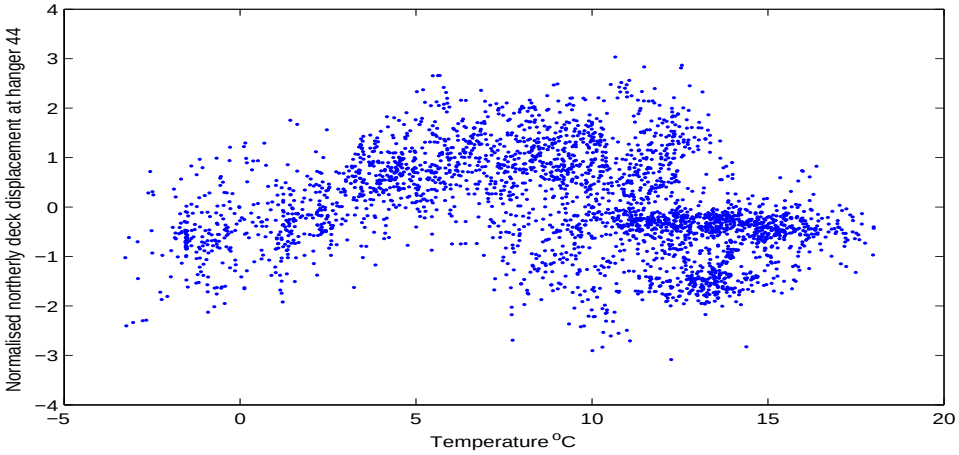


(b)

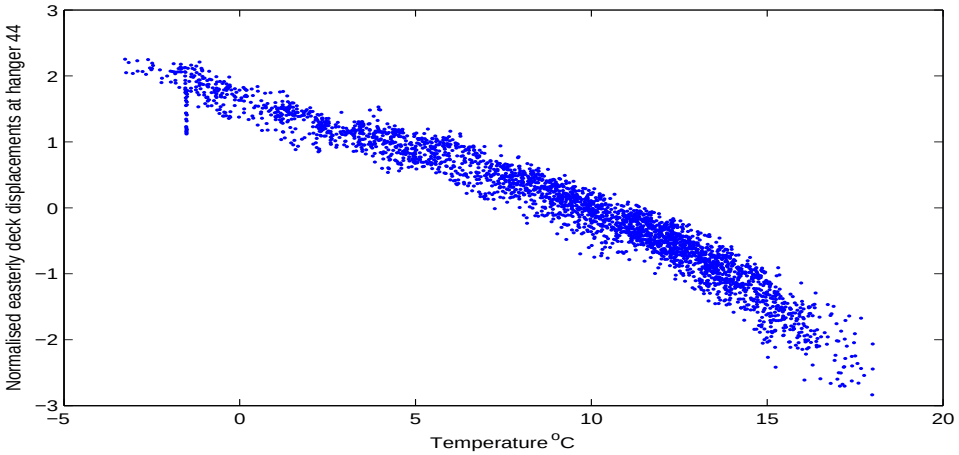


(c)

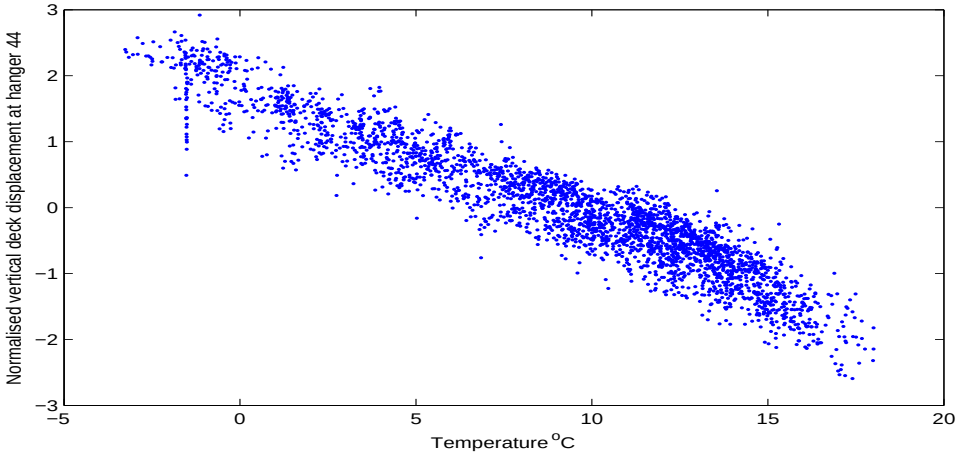
Figure 5.1: Normalised displacements at hanger 44 over the four month monitoring period, (a) in a northerly direction, (b) in an easterly direction and (c) in a vertical direction



(a)

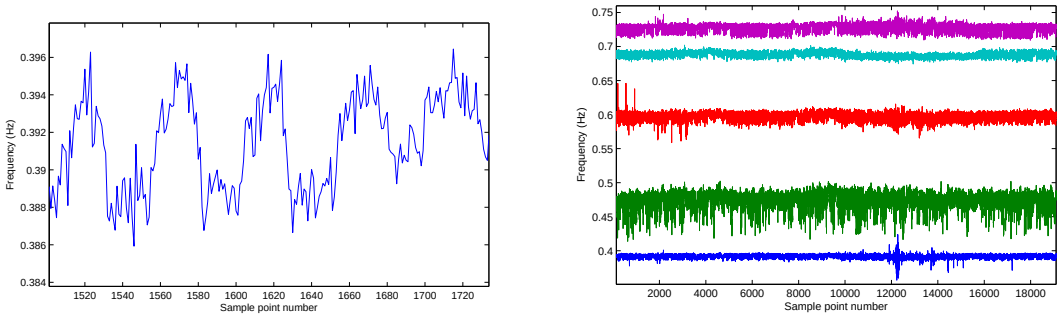


(b)



(c)

Figure 5.2: Northerly (a), easterly (b) and vertical (c) deck displacement plotted against temperature



(a) (b)

Figure 5.3: (a) Lowest deck natural frequency over a five day period (b) Lowest deck natural frequency captured over a 28 month period

5.2 Detecting novelty with extracted natural frequencies

Figure 5.3 shows two plots of the lower modal frequencies of the Tamar Bridge over two different time scales. In Figure 5.3(a) the plot shows the variation of the lowest frequency over five days, plot (b) shows the lowest five frequencies captured over a period of 28 months. From plot (a), the daily variation of the frequencies is obvious. However, although the presence of seasonal trends were demonstrated for the second mode in the previous chapter, it can be noted from (b), that, visually, none of the frequencies show obvious seasonal trends over the 28 month period. When considering long time periods (more than a year), these frequencies are stationary.

Because of this long term stationarity, novelty detection can be directly applicable without prior manipulation of the frequency data to account for the environmental/ operational trends. To demonstrate this, an outlier analysis using the Mahalanobis squared distance is carried out on the five lowest natural frequencies of the bridge deck.

Outlier analysis calculates a measure of how similar or dissimilar a sample of feature data is to other samples, this measure is called discordancy. An outlier is a (uni- or multi-variate) sample which has a large discordancy measure in comparison to the majority of the other samples in a given data set. The magnitude of the discordancy over which a sample is classified as an outlier is determined by some threshold. In this work, where multivariate features are under consideration the Mahalanobis squared

distance, shown in equation (5.1), will be used to provide a measure of discordancy D_i for each sample of data x_i.

$$D_i = (\{x_i\} - \{\bar{x}\})^T [S]^{-1} (\{x_i\} - \{\bar{x}\}) \tag{5.1}$$

In the above, $\bar{x}$ is the sample mean of some training set of observations, and S the corresponding covariance matrix. In order to label an observation as an outlier or an inlier there needs to be some threshold value against which the discordancy value can be compared. This value is dependent on both the number of observations and the number of dimensions of the problem being studied. The value also depends upon whether an inclusive or exclusive threshold is required. In this work, the threshold value is computed using a Monte Carlo method. Briefly, a matrix the same size as the data set under consideration is generated and populated with elements randomly drawn from a zero mean, unit standard deviation Gaussian distribution, for all elements the Mahalanobis squared distance is then calculated and the largest value stored. This is repeated a large number of times (10,000 times in the case of this work), each time storing the largest Mahalanobis squared distance, which are then sorted in order of magnitude. The critical values for 5% and 1% tests of discordancy can then be found from this array above which 5% and 1% of the trials occur.

In the following, the five lowest natural frequencies of the bridge deck at a given time will be considered as a single multivariate feature. The sample mean and covariance in equation (5.1), as well as the threshold are calculated from a training set of data of 1000 time samples corresponding to the first 1000 samples visible in Figure 5.3(b), this is approximately one month's worth of data. By choosing this training set, one is essentially defining with it what constitutes a normal response, against which all other measurements will be compared. The discordancy of the five natural frequencies will be calculated for the same 28 month period as shown in Figure 5.3.

Figure 5.4 shows the results of the outlier analysis described, with the (95% confidence) threshold plotted as a black (dash and dot) line. Studying Figure 5.4, one can see that the majority of discordancy measurements remain under the threshold for the 28 month period, indicating a normal response for the duration. Some outliers are visible in the earlier part of the record which may be further investigated as potential performance anomalies, the most obvious anomalous event, however, occurs

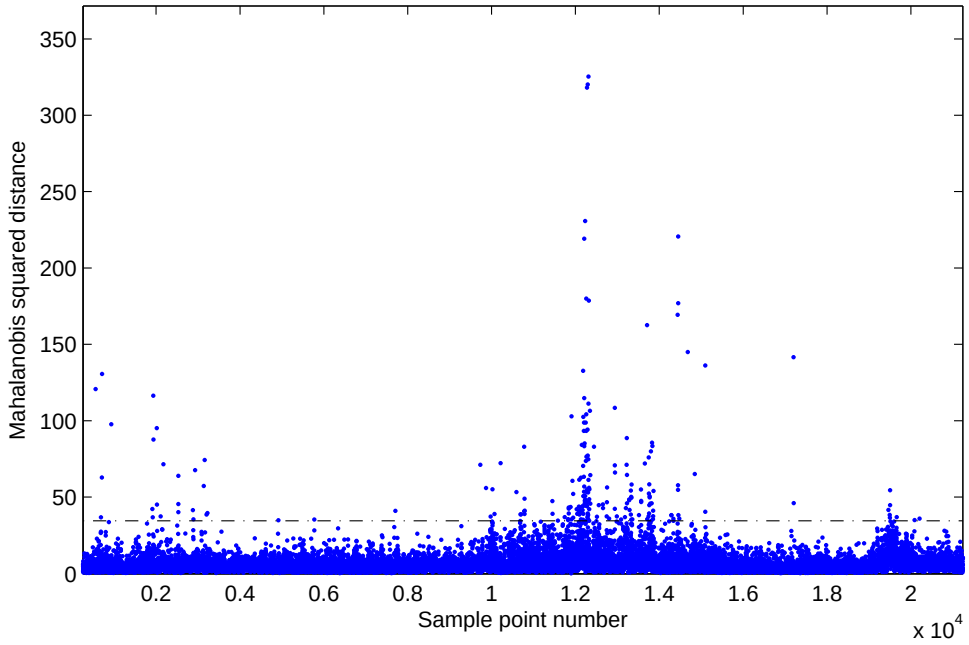


Figure 5.4: Outlier analysis on the five lowest natural frequencies of deck

in the latter third of the time period where a large number of outliers are visible. This excursion corresponds to a visible 'blip' in the lowest mode in Figure 5.3(b). On further investigation, it was found that at this time, one of the accelerometers on the deck had become waterlogged and was giving a corrupted response, which accounts for the large number of outliers.

The outlier analysis performed above illustrates the fact that novelty detection is feasible without prior manipulation of data to remove the influence of environmental and operational conditions. In this case, the daily variations in frequency caused by environmental and operational conditions, illustrated in Figure 5.3(a), have been incorporated into the definition of the normal condition of the structural response. However, although this analysis is able to detect anomalies, as evidenced by the fact that a corrupted signal is detectable, the incorporation of the daily variations into the normal condition may render some potential performance anomalies undetectable; a performance anomaly or change in structural condition that produces a variation in frequency smaller than that of the daily variation caused by environmental/operational conditions will certainly not be detected using the above approach. If one is interested in such anomalies, steps must be taken to account for environmental and operationally induced variation before novelty detection is attempted.

One idea to overcome this problem, that is very relevant to the analysis carried out in the previous chapter, is the suggestion of using model prediction errors as an anomaly detector, where the model is trained on data from the structure in its normal condition [83]. As long as the model can predict the monitored variable(s) in question with a good degree of accuracy, any large increase in model prediction error can be taken to mean that the structure has deviated from its normal condition. If the simple models used in the previous chapter in an attempt to better understand the bridge's normal condition are capable of predicting the modal frequency change to a good and most importantly consistent degree, their prediction errors would be a good candidate for an indicator of structural condition that is not affected by environmental and operational conditions. Alternatively, if the simple models produced in the previous chapter are not consistent predictors, the different and more complex modelling techniques discussed earlier [55, 57] could be utilised in a similar way. For comparison with the simply constructed models of the previous chapter, **Gaussian process regression** will be used in this chapter to model the natural frequency change with respect to measured environmental and operational conditions. Specifically in the chapter both model types will be used to attempt to predict the change in the lowest natural frequency of the deck. The use of Gaussian processes for regression is a growing area of interest in many disciplines, some brief discussion of them will be included here, but for more details readers are referred to Appendix A and [84]. Most recently in SHM, Gaussian process regression has been used for prediction of crack growth in aluminium specimens [85].

The use of Gaussian processes (GPs) is a sophisticated nonparametric Bayesian approach to regression and classification problems. Gaussian process regression, unlike classical maximum likelihood approaches, considers all possible functions that fit to a training data set and provides a predictive distribution as opposed to a single crisp prediction for a given input. From this predictive distribution a mean prediction and confidence intervals on this prediction can be obtained. Nonparametric approaches for regression have the benefit that their complexity is not limited by a set functional form. An additional benefit of using GPs lies in their compactness, the computations necessary for GP regression are simplified by the fact that a distribution directly over candidate functions can be defined, rather than over the parameters of a predefined function (as would be necessary for a Bayesian neural network for example).

In this work, the interest is in how best to model features of interest, natural fre-

quencies in this case, with respect to the environmental and operational conditions that drive their variation in the normal condition of a structure. To use Gaussian process regression, a mean and covariance function ($m(\mathbf{x})$ and $k(\mathbf{x}_p, \mathbf{x}_q)$[1] respectively), along with any hyperparameters that govern these functions, must first be specified (here p and q refer to different samples). The specification of the mean and covariance functions defines all possible candidate functions to be considered in the inference procedure. Here, a zero mean function and a squared exponential covariance function are used. The squared exponential function has the form:

$$k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_y^2 \exp\left(-\frac{1}{2l^2}|\mathbf{x}_p - \mathbf{x}_q|^2\right) + \sigma_n^2 \delta_{pq}, \tag{5.2}$$

where σ_y^2, l and σ_n^2 are hyperparameters to be determined. σ_y^2 is the signal variance (limits the vertical scale of the process), l is the length scale of the process, which defines the smoothness (determines the length between inputs before function values can change significantly), and σ_n^2 is the variance from the noise on the measurements.

Once defined, the mean and covariance functions are then conditioned on some training data. Conditioning acts to effectively discount any functions described by the original mean and covariance functions that do not match the training data. From these conditioned functions a distribution over predictions can be directly obtained.

Given a set of training data, with inputs arranged in a design matrix X and target values $\mathbf{y}$, a set of testing data with inputs arranged in a design matrix X_* and unknown target values $\mathbf{y}_*$, conditioning gives the mean (m_*) and variance (k_*) for the prediction of $\mathbf{y}_*$ as:

$$m_* = K(X_*, X)K(X, X)^{-1}\mathbf{y} \tag{5.3}$$

$$k_* = K(X_*, X_*) - K(X_*, X)K(X, X)^{-1}K(X, X_*) \tag{5.4}$$

where $K(X_*, X)$ is the calculated covariance matrix of the design matrices X_* and

[1]Note that structural dynamics notation is not used here to describe Gaussian process theory, as it would prove too cumbersome, instead the notation adopted uses bold font for vectors and upper-case letters for matrices.

X. The hyperparameters are determined by minimising the negative log marginal likelihood (again see Appendix A or [84] for more details).

For prediction of the lowest natural frequency, measurements of the temperature, traffic loading, wind speed and direction and the root-mean-squared acceleration of the deck were used as inputs to the model, which were all the original parameters considered when constructing the less complex models of the last chapter.

For comparison of the response surface models and GP regression, a training set of 3000 data points (two months' worth) will be used, which specifically is data collected from the end of October to the end of December 2007. A testing set of 9568 data points will be used to gauge model performance, which is comprised of data collected over a period of two years. After training both model types to predict the change in the lowest natural frequency both models were found to be similarly successful; the training MSE for the response surface model was 18.49, the training MSE for the GP regression was 18.39. For the testing set, the MSE of the response surface model was 21.17, while the testing MSE of the GP was 21.23. An example of the prediction of both of these models is plotted in Figure 5.5. Both models are able to predict the change in the natural frequency very well, it is interesting, also, to note that the performance of these separate models is very similar and neither can be said to outperform the other, which leads to the question of which method should be applied in such situations. Where one wishes to better understand the interactions between input and target parameters, parametric response surface models are the sensible choice as they are interpretable as shown in the previous chapter. However, where the sole aim of such modelling is for accurate prediction, the use of GPs seems a more sensible choice as the effort required by the programmer is much less as no prior manipulation of the input parameters is necessary.

As both models can predict the frequency change to a good degree of accuracy, either of the model errors should be suitable candidates for a novelty indicator. To explore the use of model errors for novelty detection, each of the models' errors are plotted in Figure 5.6, with confidence limits added at plus and minus three standard deviations (which is the 99.7% confidence interval for a null result) of the errors from the training period. Note that only the confidence limits for the GP errors are visible in Figure 5.6, as these overlay the confidence limits of the response surface model. Apart from very few anomalies, the model errors stay within the confidence interval, which demonstrates that both models could be used as a potential indicator of structural condition.

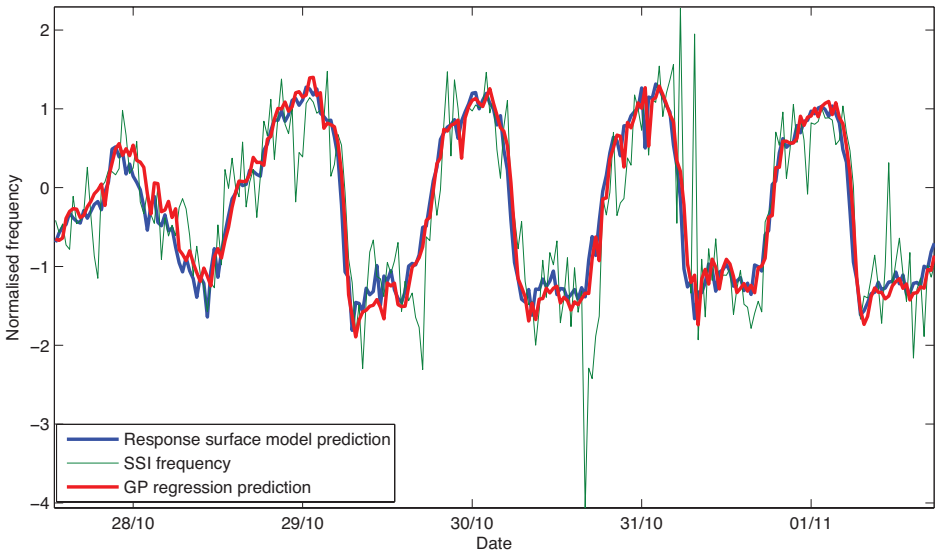


Figure 5.5: Comparison of response surface model and GP predictions for the first deck natural frequency

As always the problem lies in the fact that no data are available for validation of this statement, as the bridge has not undergone any documented structural changes. However, in this case, more data than those used for model training and validation are available from the period of time when there was a suspected sensor fault as mentioned previously. Extending the model predictions beyond the training and testing period, Figure 5.7 plots both models' prediction errors for period of time after the training and testing period. Studying Figure 5.7, the errors clearly depart significantly from the confidence interval during the time of the suspected sensor fault. Although this is somewhat a synthetic example, it does show that the model error plot is able to clearly detect (with a large number of outliers) a departure from the normal response condition.

It was previously suggested in this text that using the errors of predictive models would produce more sensitive anomaly detectors than an outlier analysis carried out on the raw measurements. At the beginning of this section, a multivariate outlier analysis including the lowest five natural frequencies of the deck was carried out. Here, to further investigate this argument, a univariate analysis is carried out to compare the number of outliers of each of the models' errors with the number of outliers of the 'raw' measured lowest natural frequency. Figure 5.8 shows the calculated discordancy of the response surface model errors, the GP model errors and the measured natural frequency for the 9568 point testing period described

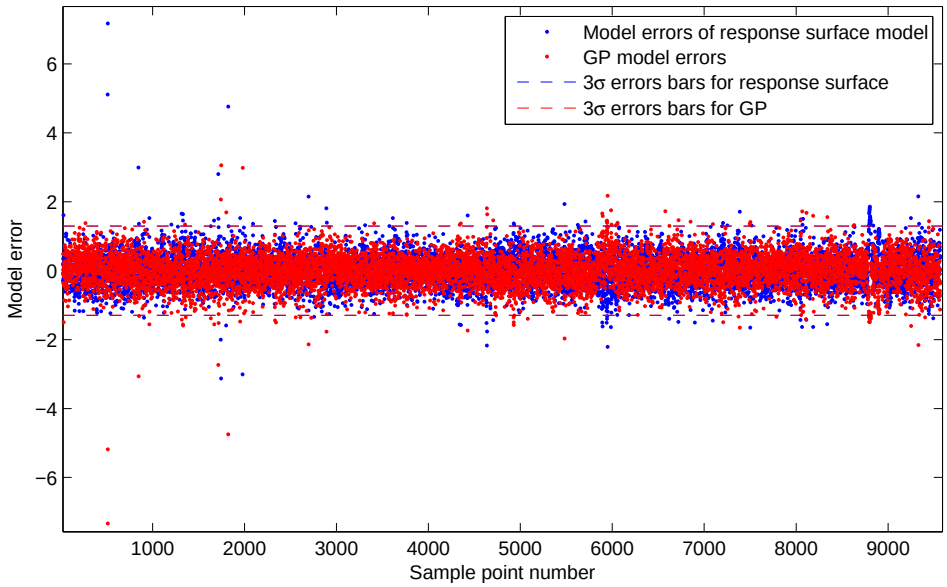


Figure 5.6: Control chart of model errors of response surface and GP models

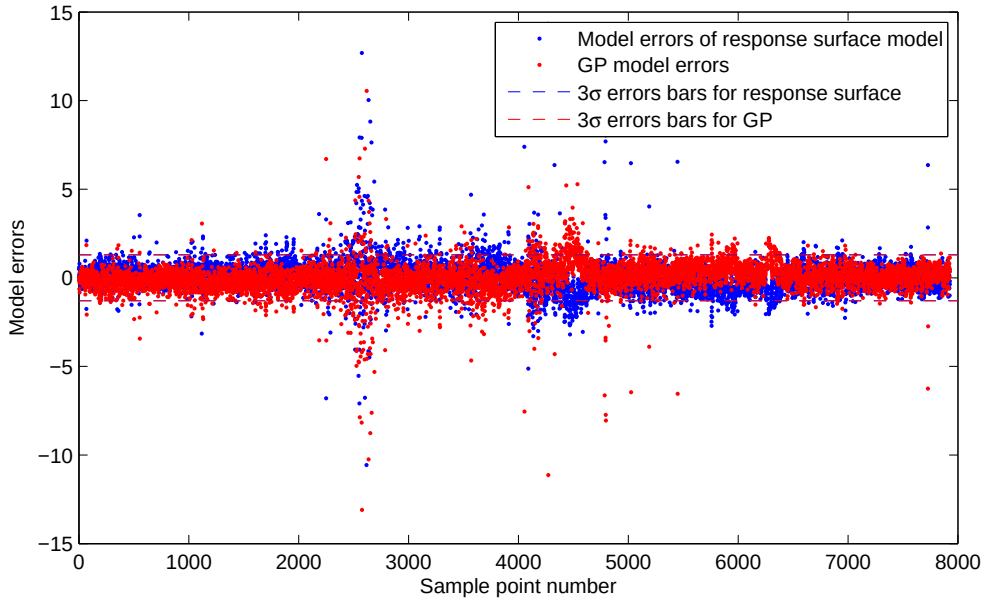


Figure 5.7: Control chart of model errors of response surface and GP models for period of time of suspected sensor fault

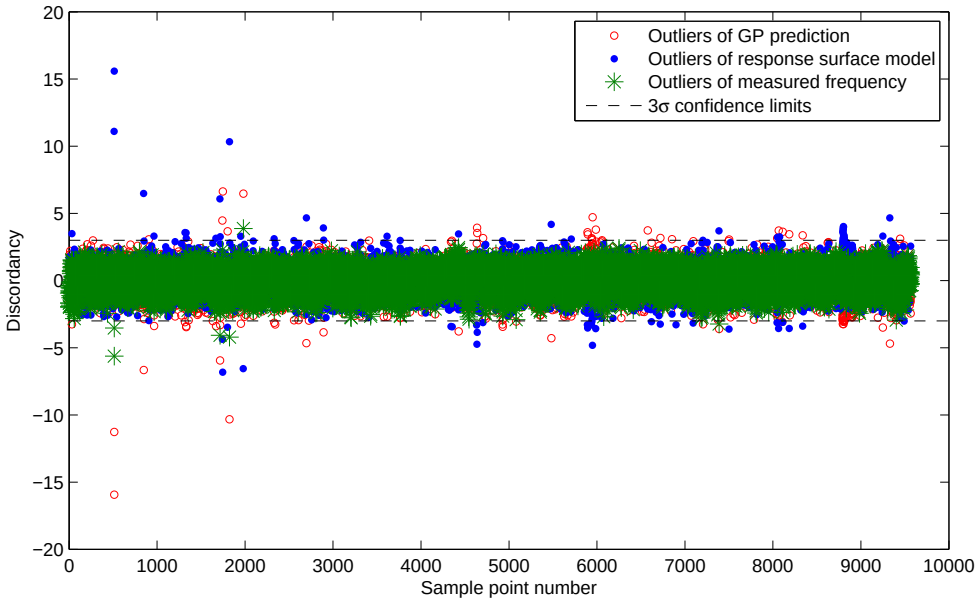
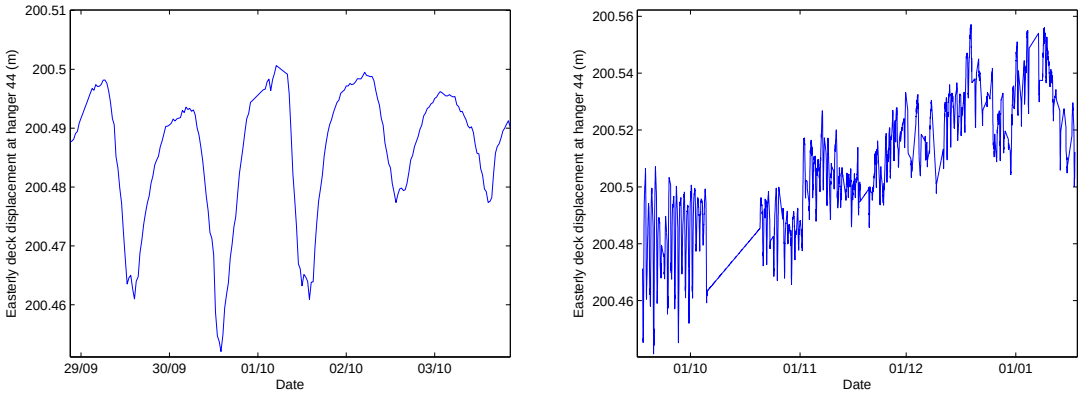


Figure 5.8: Univariate outlier analysis of measured natural frequency, GP model errors and response surface model errors

above. Studying Figure 5.8 one can see that when using model predictions, the number of outlying data points is much increased, demonstrating that, as previously suggested, the sensitivity to outliers is much increased when modelling is used to take into account operational and environmental variation.

One issue that must be discussed here is that the success of a novelty detector using model errors is reliant on an accurate predictive model. If a model is not accurate, the Gaussian error assumption that the novelty detection here is based on will likely not be valid. Although the use of Gaussian processes goes a long way to ensure that the best prediction feasible is obtained, an accurate prediction may not be possible if the latent variables driving the variation of the target are not well represented in the inputs to the model. It is the belief of the author, that although the response surface model and the GP provide a good prediction of frequency change, more accurate predictions could be obtained with a higher fidelity estimation of the traffic loading (for which only hourly counts of vehicle classes crossing the bridge are available). An improvement in model fidelity would render the novelty detection less susceptible to false-positive detections of 'anomalous' events.



(a) (b)

Figure 5.9: Easterly deck deflections at hanger 44: (a) over 5 days, (b) over four months

5.3 Detecting novelty with measured deck deflections

How novelty detection for SHM and performance monitoring could be applied to the measured deck deflections will now be explored. As previously described, the TPS data available for this study spans a four month period from September 2009 to January 2010. Figure 5.9, similarly to Figure 5.3, shows plots of the measured bridge deck deflections, in (a), the easterly (longitudinal) deflections of the deck at hanger 44 are plotted over a period of five days, in (b), the same over a four month period.

Studying Figure 5.9, it is obvious that the measured deck deflections have daily and seasonal trends, which reflects the correlation of the deck movement with temperature. The consequence of this seasonality is that the signals over the available time window are nonstationary and that novelty detection is not directly applicable to this data set, as it was with the natural frequency data. To demonstrate this, a multivariate outlier analysis is carried out on easterly displacement measurements at six locations (at hangers 44, 62, 80, 98, 112 and 123). A training period of 500 data points corresponding to the first half a month of data points of the data set shown in Figure 5.9(b) was used. The results are plotted in Figure 5.10 where the dashed line is the outlier (95% confidence) threshold. As one can see, despite including a large number of data points in the training set, the number of 'outliers'

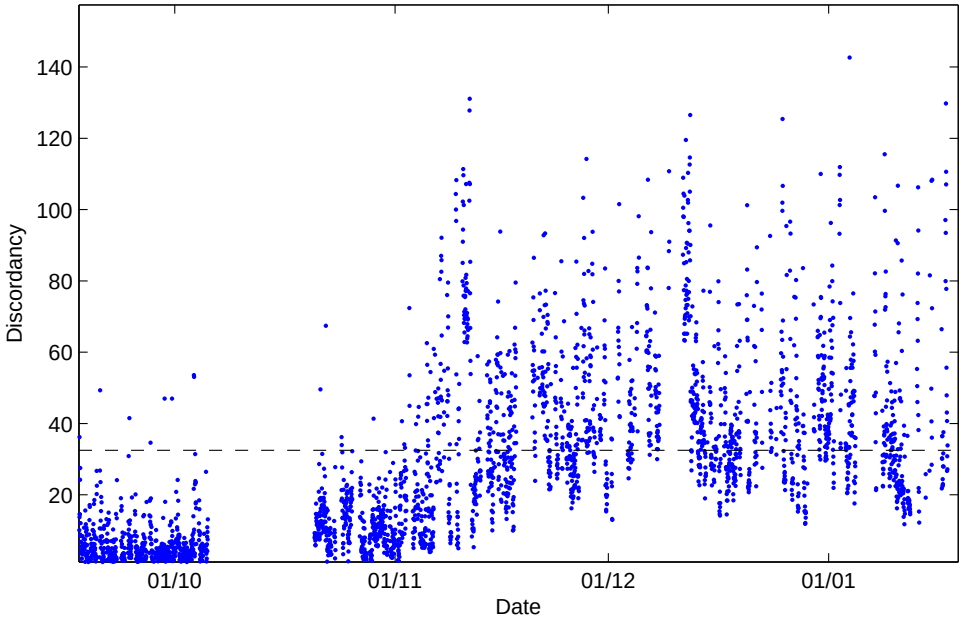


Figure 5.10: Multivariate outlier analysis of 6 measured easterly deck deflections

increases as the mean deck deflection also increases (caused by colder weather). In other words, deflection measurements from colder weather are wrongly classified as outliers.

For a successful novelty detection process without prior manipulation of the data to remove the environmental trends, most likely the training data would need to span a period of a year in order account for the seasonal trends. This would, however, significantly lower the sensitivity of the novelty detector, and may render it useless. Instead, additional steps are necessary to account for the environmentally induced trends before novelty detection is attempted.

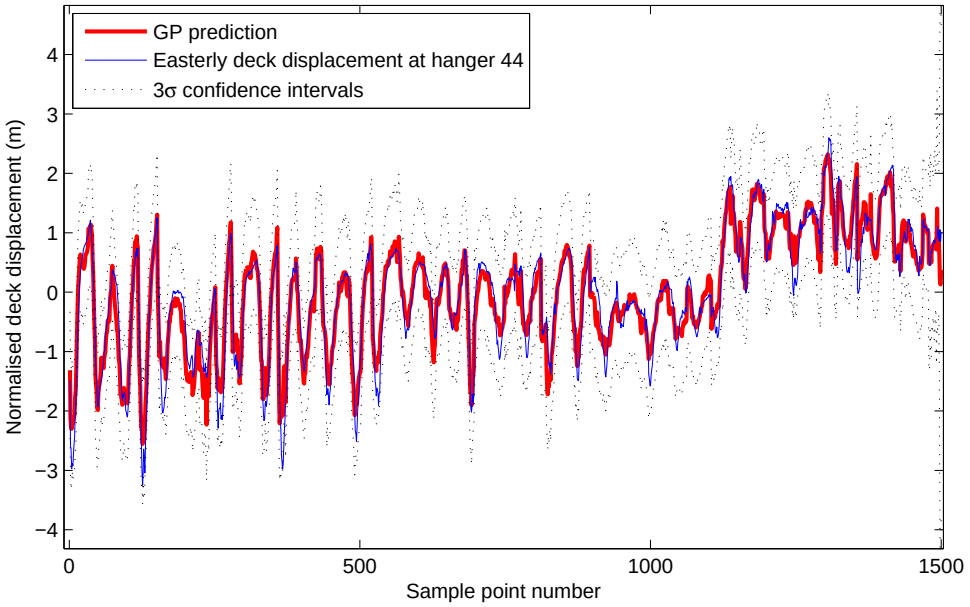
As with the natural frequency data discussed previously, a sensible choice for enhancement of novelty detection is to consider the use of data modelling. Gaussian process regression may once again be considered along with many other data modelling approaches (see for example [55, 57]). Unfortunately, the seasonality of the data introduces an additional complication to the modelling process, which is that training data may be needed from a period spanning a whole year for an effective predictive model. To demonstrate this, Gaussian process regression is used to attempt to model the easterly deck deflection measurement at hanger 44 plotted in Figure 5.9, using the available deflection data from the four month period. Before

implementing the model, the data gap visible in Figure 5.9 was first removed. Once again a zero mean function and a squared exponential covariance function were used. The training data set was chosen as every fifth point of the first half of the data set (1500 points) after the data gap had been removed, and inputs to the model were temperature, traffic loading, wind speed and RMS of deck acceleration. The GP prediction is shown in Figure 5.11, where the top plot (a) shows the GP predictions from the training period, and the lower plot (b) shows the GP predictions for the remainder of the data which corresponds to the last two months worth of data in Figure 5.9(b).

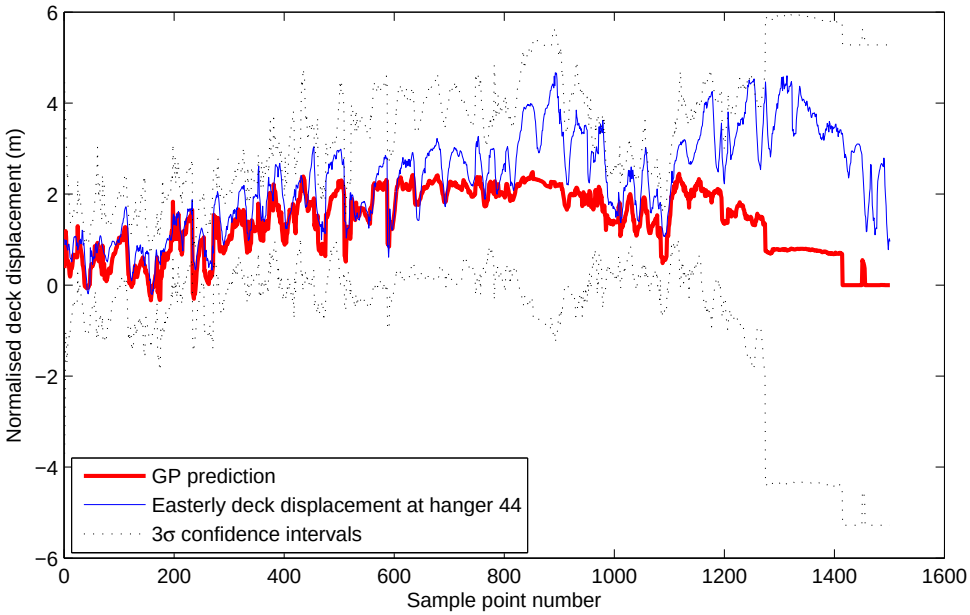
Studying Figure 5.11, one can see that the majority of the GP prediction is accurate. However, as the mean displacement starts to increase in the testing set, the predictions become less accurate, and towards the end, the confidence intervals increase as new environmental conditions not in the training set are encountered. The MSE for the training and testing period together is 39.97. The GP prediction carried out here would undoubtedly improve with an extended training period, which is currently not available for this study.

To investigate how the GP prediction may aid novelty detection here, a univariate outlier analysis of the model errors was carried out, with results shown in Figure 5.12 (training was on the first 1500 points, which corresponds to the GP training period). Another possibility for a novelty indicator, available when using Bayesian techniques, could be to use the confidence intervals that are available when applying GP regression as a statistical process control chart [86]. The intervals plotted in Figure 5.11 are plus and minus three standard deviations of the predictive distribution. Presuming an accurate model, if a measurement is outside these confidence intervals it may be counted as an outlier. This approach has the further advantage that if environmental/operational conditions occur that are very different from those in the training set, the confidence intervals will increase (as one can see in Figure 5.11(b)), meaning that a performance anomaly will not be falsely detected if new environmental/operational conditions occur. For comparison with the usual outlier analysis (Figure 5.12) the model errors are also plotted with the GP prediction confidence intervals as a control chart in Figure 5.13.

In Figure 5.12, as the model prediction ability decreases with colder temperatures a large number of erroneous outliers occur. A broader training data set is required to overcome this problem. In Figure 5.13 one can see that the GP confidence intervals expand when the prediction encounters a new area of the input space (in this



(a)



(b)

Figure 5.11: GP predictions of easterly deck deflections: (a) training period, (b) testing period

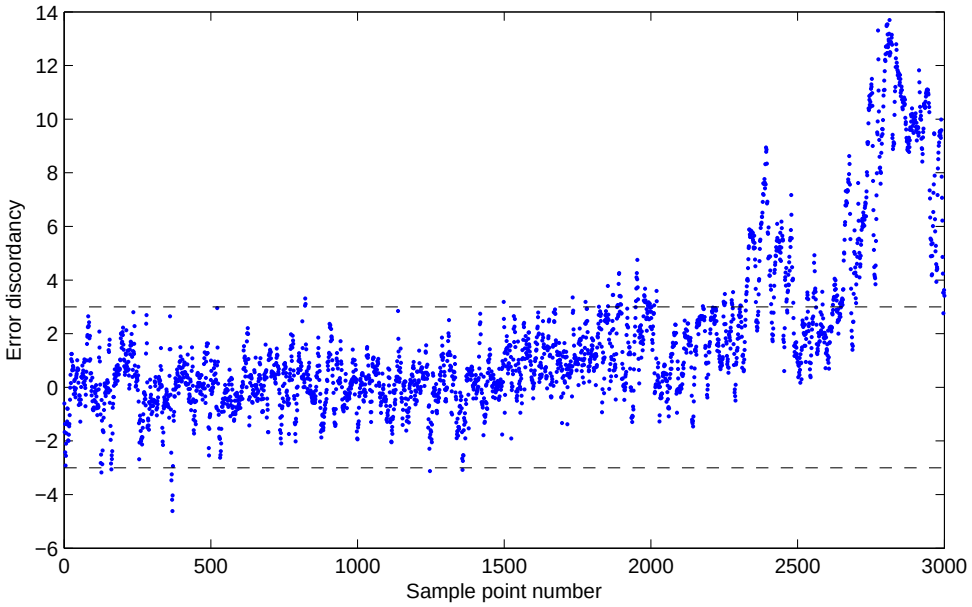


Figure 5.12: Univariate outlier analysis of GP prediction for easterly deck displacements at hanger 44

case lower temperatures than were present in the training period), which has the consequence that anomalies are not falsely identified as readily. When conservatism on identification of potential performance anomalies is necessary, this idea of using confidence intervals may prove to be very useful.

5.4 Conclusions

Novelty detection in the face of environmental and operational variations has been discussed in the context of data gathered from the Tamar Bridge monitoring campaign. Natural frequency and deck deflection measurements have been used to demonstrate two different approaches to novelty detection for features that experience variability caused by environmental and operational conditions. The first approach is to incorporate responses caused by such variable conditions into the definition of the normal condition. This is most feasible for features that do not exhibit nonstationary behaviour over long time periods, such as the natural frequencies of the bridge deck. The advantage to this approach lies in the ease of its application, a potential disadvantage is that sensitivity to performance anomalies may be lower. An alternative approach is to account for environmental/operational variation be-

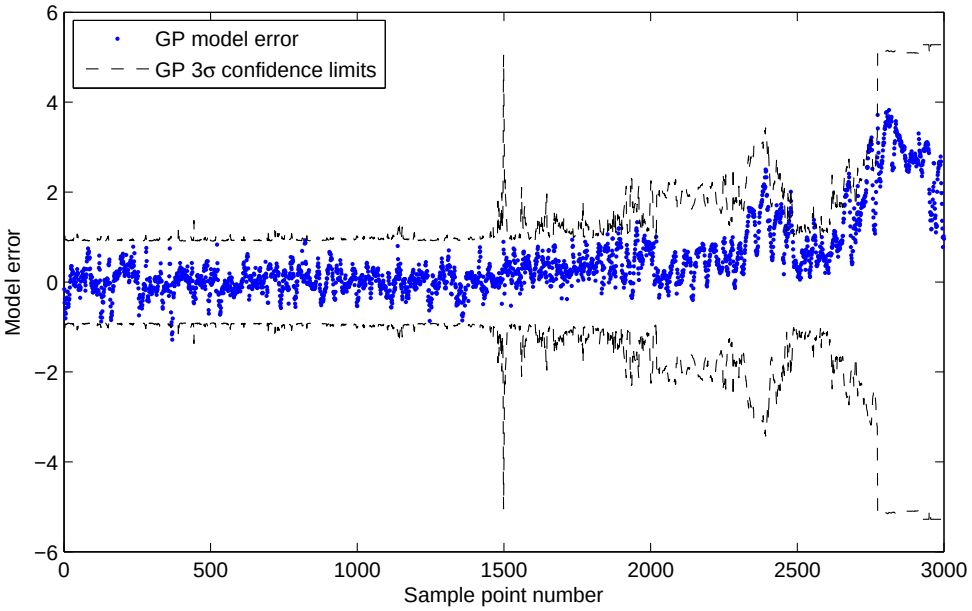


Figure 5.13: Outlier analysis of GP prediction for easterly deck displacements at hanger 44 using GP confidence intervals

fore novelty detection is attempted. In this chapter, Gaussian process regression and response surface modelling were used to model natural frequency and deck deflection with respect to the external conditions. If an accurate predictive model is available, model errors may be used as a good candidate for novelty indicators. Both parametric response surface models and Gaussian process regression were found to be a good candidates for generating accurate predictive models.

A further advantage of Gaussian processes for enhancement of novelty detection has also been highlighted, which is that the prediction confidence intervals available when using GP regression can be used to directly detect novelty in a conservative manner. The use of such confidence intervals would reduce the risk of a false-positive indication of novelty, as confidence decreases if new environmental/operational conditions are encountered.

As the chapter title suggests, this is a tentative step towards the development of a system capable of detecting changes in structural condition. The models for the first modal frequency are a success in that they can detect a departure from the normal condition. However, there is still a long way to go as far as practical SHM is concerned. One condition indicator based on one modal frequency, for example, would evidently not be sufficient to reliably assess the state of such a structure, especially

as modal frequencies can be insensitive to localised damage scenarios. Any credible system put in place for real anomaly detection would need a number of such predictive models taking into account different response measurements, not just global modal parameters. Furthermore, although the detection of a possible sensor fault with the frequency prediction model was a useful exercise to show how a departure from the normal condition could be detected, it illustrates perfectly another challenge that must be met with before any SHM system can be relied upon, which is that it must be able, not only to distinguish between response fluctuations caused by environmental and operational conditions but to distinguish between sensor faults and real structural degradation. If model prediction errors are to be used as an indicator of structural condition, sensor faults must be detected before any data is fed into the model. For further reading on detecting sensor faults see, for example [87].

Chapter 6

Cointegration for the data normalisation problem

This chapter introduces the concept of *cointegration*, a tool for the analysis of non-stationary time series, as a promising new approach for dealing with the problem of environmental variation in monitored features. If two or more monitored variables from an SHM system are cointegrated, then some linear combination of them will be a stationary residual purged of the common trends in the original data set. The stationary residual created from the cointegration procedure can be used as a damage sensitive feature that is independent of the normal environmental and operational conditions.

6.1 Introduction to cointegration

As discussed in the previous chapters, before SHM technologies can be reliably implemented on structures outside laboratory conditions, the problem of environmental variability in monitored features must first be addressed. In the previous two chapters data modelling and novelty detection techniques have been applied in an attempt to address this issue, in this chapter a new approach for data normalisation is discussed. As summarised in Chapter 2 of this thesis, several different ways to deal with the problem of environmental and operational variations in features have already been suggested, perhaps the most promising of which involve the projection

of the feature data onto new axes (using PCA [63] or factor analysis [65]). It transpires, however, that the projection approach had actually been anticipated, but in the literature of econometrics; as described by [88, 89]. In fact, the PCA approach in econometrics actually belongs to a larger class of algorithms being developed in the literature; these algorithms are associated with the concept of *cointegration.*

In the current work, cointegration is suggested as a suitable methodology for removing environmental trends from SHM data. Recently the idea has been applied to statistical process control in [90].

Cointegration is a property of some nonstationary time series; briefly, two or more nonstationary variables are *cointegrated* if some linear combination of them is stationary. Econometricians traditionally test for cointegration between two or more economic variables as a means of establishing whether there is a statistically significant relation between them. Although engineers may well be interested in problems of a similar nature, they may find the stationary linear combination created during the cointegration process of more practical interest. If a number of variables from some process under investigation are cointegrated, the stationary linear combination of them found during the cointegration process will be purged of all common trends in the original data sets, leaving a residual equivalent to the long run dynamic equilibrium of the process [90]. In terms of SHM data, the common trends removed by the cointegration process will be those caused by the latent variables driving the response of the structure, i.e. the environmental and operational conditions.

In theory then, the cointegration process is ideally suited to remove environmental and operational trends from SHM data. This idea is explored using a simulated scenario in the next section. Subsequently the, sometimes complex, mathematics behind the cointegration process will be introduced.

6.2 Illustrative Example

The current section is intended as an illustrative and intuitive example of the approach taken towards removing environmental trends from damage sensitive monitored variables in this work. In essence, the following section describes how cointegration can be utilised for SHM; any formal reference to the cointegration procedure is, however, left to later sections. To begin, the dynamic response of a ten degree of

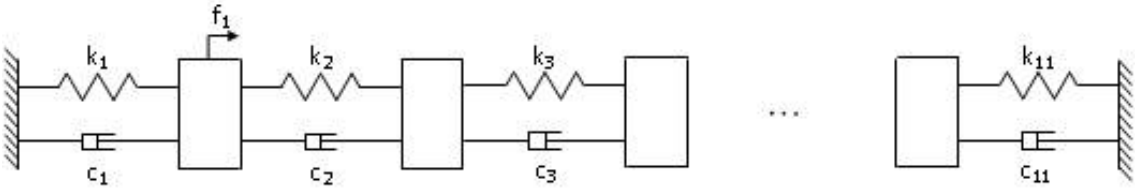


Figure 6.1: Simulated 10 degree of freedom lumped mass model

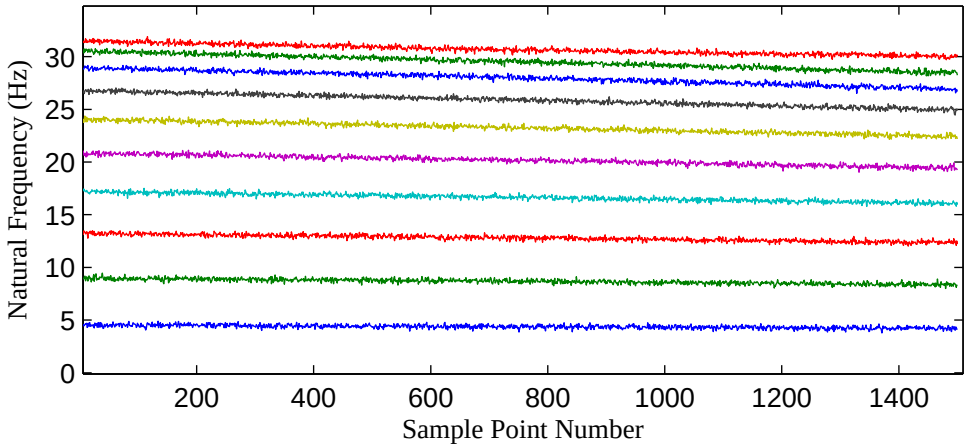


Figure 6.2: Extracted natural frequencies of the simulated system

freedom system, shown in Figure 6.1, to a random excitation is simulated.

To imitate a dependency on environmental conditions and introduce nonstationarity to the dynamic response of the system, the spring stiffnesses were allowed to vary with time. A simulated temperature field was then added with the highest temperature at the first mass and affecting the subsequent masses successively less. Each stiffness was set as a linearly decreasing function of temperature, and over the duration of the simulation this temperature was steadily decreased. To serve as damage sensitive features, the natural frequencies of the system were extracted by solving the eigenvalue problem at each time instant in the simulation as shown in Figure 6.2, a small amount of Gaussian noise has also been added to simulate instrument noise.

In their current state, each natural frequency is dependent on temperature and therefore less than ideal as a choice for a damage sensitive feature to monitor. However, as each frequency here is driven by the same temperature field, the correct parameter choice for a simple linear combination of a number of the frequencies would result in a stationary residual, purged of any dependence on that temperature field.

In this example, the first two natural frequencies are combined as follows;

$$\alpha\omega_1 + \beta\omega_2 = \varepsilon, \tag{6.1}$$

where ω_i denotes the ith natural frequency, ε the residual sequence and α, β are constants to be determined. On the correct choice of the parameters α, β the residual sequence will become stationary, as shown in Figure 6.3 for this example. Now, regardless of the temperature, this residual will continue to be stationary all the time the system is operating in its normal condition. Upon the occurrence of an event that changes the relationship between the variables included in the combination the residual will become nonstationary. In other words, the residual sequence will become nonstationary if the system begins to operate outside of its normal condition. For structural health monitoring, then, the residual sequence seems a sensible choice for a damage sensitive feature.

For the current example, it remains to test the created stationary residual's sensitivity to damage. In order to do this, the system was re-simulated, again with temperature variation but with a different excitation sample; in this case, the second spring stiffness was abruptly reduced to 50% of the healthy value mid-way through the simulation. Figure 6.4 illustrates the projection of the newly simulated features onto the established combination (6.1). Clearly the residual becomes nonstationary at the midway point, corresponding to the introduction of damage to the system. The upper and lower limits are computed from the undamaged residual and represent the mean plus or minus three standard deviations. This plot is essentially a Statistical Process Control (SPC) 'X-chart' [86], which clearly indicates that damage has occurred shortly after the mid-point of the time record.

As mentioned previously, the point of this example has been to illustrate the suggested approach in this work to removing unwanted environmental trends from damage sensitive features. The problem has been reduced to finding the correct parameters with which to combine the nonstationary variables into one stationary residual sequence. This approach, however, is evidently only valid for variables that share common trends that can be removed by a linear combination of the original variables. From econometrics, this is in fact the definition of cointegrated variables; one can therefore draw on the considerable and sophisticated research carried out in that field to discover the best way to approach the problem of finding the parameters that will create the most stationary residual sequence.

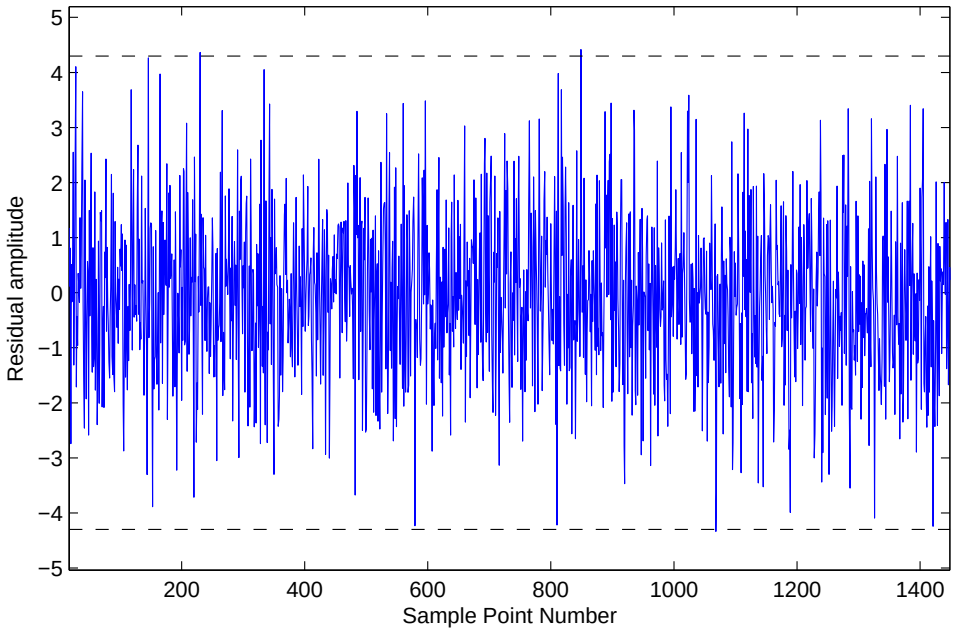


Figure 6.3: Stationary residual created from equation (6.1)

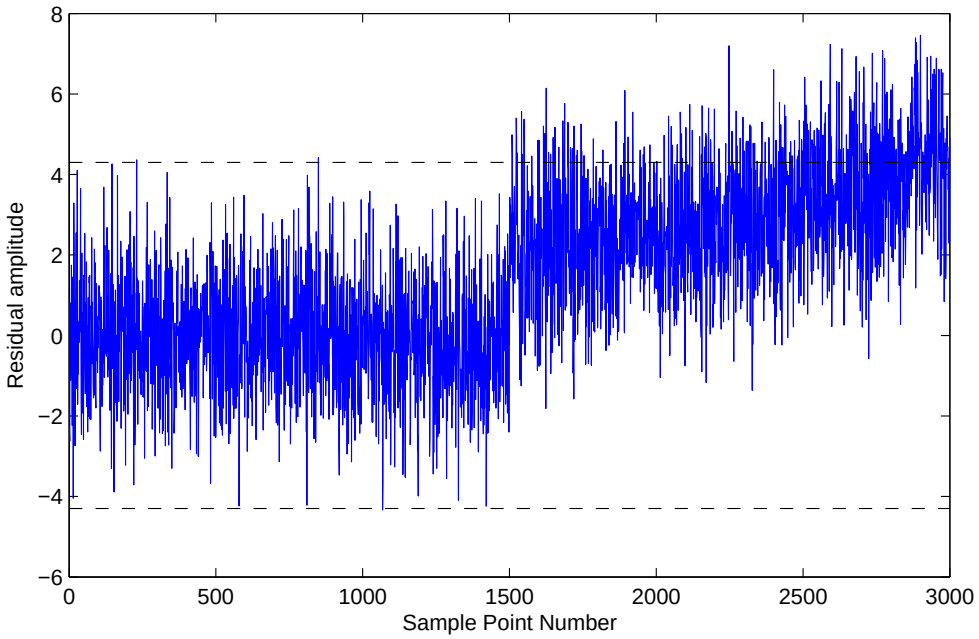


Figure 6.4: Residual sequence of the natural frequencies in the 'damaged' case

6.3 The theory of cointegration

6.3.1 Overview

The current section aims to introduce the concept of cointegration in a more rigorous manner than attempted in the previous two sections of this chapter. All of the following theory is known from the econometrics literature; however, it is considered useful here to present it in a form tailored to structural dynamicists and in appropriate notation. In the following some familiarity with auto-regressive and vector auto-regressive models is assumed, however, Appendix B provides a short introduction to the points considered important. As noted above, this section draws from a number of key texts from the econometrics literature [91–94], which may be referred to for a more mathematically rigorous treatment of the material. Here, to begin, a simple definition of cointegration is introduced.

Definition 1. Two or more nonstationary time series are *cointegrated* if a linear combination of them is stationary.

In the following equation, where the nonstationary time series are modelled as a vector-autoregressive process (VAR) $\{y_i\}$, the series are cointegrated if a vector $\{\beta\}$ exists such that z_i is stationary, where

$$z_i = \{\beta\}^T \{y_i\}. \tag{6.2}$$

If this is the case, $\{\beta\}^T$ is called a *cointegrating vector*. If $\{y_i\}$ includes a total of n variables, there may be as many as $n-1$ linearly independent cointegrating vectors. Clearly for the time series to be cointegrated they must have shared/common trends to begin with. There is one further restriction, which is that all times series must be *integrated* of the same order.

Definition 2. If a nonstationary process variable y becomes stationary after differencing d times, it is said to be *integrated of order* d, which is denoted $y \sim I(d)$.

In other words, the time series must have the same 'degree of nonstationarity' if they are cointegrated.

For the purposes of structural health monitoring the intent would be to use mon-

itored variables that are cointegrated and find the cointegrating vector to create a stationary residual sequence for damage detection. From an engineering point of view, monitored variables from the same process or system are more than likely to share common trends on account that each process variable will be driven by the same latent influences. This cannot be said, however, of the order of integration of each monitored variable, this must be ascertained before any attempt is made to find the cointegrating vector.

The order of integration of a time series is ascertained in econometrics by employing a stationarity test, which is often analogous to testing for a unit root in a time series model. The stationarity test employed here is called the Augmented Dickey-Fuller test and will be described later in this section.

Once it has been ascertained to what order all process variables of interest are integrated to, it remains to find the cointegrating vector that will result in the most stationary combination of the variables. There are two common approaches to this problem in econometrics; the first is the Engle-Granger two step estimation procedure [95] often employed when there are only two process variables included in the analysis, the second is the Johansen procedure [96], a more complex maximum likelihood multivariate estimation procedure. Due to its increased sophistication the Johansen procedure will be employed here, its sometimes complex mathematics will be described, but a summary of the process will also be provided for quick reference.

6.3.2 The Augmented Dickey Fuller Test

The first step is to test that all variables under consideration are integrated of the same order, this is achieved by the Augmented Dickey Fuller test [97, 98]. Like many econometric stationarity tests, the Augmented Dickey Fuller (ADF) test is based on a unit root test for a time series model. If a time series model has a characteristic root on the unit circle it will be inherently nonstationary. In this work only real valued roots will be considered (see Figure 6.5). The idea is perhaps best illustrated by looking at a first-order auto-regressive model $AR(1)$, which takes the form

$$y_i = a_1 y_{i-1} + \varepsilon_i \tag{6.3}$$

where ε_i can be considered to be a Gaussian white noise process. In this case, the

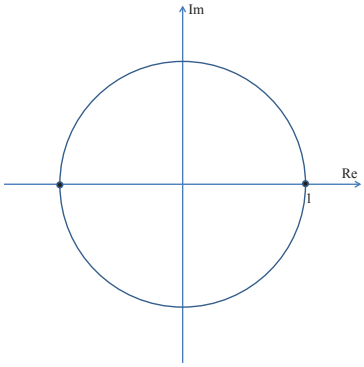


Figure 6.5: The unit circle, with real unit roots highlighted

value of a_1 defines the root of the characteristic equation of the process (see Appendix B for more details). The roots of the characteristic equation of any process determine its stability and therefore its stationarity. In this example, the process y_i will be stationary if a_1 is less than one in magnitude and nonstationary if it is larger or equal to one in magnitude. In the case that a_1 is equal to one, the process will have a *unit-root*, and equation (6.3) becomes

$$y_i = y_{i-1} + \varepsilon_i \implies \triangle y_i = \varepsilon_i \tag{6.4}$$

The process will be nonstationary but its first difference will be stationary, in econometrics terminology it will be integrated order one, denoted $y_i \sim I(1)$.

When fitting a process to an AR(1) model then, information on the stationarity of the process is obtained from the parameters defining the characteristic root. This is normally achieved by testing a null hypothesis of $a_1 = 1$. The most obvious way of going about this would be to carry out a t-test on the parameter a_1, however, under the assumption of nonstationarity, the least-squares estimate of the parameter will not be distributed around unity. Rather than carrying out a traditional t-test, the t-test statistic will normally be compared with critical values constructed by Dickey and Fuller, found in [97].

The Augmented Dickey Fuller test follows the same premise as described above but involves fitting the data to a more complex time series model as described by the following equation:

$$\triangle y_i = \rho y_{i-1} + \sum_{j=1}^{p-1} b_j \triangle y_{i-j} + \varepsilon_i \tag{6.5}$$

Here the difference operator $\triangle$ is defined as $\triangle y_{i-j} = y_{i-j} - y_{i-j-1}$. A suitable number of lags p should be included to insure that ε_i becomes a white noise process [99]. To convert from the more traditional *AR(p)* model to the model form (6.5), the following substitutions should be made; let $a_1 = 1 + \rho + b_1$, $a_n = -b_{n-1} + b_n$, for $n = 2 \ldots p-1$, and $a_p = -b_{p-1}$, where a_j are the AR model coefficients.

Using these substitutions the characteristic equation of (6.5) can readily be obtained from the characteristic equation of an *AR(p)* process as

$$1 - \lambda^{-1} - \rho\lambda^{-1} - \sum_{j=1}^{p-1} b_j(1 - \lambda^{-1})\lambda^{-j} = 0, \tag{6.6}$$

where the λ are the roots of the characteristic equation. With this more complex form of time series model there could be as many as p independent roots. As an explosive process would be obvious to the analyst from the outset, the scenarios that remain of interest here are those where all roots are smaller than or equal to unity. If at least one root of the characteristic equation is unity, it follows from (6.6) that ρ must equal zero. Assume for the moment that there is a single unit root and consider the remaining $(p-1)$ roots of the characteristic equation. With $\rho = 0$, equation (6.6) becomes

$$\begin{aligned} &\left(1 - \lambda^{-1}\right)\left(1 - \sum_{j=1}^{p-1} b_j \lambda^{-j}\right) = 0 \\ &\Rightarrow 1 - \sum_{j=1}^{p-1} b_j \lambda^{-j} = 0 \end{aligned} \tag{6.7}$$

It is clear that if y_i has a one unit root, all remaining roots are smaller in magnitude than one. Furthermore, on closer inspection, (6.7) is the characteristic equation of the AR process of the differenced time series:

$$\triangle y_i = \sum_{j=1}^{p-1} b_j \triangle y_{i-j} + \varepsilon_i \tag{6.8}$$

As equation (6.7) must have all roots smaller than one in magnitude, the first difference Δy_i must be stationary. If y_i is nonstationary, but its first difference is stationary, as in the case above, the process is integrated order one.

To summarise, for y_i to be integrated of order one it is necessary that $\rho = 0$ in equation (6.5). The ADF test procedure is therefore to estimate the parameters in (6.5) by least-squares methods and then test the null hypothesis $\rho = 0$. The test statistic

$$t_\rho = \frac{\widehat{\rho}}{\sigma_\rho} \tag{6.9}$$

where $\widehat{\rho}$ is the least squares estimate of ρ, and σ_ρ the variance of the parameter should be compared with the critical values from the Dickey-Fuller (DF) tables (which are available in [92]). The hypothesis is rejected at level α if $t_\rho < t_\alpha$. If the hypothesis is accepted, the time series has a unit root and is *I(1)*. If the hypothesis is rejected, the test is repeated for $\triangle y_i$, if the hypothesis is then accepted y_i is an *I(2)* nonstationary sequence. This can be continued until the integrated order of the time series is ascertained.

Additional hypotheses and test statistics are needed if the model form used is extended to include shifts or deterministic trends (or both). For the extended time series model form

$$\triangle y_i = \rho y_{i-1} + \sum_{j=1}^{p-1} b_j \triangle y_{i-j} + \mu + \nu t + \varepsilon_i, \tag{6.10}$$

the null hypothesis for the time series to be integrated order one should be extended to include $\mu, \nu = 0$. More details for these specific cases can be found in [92, 98].

Having ascertained the degree of nonstationarity of each process variable of interest, an attempt to create a stationary residual through combination of those variables integrated to the same order can be made. The Johansen procedure is outlined below for this purpose.

6.3.3 The Johansen Procedure

The Johansen procedure is traditionally used to test if a number of *I(1)* economic variables are cointegrated, and if they are, to establish the number of independent cointegrating vectors and also determine which of the cointegrating vectors will create the most stationary linear combination of the variables in question. For completeness, the majority of the theory behind the procedure will be laid out here, although the real interest to SHM practitioners will most likely be in how to find the best cointegrating vector for a given set of monitored variables.

The premise of the Johansen procedure is to use a maximum likelihood approach to estimate the parameters of a Vector Error-Correction Model (VECM) of the variables under consideration. A VECM takes the form

$$\{\triangle y_i\} = [\Pi]\{y_{i-1}\} + \sum_{j=1}^{p-1} [B_j]\{\triangle y_{i-j}\} + [\phi]\{D(t)\} + \{\varepsilon_i\}, \tag{6.11}$$

where $\{y_i\}$ denotes an n-vector including all n variables to be analysed, with the subscript i relating to time, $i = 1, \ldots N$, p represents the model order, or the number of lags to be included in the model, and $\{\varepsilon_i\}$ is a normally distributed noise process; $\{\varepsilon_i\} \sim N(0, [\Sigma])$. A term to describe a deterministic trend $\{D(t)\}$ has also been included. Equation (6.11) is the multivariate analogue of equation (6.5).

Error-correction models are common in econometrics and are closely linked with the idea of cointegration. In fact, the existence of an error-correction model implies that the included variables are cointegrated and vice-versa, this is called the *Granger Representation Theorem* [95]. If a true error correction model exists (i.e. where $\{\varepsilon_i\} \sim N(0, [\Sigma])$), the parameters in $[\Pi]$ would describe the long-run equilibrium between variables, and the parameters $[B_j]$ would account for short run adjustments needed to return the process to equilibrium after any drifts.

The Johansen procedure uses the maximum likelihood of observing the correct $\{\varepsilon_i\}$ to estimate the parameters $[\Pi], [B_j], [\phi]$. A summary of the necessary points for calculation of the best cointegrating vector will be provided at the end of this section. The derivations of these points are provided below.

Under the assumption that (6.11) is a true error correction model and that the

variables under consideration are $I(1)$ (which implies that $\{\triangle y_i\}$ and $\{\triangle y_{i-j}\}$ are stationary) the parameter matrix $[\Pi]$ must be rank-deficient, say of rank r ($r < n$), and can therefore be decomposed into two matrices;

$$[\Pi] = [\alpha]\,[\beta]^T \tag{6.12}$$

where $[\alpha]$ and $[\beta]$ are both $n \times r$ matrices. From basic linear algebra theory, the r rows of $[\beta]^T$ will span the row space of $[\Pi]$. Now as the original matrix $[\Pi]$ described the long-run equilibrium relations between the variables, $[\beta]$ can be taken as the desired cointegrating vector to be found.

As previously indicated, parameter estimation is achieved by maximising the likelihood of observing the correct $\{\varepsilon_i\}$. If $\{\varepsilon_i\} \sim N(0, [\Sigma])$, its probability density function will be

$$p(\{\varepsilon_i\}) = \frac{1}{\sqrt{(2\pi^n)\,|\Sigma|}} \exp\left(-\frac{1}{2}\{\varepsilon_i\}^T[\Sigma]^{-1}\{\varepsilon_i\}\right) \tag{6.13}$$

where $|\Sigma|$ is the determinant of the estimated covariance of $\{\varepsilon_i\}$. It follows that the likelihood of observing the entire correct sequence of $\{\varepsilon_i\}$ will equal $\prod_{i=1}^{N} p(\{\varepsilon_i\})$.

On closer inspection of (6.13), each individual term is bounded above by the fractional term preceding the exponent, therefore the likelihood function is bounded above by $((2\pi)^n\,|\Sigma|)^{-\frac{N}{2}}$, and so

$$\mathcal{L}_{MAX} = ((2\pi)^n\,|\Sigma|)^{-\frac{N}{2}} \tag{6.14}$$

In other words, the maximum likelihood parameter estimates will correspond to the parameters that maximise $|\Sigma|$. This point will be returned to later, however, for now, the effort will be focused on manipulating the maximised likelihood function in order to express all parameter estimates in terms of $[\beta]$. Before going any further however, following [91], some new notation will be introduced to simplify the VECM expression (6.11).

Let

- $\{z_{0i}\} = \{\Delta y_i\}$,

- $\{z_{1i}\} = \{y_{i-1}\}$
- $\{z_{2i}\} = \{\{\Delta y_{i-1}\}^T, \{\Delta y_{i-2}\}^T, \ \ldots, \{\Delta y_{i-p}\}^T, \{D\}^T\},^T$ and
- $[\Psi] = [[B_1], \ [B_2], \ \ldots, \ [B_{p-1}], [\phi]]$,

then (6.11) will take on the simplified form:

$$\{z_{0i}\} = [\Pi]\{z_{1i}\} + [\Psi]\{z_{2i}\} + \{\varepsilon_i\} \tag{6.15}$$

Referring to this simplified form the log likelihood function L, where $L(-) = \ln \mathcal{L}(-)$, is first used to estimate $[\Psi]$ by calculating

$$\frac{\partial L}{\partial [\Psi]} = 0 \tag{6.16}$$

After the necessary matrix calculus and some careful rearrangement, the estimate for $[\Psi]$ can expressed as

$$\left[\widehat{\Psi}\right] = [M_{02}][M_{22}]^{-1} - [\alpha][\beta]^T[M_{12}][M_{22}]^{-1} \tag{6.17}$$

where $[M_{nm}]$ are product moment matrices defined by

$$[M_{mn}] = \frac{1}{N}\sum_{i=1}^{N}\{z_{mi}\}\{z_{ni}\}^T \qquad m, n = 0, 1, 2. \tag{6.18}$$

Substituting (6.17) back into equation (6.15), $\{\varepsilon_i\}$ may now be expressed as

$$\{\varepsilon_i\} = \{z_{0i}\} - [\alpha][\beta]^T\{z_{1i}\} - [M_{02}][M_{22}]^{-1}\{z_{2i}\} + [\alpha][\beta]^T[M_{12}][M_{22}]^{-1}\{z_{2i}\}. \tag{6.19}$$

This expression can be further simplified by defining the residuals $\{R_{0i}\}$ and $\{R_{1i}\}$ from the following regressions;

$$\begin{aligned}\{z_{0i}\} &= [C_1]\{z_{2i}\} + \{R_{0i}\} \\ \{z_{1i}\} &= [C_2]\{z_{2i}\} + \{R_{1i}\}\end{aligned} \tag{6.20}$$

where the coefficient matrices are found by ordinary least-squares; $[C_1] = [M_{02}][M_{22}]^{-1}$, $[C_2] = [M_{12}][M_{22}]^{-1}$. Finally, equation 6.19 becomes

$$\{\varepsilon_i\} = \{R_{0i}\} - [\alpha][\beta]^T\{R_{1i}\} \tag{6.21}$$

$\{\varepsilon_i\}$ has now been expressed in terms of the residuals of regressions of $\{z_{0i}\}$ and $\{z_{1i}\}$ on $\{z_{2i}\}$, and $[\alpha], [\beta]$, which are still to be found. To use econometrics terminology, the term $\left[\widehat{\Psi}\right]\{z_{2i}\}$ has been 'concentrated out'.

It now remains to find the maximum likelihood estimates of $[\alpha]$ and $[\Sigma]$ in terms of $[\beta]$. Assuming a fixed $[\beta]$, these are found to be

$$[\widehat{\alpha}] = [S_{01}][\beta]\left([\beta]^T[S_{11}][\beta]\right)^{-1} \tag{6.22}$$

$$\left[\widehat{\Sigma}\right] = [S_{00}] - [S_{01}][\beta]\left([\beta]^T[S_{11}][\beta]\right)^{-1}[\beta]^T[S_{10}] \tag{6.23}$$

where similarly to equation (6.18), $[S_{nm}]$ are product moment matrices defined by

$$[S_{mn}] = \frac{1}{N}\sum_{i=1}^{N}\{R_{mi}\}\{R_{ni}\}^T \qquad m,n = 0,1. \tag{6.24}$$

All free parameters of $\{\varepsilon_i\}$ have now been expressed in terms of $[\beta]$, which is still to be estimated. The estimation of $[\beta]$ is achieved using the previously ascertained fact that the maximum likelihood parameter estimates will correspond to the parameters that maximise $|\Sigma|$ (equation (6.14)). From equation (6.23) then, the maximum likelihood estimate of $[\beta]$ corresponds to the $[\beta]$ which maximises

$$|\Sigma| = \left|[S_{00}] - [S_{01}][\beta]\left([\beta]^T[S_{11}][\beta]\right)^{-1}[\beta]^T[S_{10}]\right|. \tag{6.25}$$

Using a matrix lemma from [91], this can be re-expressed as

$$|\Sigma| = \frac{|[S_{00}]| \cdot \left| [\beta]^T \left([S_{11}] - [S_{10}] [S_{00}]^{-1} [S_{01}] \right) [\beta] \right|}{\left| [\beta]^T [S_{11}] \ [\beta] \right|} \tag{6.26}$$

Now, since $[S_{00}]$ is fixed, $|\Sigma|$ is maximised by maximising

$$\frac{\left| [\beta]^T [M] [\beta] \right|}{\left| [\beta]^T [N] \ [\beta] \right|} \tag{6.27}$$

where $[M] = [S_{11}] - [S_{10}] [S_{00}]^{-1} [S_{01}]$, $[N] = [S_{11}]$. Utilising a second lemma from [91] (Lemma A.8), for $[M]$, $[N]$ symmetric and positive definite, the ratio (6.27) is maximised by $[\widehat{\beta}] = (\{v_1\}, \{v_2\}, \ldots, \ \{v_r\})$, with the maximal value equal to $\prod_{i=1}^{r} \lambda_i$, where $\{v_i\}$ and λ_i are the solutions of the generalised eigenvalue problem

$$(\lambda [N] - [M]) \{v\} = 0. \tag{6.28}$$

From the same lemma, any $[NS][\hat{\beta}]^T$ can be chosen as the maximising argument of (6.28), where $[NS]$ is any non-singular $r \times r$ matrix. This allows normalisation of the cointegrating vectors found.

By substituting the relevant $[M]$ and $[N]$ into (6.28) the final generalised eigenvalue problem to be solved is

$$\left(\lambda_i' [S_{11}] - [S_{10}][S_{00}]^{-1}[S_{01}] \right) \{v_i\} = 0 \tag{6.29}$$

where $\lambda_i = 1 - \lambda_i'$. It is also interesting to note that with appropriate normalisation $\{v_i\} [S_{11}] \{v_j\} = \delta_{ij}$. With this appropriate normalisation for $[\widehat{\beta}]$ the following hold true;

$$[\hat{\beta}]^T [S_{11}] [\hat{\beta}] = [I] \tag{6.30}$$

and

$$[\beta]^T [S_{10}] [S_{00}]^{-1} [S_{01}] [\beta] = diag(\lambda_1, \ldots, \lambda_r). \tag{6.31}$$

Upon solving equation (6.29), the relevant cointegrating vectors are found. It turns out that the choice for the 'best' cointegrating vector corresponds to the largest eigenvalue of (6.29). The eigenvalues λ_i measure how strongly the cointegrated relation correlates with the stationary part of the process. The larger the eigenvalue, the 'more stationary' the cointegrated relation.

6.3.4 The Johansen test statistic

Upon solving the eigenvalue problem (6.29) the required cointegrating vector has been found. The final step, however, is Johansen's test for cointegration. For econometricians, this test is the key point to the procedure, as it is that which verifies whether the variables under consideration are in fact cointegrated or not. From an engineering perspective, the relationships between a set of monitored variables are often much better understood, which means that the question of whether they are cointegrated or not is less important than the one of, 'will the residual created stay within a set of limits whilst in normal condition' or not, which can more often than not be verified visually. Having said that it is always useful to have a way to quantify and clarify any analysis carried out.

Johansen's test for cointegration, then, depends upon the rank of the matrix [Π] in equation (6.11). Recall that, for $I(1)$ variables, the matrix [Π] in (6.11) is required to be rank deficient if the error correction model is to hold true. If [Π] has full rank the $I(1)$ variables cannot be cointegrated. Therefore, a test for cointegration can be based on the rank of [Π]. Johansen's approach to this is to use a likelihood ratio test where the hypothesis $H(r)$, of [Π] having rank r is tested against the hypothesis $H(n)$, of [Π] having full rank. Recall from (6.14) that the maximised likelihood depends on the estimated covariance matrix of the error residual in the error correction model. Using relations (6.30) and (6.31), and the lemma (A8) from [91], (6.14) can be re-expressed as

$$\begin{aligned}(2\pi)^{-n}\mathcal{L}_{MAX}^{-\frac{2}{N}} &= |\Sigma|_{MAX} \\ &= |[S_{00}]| \prod_{i=1}^{r}(1-\lambda_i)'\end{aligned} \tag{6.32}$$

The likelihood ratio test, Q, then takes the form

$$\begin{aligned}Q(H(r)|H(n)) &= \frac{|\Sigma(\beta_{rank=r})|}{|\Sigma(\beta_{rank=n})|} \\ &= \frac{\mathcal{L}_{MAX(rank=r)}^{-\frac{2}{N}}}{\mathcal{L}_{MAX(rank=n)}^{-\frac{2}{N}}} \\ &= \frac{\prod_{i=1}^{r}(\lambda_i'+1)}{\prod_{i=1}^{n}(\lambda_i'+1)}\end{aligned} \tag{6.33}$$

This leads directly to the 'trace statistic', λ_{trace}, used to test the hypothesis that there are at most r cointegrating vectors.

$$\lambda_{trace} = 2\log Q(H(r)|H(n)) = -N\sum_{i=r+1}^{n}\log(1-\lambda_i') \tag{6.34}$$

The asymptotic distribution of this test statistic depends on the type of deterministic trend in the model (6.11). The relevant table with critical values can be found in [91].

Juselius [94] recommends using the trace test statistic in the following way; to begin the hypothesis of $r = 0$ is tested against one of $[\Pi]$ having full rank. If the test statistic is smaller than the relevant critical value accept the hypothesis of $r = 0$, i.e. no cointegration. If it is larger, reject the hypothesis and move on to test $r = 1$. In general, if the test statistic is larger than the relevant critical value, one should reject that there are as few as r cointegrating vectors, and move on to test the next largest possible rank r. In this way, the hypothesis of *r=0* up to *r=n* (i.e. full rank) is tested.

This concludes the cointegration theory section. The next section provides a summary of this theory which omits any derivation, and only includes the steps needed

if one were to implement the cointegration test procedure for SHM purposes.

6.4 Summary of the cointegration process for SHM

The previous sections have outlined the sometimes complex mathematics around the theory of cointegration. For the purposes of this work, the Johansen procedure is used to find the most stationary linear combination of a set of monitored variables. Although many steps are necessary in the derivation of the Johansen procedure, as outlined in the preceding section, its application to the kinds of problem in question can be achieved in just a small number of steps. The procedure and application of cointegration for the creation of damage sensitive features independent of the effects of the environmental and operational conditions is summarised below.

1. The suitability of the monitored variables to application of the cointegration process should first be assessed. To be included in the analysis each monitored variables should be integrated of the same order, in other words they should have the same 'degree' of nonstationarity. Furthermore, for application of the Johansen procedure, each monitored variable should be integrated order one $\sim I(1)$, i.e. a nonstationary variable with first difference stationary. Information of this kind is obtained by an ADF test on each variable. To carry out an ADF test:

   (a) Fit each variable in question to the following time series model using a least-squares approach.

   $$\triangle y_i = \rho y_{i-1} + \sum_{j=1}^{p-1} b_j \triangle y_{i-j} + \varepsilon_i \tag{6.35}$$

   where the difference operator $\triangle$ is defined as $\triangle y_{i-j} = y_{i-j} - y_{i-j-1}$. A suitable number of lags p should be included to insure that ε_i becomes a white noise process ([99]).

   (b) The least-squares estimate of the parameter ρ is used to infer the degree of nonstationarity of the variable. If ρ is statistically close to zero the process will be nonstationary and integrated order one. The null hypothesis of $\rho = 0$ is tested by comparing the value of the test statistic;

$$t_\rho = \frac{\widehat{\rho}}{\sigma_\rho} \tag{6.36}$$

where $\widehat{\rho}$ is the least squares estimate of ρ, and σ_ρ the variance of the parameter, with the critical values from the Dickey-Fuller (DF) tables (see [92]). The hypothesis is rejected at level α if $t_\rho < t_\alpha$. If the hypothesis is accepted, the time series has a unit root and is $I(1)$. If the hypothesis is rejected, the test should be repeated for $\triangle y_i$, if the hypothesis is then accepted y_i is an $I(2)$ nonstationary sequence. This can be continued until the integrated order of the time series is ascertained. Additional hypotheses and test statistics are needed if the model form needs to be extended to include shifts or deterministic trends (or both).

2. The Johansen procedure can now be applied to monitored variables found to be integrated order one. If the variables are cointegrated the Johansen procedure will find the linear combinations of them that will result in a stationary residual sequence purged of the common trends shared in the variable set.

   (a) Fit the variables in question to a vector auto-regressive model:

$$\{y_i\} = [A_1]\{y_{i-1}\} + [A_2]\{y_{i-2}\} + \cdots + [A_p]\{y_{i-p}\} + \{\varepsilon_i\} \tag{6.37}$$

   and determine for those variables the most suitable model order p using the AIC criterion or similar (See for example [100]).

   (b) The best linear combination of the variables, or cointegrating vectors, is found as the parameter $[\beta]$ in the vector error correction model (VECM) of the variable set which takes the form:

$$\{z_{0i}\} = [\alpha][\beta]^T\{z_{1i}\} + [\Psi]\{z_{2i}\} + \{\varepsilon_i\} \tag{6.38}$$

   Where $\{z_{0i}\} = \{\Delta y_i\}$, $\{z_{1i}\} = \{y_{i-1}\}$,
   $\{z_{2i}\} = \{\{\Delta y_{i-1}\}^T, \{\Delta y_{i-2}\}^T, \ldots, \{\Delta y_{i-p}\}^T, \{D\}^T\}^T$, p is model order ascertained previously in (a) and $\{D(t)\}$ is a deterministic trend.

   To find $[\beta]$, the cointegrating vector, first establish the residuals $\{R_{0i}\}$ and $\{R_{1i}\}$ of the following regressions:

$$\begin{aligned} \{z_{0i}\} &= [C_1]\{z_{2i}\} + \{R_{0i}\} \\ \{z_{1i}\} &= [C_2]\{z_{2i}\} + \{R_{1i}\} \end{aligned} \tag{6.39}$$

(c) Next, define the product moment matrices:

$$[S_{mn}] = \frac{1}{N}\sum_{i=1}^{N}\{R_{mi}\}\{R_{ni}\}^T \qquad m,n = 0,1. \tag{6.40}$$

Now, the required cointegrating vectors are found as the eigenvectors of the generalised eigenvalue problem:

$$\left(\lambda_i'[S_{11}] - [S_{10}][S_{00}]^{-1}[S_{01}]\right)\{v_i\} = 0 \tag{6.41}$$

The cointegrating vector that will result in the most stationary combination of the original variables will be the eigenvector with the corresponding largest eigenvalue.

3. Once a suitable cointegrating vector has been found, new data from the monitored variables should be projected onto it. If the cointegrating vector was established on data from the normal condition of the structure, the residual sequence from the linear combination will continue to be stationary all the time the structure continues to operate in its normal condition. The residual sequence should therefore be continually monitored and deviations from stationarity taken to indicate a deviation from the normal condition of the structure.

4. Finally a trace test can be carried out to indicate the number of cointegrating vectors possible for a set of variables. To test the hypothesis that there are at most r cointegrating vectors, the trace test statistic is used;

$$\lambda_{trace} = 2\log Q\left(H(r) \mid H(n)\right) = -N\sum_{i=r+1}^{n}\log(1-\lambda_i') \tag{6.42}$$

Normal procedure is to first test the hypothesis of *r=0* (no cointegration). If

the test statistic is smaller than the relevant critical value accept the hypothesis of *r=0*. If it is larger, reject the hypothesis and move on to test *r=1*. In general, if the test statistic is larger than the relevant critical value, one should reject that there are as few as r cointegrating vectors, and move on to test the next largest possible rank r. In this way, the hypothesis of *r=0* up to $r=n$ (i.e. full rank) is tested.

When employing the steps described above in coded form the choices made by an operator are very small in number. A range of model orders (number of lags in (6.37)) under investigation should be specified, although proper use of model fitness measures such as the AIC criterion will allow one to check that a suitable model order has been found. Another choice that arises concerns the selection of the significance level employed when using the ADF test statistic. The final and most important choice available to one wishing to implement the above steps is the selection of the training period on which the cointegrating vectors will be established. The training data should come from a period where confidence is high that the structural response is normal.

6.5 Conclusions

In this chapter the, sometimes complex, theory of cointegration from the field of econometrics has been introduced. The concept for its use for SHM is to exploit the cointegrated property of nonstationary damage sensitive features in order to find a stationary linear combination of them. A stationary linear combination of cointegrated variables can be found using the Johansen procedure and if this is achievable using feature data from a structure operating in its normal condition, the stationarity of the linear combination can then be used as an indicator of structural condition. If the way feature variables interrelate change, as might be expected if damage had been introduced to the structure, the cointegrated linear combination will become nonstationary. In the next chapter this theory is applied to data from the Tamar monitoring campaign. Chapter 7 also includes a discussion of why some of the econometric theory described in this chapter is applicable to engineering data.

Chapter 7

Applying cointegration for the data normalisation problem

The previous chapter introduced how the concept of cointegration could be used to remove environmental and operational trends from data. This chapter begins with some general discussion about why cointegration is applicable to engineering data, the chapter then concludes with the application of the theory to data from the Tamar monitoring campaign.

In the following discussion the author will argue that measured responses from healthy structures exhibit nonstationary behaviour over relatively short time periods, but should generally be stationary in the long term. As one would wish to detect any occurrence of structural degradation swiftly, these shorter time periods are of great interest to SHM. The discussion below will argue that such nonstationarity exhibited by structural response variables may be well represented by econometric theory, which models nonstationarity with unit root processes or with a deterministic time trend (or both). While under the time periods of interest such models of nonstationarity are valid, and therefore, too, the applied theory of cointegration in the previous chapter, the acceptance of a unit root generating process for a structural response is not a comfortable one. The author will argue that so long as the econometric models fit the data well, the philosophical question of whether a unit root assumption is valid is not really of interest. Indeed, to the engineer, so long as the combination of variables found via the Johansen procedure, or any other cointegrating method, does the job of removing confounding influences, the underlying

assumptions of the process used to arrive at this combination may not be of any interest. None the less, a good understanding of such assumptions may help to avoid any suboptimal use of the sophisticated theory introduced in the previous chapter.

7.1 The assumptions behind cointegration applied to engineering data

In this thesis the idea of cointegration has been adopted from econometrics for the purposes of removing environmental and operational trends in damage sensitive features. The theory presented in the previous chapter assumed that the variables of interest were nonstationary and, for the Johansen procedure, integrated of order one (i.e. a unit root process). This section aims to discuss whether the unit root assumption is valid in the context of an engineering application and whether structural response can be nonstationary in the long term.

A nonstationary time series is defined as one whose statistics change with time. A weakly stationary process has constant mean and variance, and an autocovariance that depends only on the lag length considered. In the presence of nonstationarity standard regression techniques and inferences made on such regressions have been found to be unreliable, and since Yule's seminal paper [101], much research has been carried out on nonstationary processes.

Of the variables of interest to SHM, many of them appear to exhibit nonstationarity, some examples of which have already been introduced in the thesis. The most relevant example to this chapter is the easterly displacement of the deck at hangar 44 of the Tamar bridge, plotted in Figure 7.1. Studying Figure 7.1, the time series certainly appears nonstationary over the time of available measurements.

To an engineer, when considering a dynamic response the usual modelling approach would be to employ the classic second order differential equation of motion. The nonstationarity of a response is then explained by the changing physical parameters such as mass and stiffness, or a change in the nature of the excitation. In the context of applying cointegration theory, the interest is in modelling the entire process (response) with fixed parameters that cannot change. For this task, the most comfortable way of thinking about a structural response is probably as a function of a number of external conditions which are themselves fluctuating. One might expect

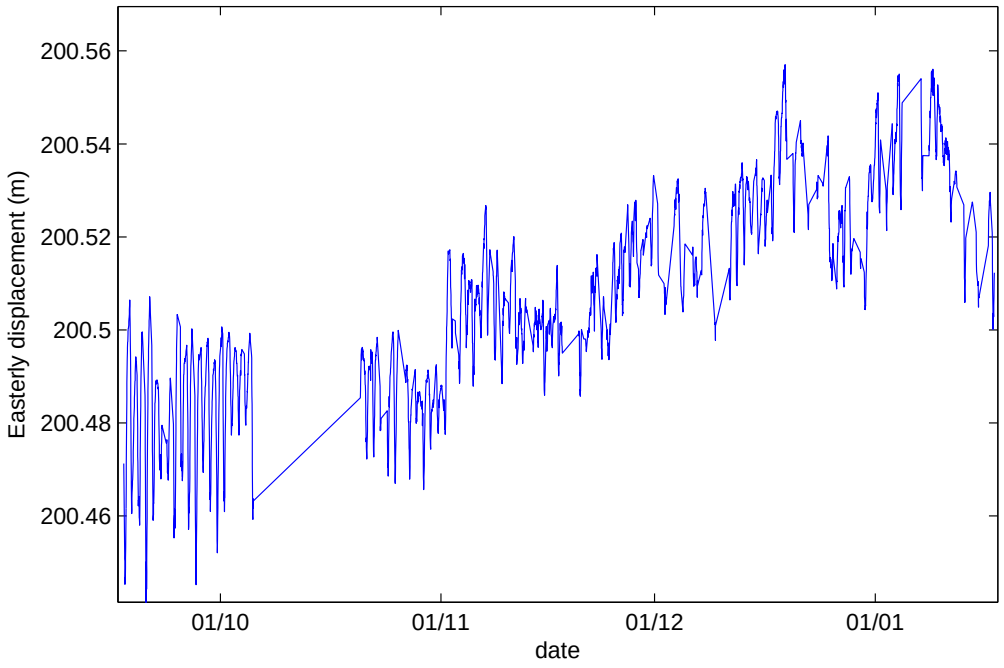


Figure 7.1: Change of easterly deck deflection over a number of months

that if all external conditions could be accounted for, a structural response would be stationary around the trends introduced by these external conditions (indeed, this assumption is implicit to the regression applied in Chapter 4). Unfortunately, it may be very infeasible to account for all external conditions driving the response of a structure, as it certainly is in the field of econometrics where relations between variables of interest are uncertain. This is where an AR type model can be useful and why they are regularly applied to econometric time series, as one no longer needs information about each driving factor to be able to describe a dynamic process. The cointegration theory in the previous chapter was developed in the context of such models, and so to utilise this theory one must adopt them too.

In the theory described in the previous chapter, an error correction model was used to model a process variable, y_i of interest:

$$\triangle y_i = \rho y_{i-1} + \sum_{j=1}^{p-1} b_j \triangle y_{i-j} + \mu + \nu t + \varepsilon_i, \tag{7.1}$$

When considering such models, nonstationarity of a time series can be prescribed to two different mechanisms; either a deterministic trend νt or a unit root. Where a

deterministic trend is the cause of nonstationarity, econometricians refer to the time series as a trend stationary process; fluctuations around this deterministic trend are stationary in nature. Where a unit root is the cause of nonstationarity, the time series is referred to as difference stationary processes, due to the fact that difference operations will render the series stationary.

Now considering structural response variables that exhibit nonstationarity, any term in a descriptive model that continually increases or decreases in time is unlikely to mimic the behaviour of a stable structural response. This implies that to continue using an autoregressive type model, the nonstationarity of the variable must be described by a unit root process.

Many debates surround the idea of unit root modelling and its suitability to real life applications, within and outside economics (for example, there is currently much debate around using unit root processes to model the Earth's temperature change (see for example [102–104])). The idea of a variance that increases with time, which is inherent to a unit root process, certainly sits uncomfortably when considering a structural response over time. However, while it may be that unit root processes do not ideally suit the dynamics of a structural response, the framework in which the cointegration theory has been borrowed from has been established on these types of models, and for the present it seems sensible to utilise what developed theory one can. If the ECM model structure can describe the variables of interest in SHM well enough, then the philosophical question on unit roots becomes unimportant, and the theory in the previous section can be applied without further ado. As, for the present, the interest is not in accurate forecasting of a structural response variable a potential model misspecification is less important.

For the purposes of utilising the cointegration theory, the author suggests that so long as the variables of interest to an engineer follow a similar behaviour to a unit root process over a given time interval, then the theory described in the previous chapter is applicable. In this case, the theory can happily be applied to process variables which are nonstationary with first difference stationary, which at the very least may be roughly checked visually with ease.

Returning to the previous example, the first difference of the easterly deck displacements are plotted in Figure 7.2. Studying this figure it is clear that the first difference is a stationary process. As the variable is nonstationary with first difference stationary, the Johansen procedure is applicable.

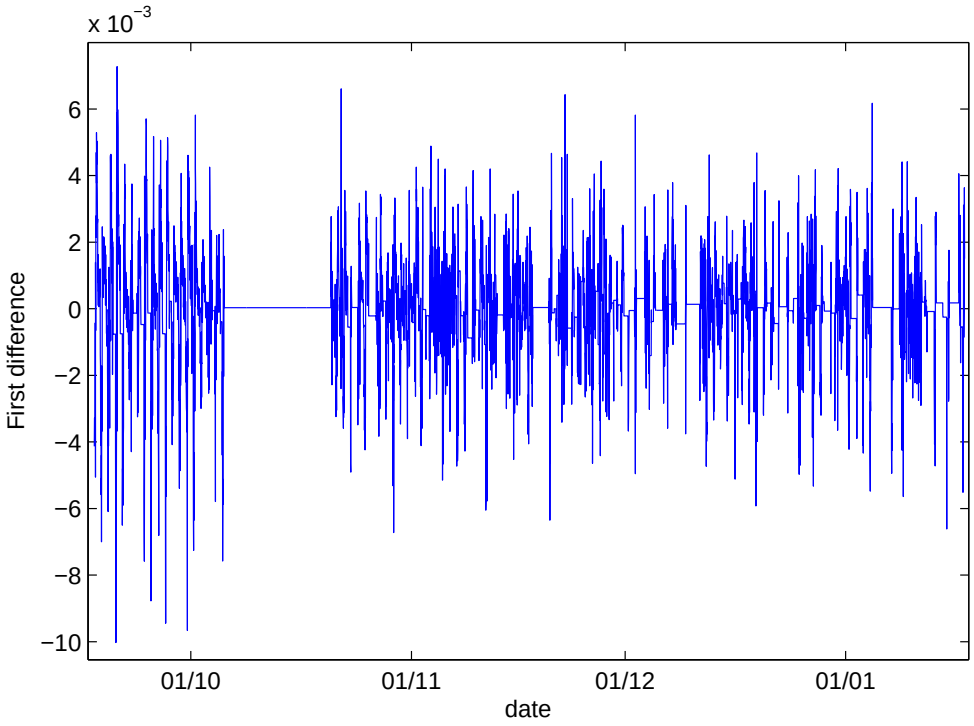


Figure 7.2: First difference of easterly deck displacements as plotted in Figure 7.1

The author believes that, although SHM variables may exhibit nonstationary behaviour over observable windows, in the nature of engineering, these response variables should be stationary when considering long periods of time. By design, engineering variables from stable healthy structures should be describable by a long run mean and a variance. In the long term one can therefore describe them as stationary, which enables one to use the regression techniques applied in Chapter 4, for example. The interests of SHM, however, are focused on response variables over shorter time periods, as any useful detection of structural degradation should be swift. One must therefore rely upon nonstationary theory and, in this work, cointegration. In the circumstance where a set of variables of interest include stationary and nonstationary measurements, the inclusion of the stationary variables doesn't invalidate the Johansen Procedure.

7.2 Application of cointegration to data from the Tamar bridge

As described in previous chapters, the Vibration Engineering Section (VES) at the University of Sheffield have, over the last few years, developed various monitoring systems for the Tamar Suspension Bridge in Southwest England. Modal parameters are automatically identified by data-driven stochastic subspace identification (SSI); environmental and operational conditions are monitored by a large network of sensors, and most recently, a total positioning system (TPS) has been installed to measure reliably the movement of the bridge deck and towers. Up to three years of dynamic, static and environmental data are now available for analysis. In the following, areas where cointegration could be of use for the Tamar monitoring campaign are highlighted.

One of the newest additions to the Tamar monitoring campaign is a TPS, which uses a robotic total station (RTS) for precise monitoring of the displacement of the bridge deck and towers (accurate to within 2 or 3mm). Figure 7.3 displays a plot of a number of longitudinal deflections of the deck and tower measured over a time span of just over three days (77 and half hours). From this plot the daily fluctuations in deck and tower displacements caused by operational and environmental conditions are obvious, but without further investigation the plot is somewhat unenlightening and uninformative for making any judgement about the structure's performance or condition. However, if the variables in Figure 7.3 are cointegrated, the stationary linear combination of the variables that can be found (via the Johansen procedure) could be used as a measure of normality of structural response. As described in the previous chapter, the procedure would be to find the best cointegrating vector of a training set of data from a period where one has high confidence that the structure's responses are normal and representative. All new data would then be projected onto the established linear combination, which should remain stationary all the time the structure operates in its normal condition. In the event that the cointegrated residual becomes nonstationary (which could be monitored with a statistical process control (SPC) chart, see [86]) one has an indication that the way the variables interrelate has changed, and further investigation into the cause can be carried out.

As perhaps expected from the plot, the variables in Figure 7.3 admit to a stationary linear combination, which is shown in Figure 7.4, along with error bars at plus and

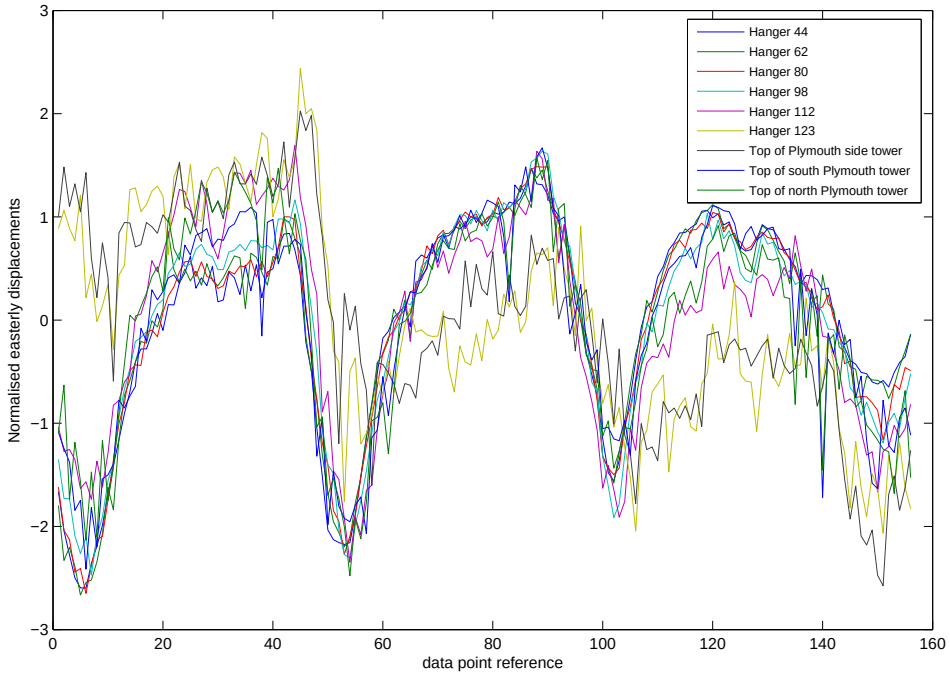


Figure 7.3: TPS measurements of bridge displacements (Eastings) over three days

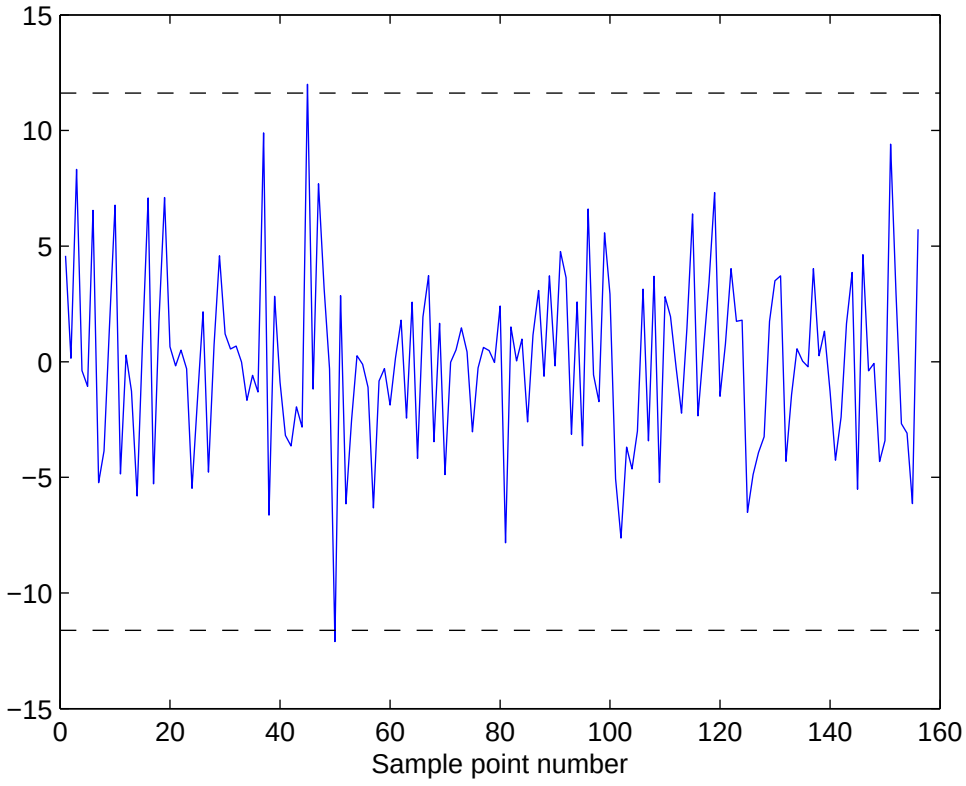


Figure 7.4: Cointegrated residual of variables in Figure 7.3

minus three standard deviations of the residual added to act as an SPC control chart. From 7.4 it is clear that the environmental or operationally induced trends have been purged. As the data in Figure 7.3 is considered to have come from normal operating conditions, the stationarity of the linear combination can be considered to represent a normal operating condition.

This can be tested by looking at a longer period of displacement data; Figure 7.5 shows the same easterly displacements plotted in Figure 7.3 over a two month period (with any gaps in the record removed), a seasonal trend is clearly visible resulting from colder temperatures. Figure 7.6 is a plot of this data projected onto the linear combination established on the data shown in Figure 7.3, the same error bars as in Figure 7.4 have been added to continue to act as an SPC control chart. The residual continues to remain stationary for the duration of the two months, which indicates that the relationships between the considered variables remain the same. Importantly one can clearly see that the seasonal trend visible in Figure 7.5 has been purged, which happily demonstrates how useful cointegration may be to the SHM community.

While the above example has demonstrated that cointegration can be used to remove environmental and operational variation from displacement monitoring data, the aim is to use the residual as a condition or performance indicator, and so in order to test this a known performance anomaly is needed. In fact, despite its age the bridge performs so well that neither VES nor the bridge operators have found such an anomaly during the monitoring, and so the only alternative is to fabricate one. For illustrative purposes, one could manipulate the data in an attempt to mimic/simulate a potential abnormal change in structural response and see the effect on the trained cointegrated residual. Here, in a crude attempt to simulate the possible effect of a loss in tension in one of the stay cables (specifically the stay cable labelled P2 in Figure 3.6 in Chapter 3), additions are made to some of the displacement channels of the data shown in Figure 7.5. After an arbitrary number of data points (500 in this case), an additional 5mm was added to the longitudinal displacements of the deck 44m away from the Saltash end of the bridge (the reflector placed between the stay cables labelled S2 and S4 in Figure 3.6), 4mm was added to each reading of the deflections at 62m away from the Saltash end of the bridge (the centre of the bridge), and 3mm was added to each reading of longitudinal displacement 80m away from the Saltash end of the bridge (between the stay cables labelled P4 and P2 in Figure 3.6). The corrupted data when plotted on the same scale as

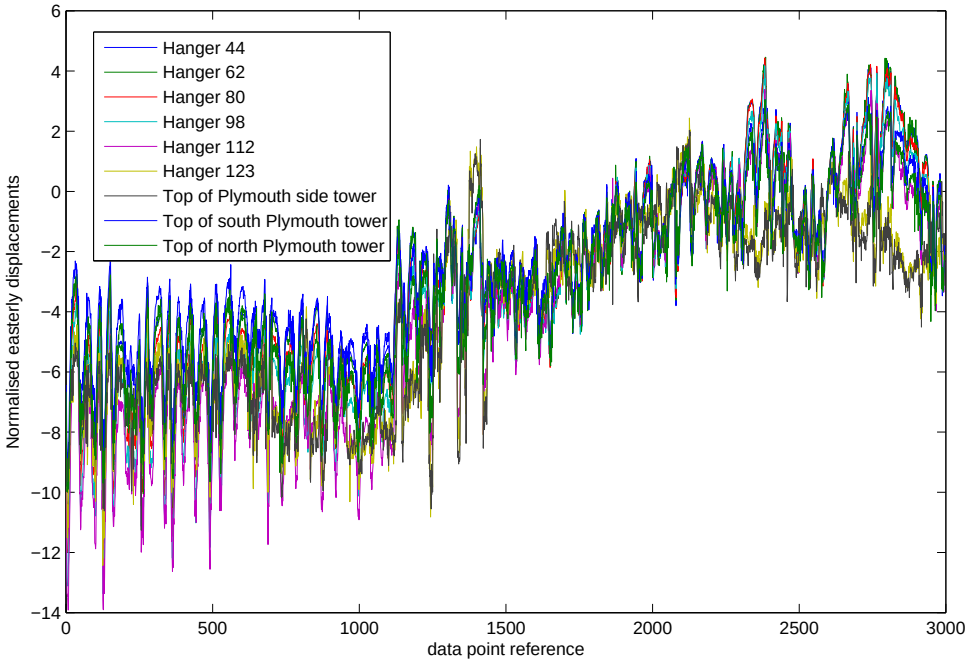


Figure 7.5: TPS measurements of bridge displacements (Eastings) over a two month period

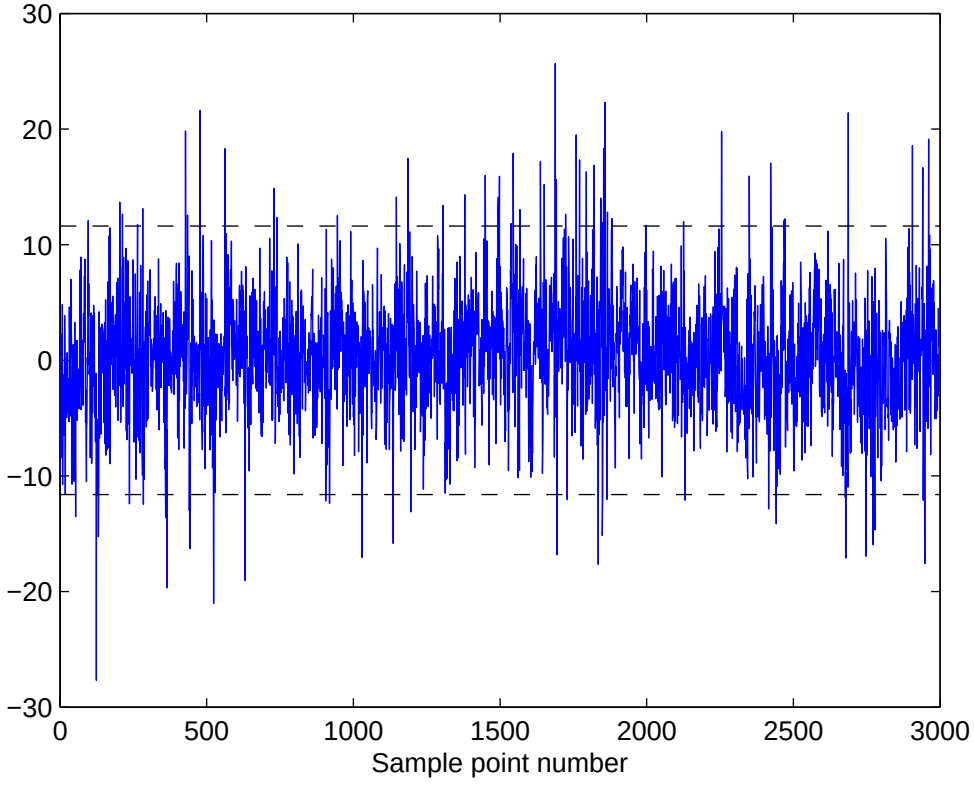


Figure 7.6: Data collected over two months projected on cointegrated residual established in on data in Figure 7.3

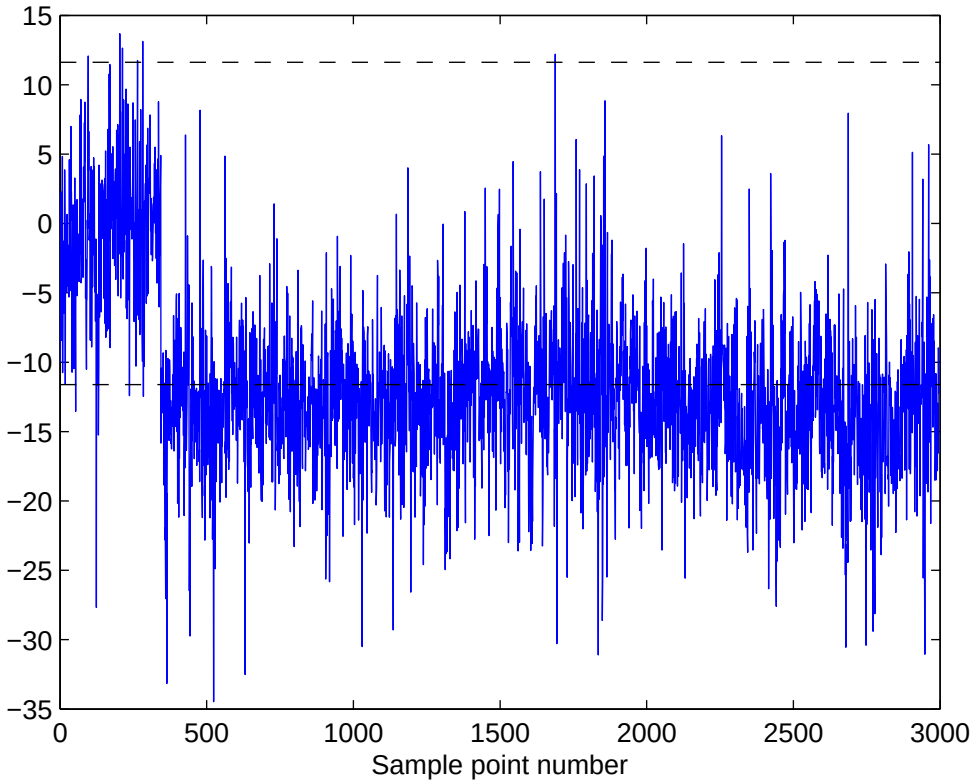


Figure 7.7: Cointegrated residual with manipulated data

in Figure 7.5 is indistinguishable from Figure 7.5, as so is not included here. The projection of this corrupted data onto the stationary linear combination established with the (un-tampered with) data from Figure 7.4 is shown in Figure 7.7. After 500 data points there is a clear shift in the residual, which could clearly be taken as an indication that the bridge is no longer operating in its normal condition.

As a further example of the application of cointegration to this monitoring campaign the natural frequencies of the deck are considered. Figure 7.8 shows a plot of the lowest five natural frequencies recorded over the duration of one month, again operational and environmentally induced trends are clearly visible. Interestingly, one can see the long term stationarity of the natural frequencies along with the (daily) environmental and operational trends. The aim here when applying the Johansen procedure is to create a useful variable free of the trends visible in Figure 7.8. The linear combination of these variables found by the Johansen procedure is plotted in the higher plot in Figure 7.9. In this instance the Johansen procedure was trained on the first 500 data points of the set plotted in Figure 7.8. Over the month long

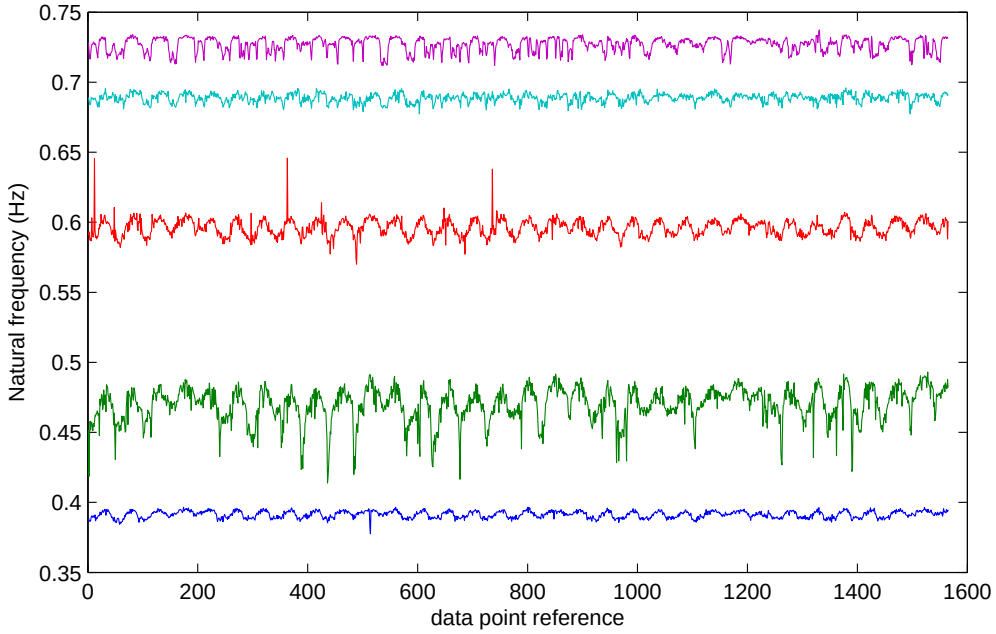


Figure 7.8: First five deck natural frequencies recorded over a one month period

period, the linear combination is stationary, as would be expected, the daily variations are now no longer clear. As before, this stationary linear combination can be used as a measure of normality; one would expect the linear combination to remain stationary all the time the bridge responds in a normal condition. The lower plot in Figure 7.9 is of new data projected onto the established cointegrating vector with error bars at plus and minus three standard deviations of the residual during the 500 point training period. This data was collected over a period of six months almost a year and a half later than the data used to choose the cointegrating vector. From observation, the residual remains fairly stationary over the 6 month period indicating that relationships between the frequencies have remained constant. There is, however, one large anomalous area in the plot, where, around data point 4000, the deviation of the residual becomes much larger, and the error bars are exceeded many times consecutively. This large anomaly indicates that something untoward happened during this period of the monitoring campaign. On further investigation, this is the anomaly encountered earlier in the thesis (see Chapter 5), which occurs at a time when one of the accelerometers became waterlogged and was therefore providing a corrupted signal. This event temporarily destroyed the equilibrium of the linear combination, which in itself is testament to the method's ability to detect anomalous events.

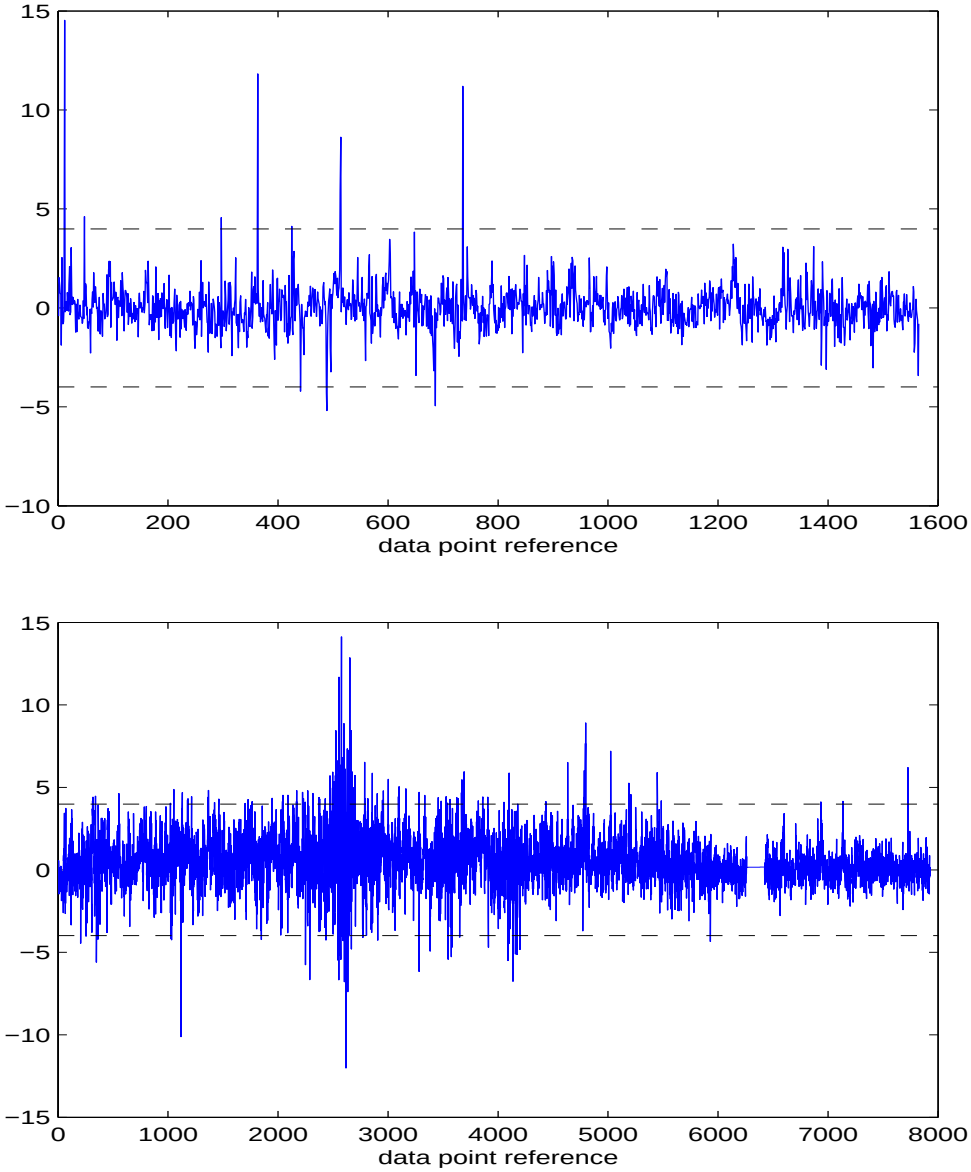


Figure 7.9: **Above**. Cointegrated residual of first five natural frequencies plotted in Figure 7.8 (cointegrating vector established on first five hundred points only). **Below**. Data from later on in the year projected onto the same cointegrating vector.

7.3 Conclusions

Econometricians understand nonstationarity in terms of trend stationary and difference stationary processes. This chapter has argued that although neither approach may be philosophically suited for SHM feature variables, these feature variables do mimic the behaviour of difference stationary processes, for which the Johansen procedure was developed for. With this in mind, no problems should arise with applying the theory presented in the previous chapter to SHM data.

The second half of the chapter applied cointegration theory to data from the Tamar monitoring campaign, where it was able to remove environmental and operational trends for deck displacement and natural frequency data. The environmentally insensitive features created were able to detect a simulated damage scenario as well as a sensor malfunction.

Chapter 8

A comparison of cointegration and PCA

Previously in the literature the idea of using projection of data onto the minor components obtained in a PCA to remove unwanted environmental and operational trends has been investigated [63]. This chapter explores the similarities and differences between PCA and cointegration, which on the surface seem to be similar ideas. Both approaches are investigated here in the context of data from the Brite-Euram project DAMASCOS (BE97 4213), which was collected from a Lamb-wave inspection of a composite panel subject to temperature variations in an environmental chamber. Original results for the application of PCA to this benchmark were presented in [63]. This chapter is intended to build on the work covered in [63], discussion of the results presented in this paper will be significantly expanded here, the results gained using PCA and outlier analysis will be compared with new results gained using cointegration.

This chapter will apply both cointegration and PCA to the Damascos data set with the aim of removing a temperature dependent trend from damage sensitive features. As previously described in Chapter 4, PCA projects data onto a new set of orthogonal axes (or principal components) which are linear combinations of the originals but ordered according to the proportion of the variance of the data each accounts for. In the case of an undamaged structure subject to changing environmental conditions, information on the response of a set of monitored variables to the environmental variation will be contained within the principal components

that account for significant amounts of variance in the data set [63]. The idea explored here is to project a data set onto its *minor* components, i.e. those which account for less variance in the data, and therefore discard the dimensions of the data that carry any dependence on environmental factors. In theory, so long as damage does not manifest as variance along an axis in the same direction as any of the major components disregarded, the feature created using the minor components will be insensitive to environmentally induced structural responses but still sensitive to damage.

As in the previous chapter, when applying cointegration, the Johansen procedure is used to find the most stationary linear combination of variables possible, which can then be used as a univariate damage sensitive feature.

8.1 Experimental Data

The methods outlined in the previous section will be explored in this chapter in the context of data collected from the Brite-Euram project DAMASCOS (BE97 4213), which studied the damage detection capabilities of Lamb-wave propagation within composite structures. The data used here comes from a Lamb-wave inspection of a composite panel subject to temperature variations in an environmental chamber, of which the test set up is illustrated in Figure 8.1.

Identical piezoceramic discs were bonded at the plate edges to minimise reflections from these edges and at the mid-point of these edges to allow for greater discrimination between the direct propagating mode and its reflections from the side edges. The plate material was Carbon Fibre Reinforced Plastic (CFRP) with a 0°/90° lay-up. Fundamental symmetric (S_0) and anti-symmetric (A_0) Lamb-waves were launched by driving the transmitter with a 5 cycle toneburst from a signal generator at 300kHz and 80kHz respectively. The signals resulting at the sensor were monitored by digital storage oscilloscope then transferred to PC. Figure 8.2 shows a typical signal in the time and frequency domains.

For this particular test, Lamb-wave signals were recorded every minute. For the first 1355 signals (a period of approximately 22 $\frac{1}{2}$ hours) the chamber temperature was held at a constant 25°C. The temperature within the chamber was then decreased to 10°C before being ramped to 30°C over a three hour period then back to 10°C,

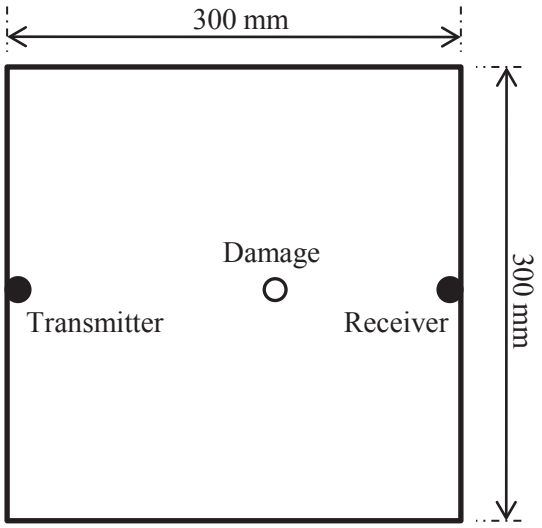


Figure 8.1: 3mm thick composite plate instrumented with piezoceramic transmitter

again over a period of three hours. This cycling was repeated for more than three further cycles. After approximately 41 hours (signal number 2483), the chamber was opened, a 10mm hole was drilled in the plate between the two sensors then the chamber was closed. This essentially means that there were three different phases to the test: signals 1 - 1355 are from the undamaged panel held at a constant 25°C, signals 1356 - 2482 are from the undamaged panel with temperature cycling and signals 2483 - 2944 are from the damaged panel with temperature cycling.

For the purposes of this work it was necessary to sub-sample the data collected from the test described above. 50 spectral lines from the area around the peak of the frequency spectrum, an example of which is plotted in Figure 8.2, are selected here as an area of interest (these are lines 46-95). The feature that will be studied here, then, is the amplitude of each of these 50 spectral lines for each of the 2944 signals recorded in the test. A time history of these spectral line amplitudes is plotted in Figure 8.3.

In order to understand the feature data better, preliminary outlier and principal component analyses were carried out (see Chapter 4 for more details on PCA and Chapter 5 for outlier analysis). For both of these a training data set was chosen as every second data point recorded when the temperature of the plate was held constant, in other words, taking the plate under constant temperature as the normal

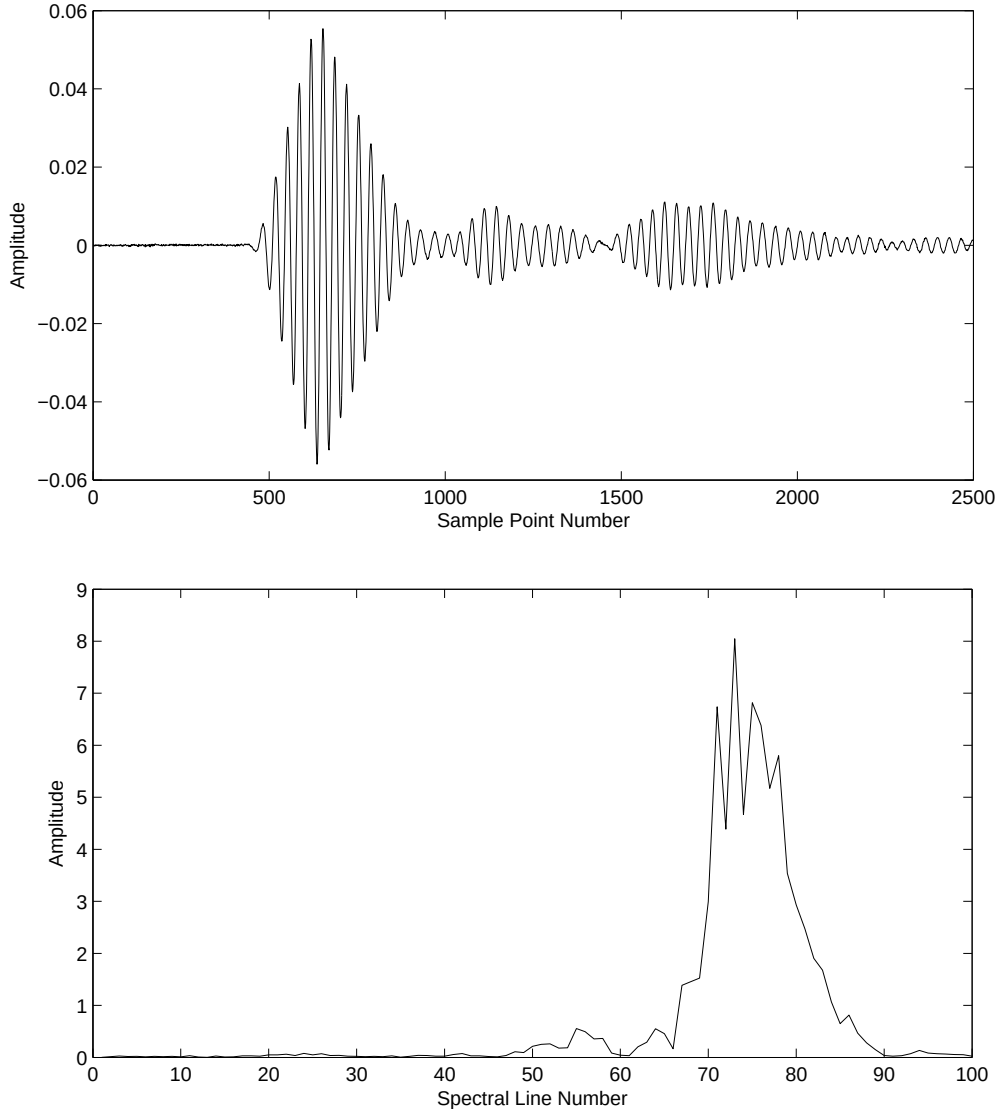


Figure 8.2: Typical Lamb-wave signal in time and frequency domains

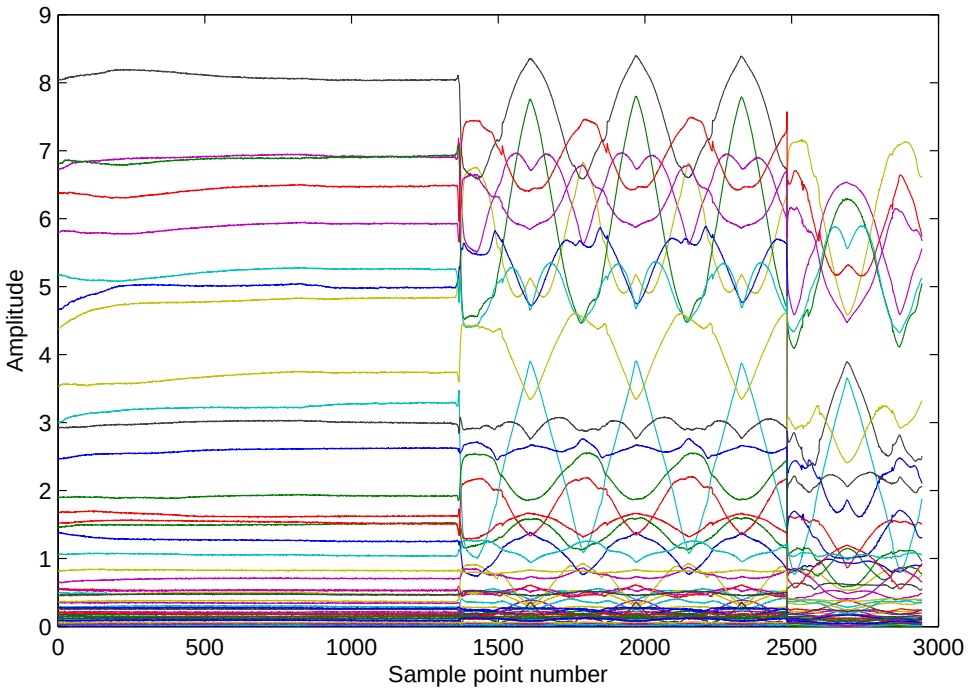


Figure 8.3: Time history of 50 dimensional feature

condition. For the outlier analysis the mean, $\bar{x}$, and covariance matrix, S, were calculated for the 678 training set samples. All feature samples were then in turn designated x_i and values for D_i, the novelty index, were calculated using equation (5.1). Figure 8.4 shows the results of this analysis with novelty index being plotted on a log scale (note that novelty index of the samples in the training set are also plotted). The horizontal dotted line represents the threshold value which is the critical value for a 1% test of discordancy (calculated using the training data), whilst the vertical lines separate the three regimes.

Not surprisingly, almost all of the novelty indices from samples in the constant temperature regime are below the threshold. Meanwhile, the features from the temperature cycling period and the damage set are all substantially over the threshold, indicating an abnormal response from the plate for the majority of the testing period. This is clearly an undesirable situation; if the outlier analysis was to be intended as a damage detector, responses from the plate under a changing temperature would be wrongly classified as such.

Principal component analysis is also carried out here to better understand the underlying structure of the response data from the three different regimes. A plot of

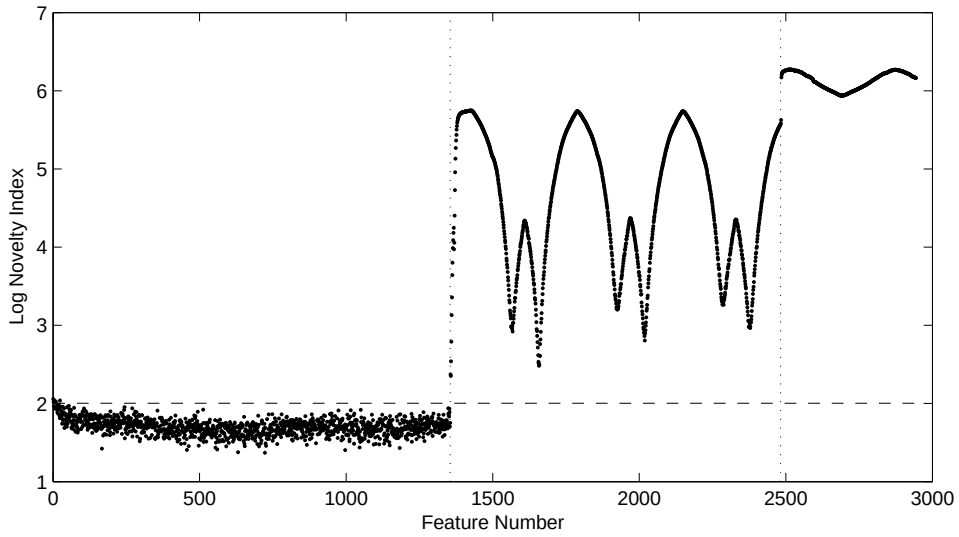


Figure 8.4: Results of outlier analysis using basic feature

the first two principal component scores is shown in Figure 8.5, where one can see that data from the three different regimes cluster separately, with very little overlap. The most important thing to note is that data from the undamaged response at a constant temperature does not overlie the 'undamaged' data from the temperature cycling period. The consequence of this, as for the outlier analysis carried out previously, is that a reliable damage indicator may not fabricated from the constant temperature measurements alone.

Having now a clear view of the data, the next section will explore how the effects of temperature can be dealt with to create a working novelty detector.

8.2 Results

Studying Figures 8.4 and 8.5 gives one insight into how badly a novelty detector would work if the constant temperature data were considered to define the normal condition; the temperature fluctuations lead to a false-positive detection of damage which is very undesirable. An obvious improvement should come from including data from the undamaged plate when the temperature was fluctuating in the training set. Figure 8.6 shows the results of the same outlier analysis carried out in the previous section, this time with the training data extended to include data from the fluctuating temperature regime (training data was specifically every second data

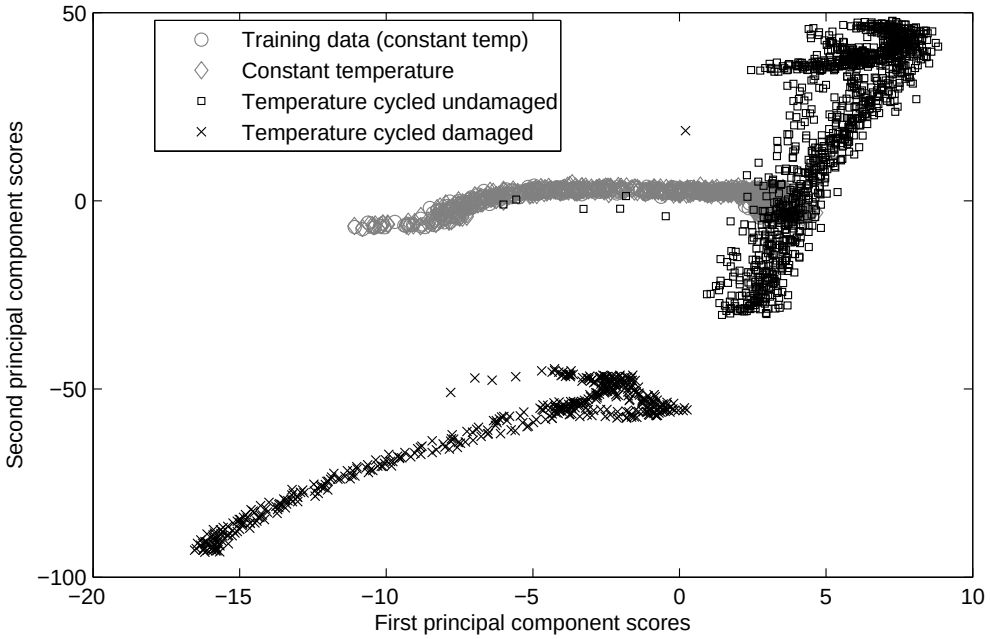


Figure 8.5: Plot of first two principal component scores, trained on constant temperature data

point up to data point 2000, this includes data from just under two full cycles of temperature fluctuation). On inspection of Figure 8.6, redefining the normal condition to include data points from the temperature fluctuating regime of the test has certainly decreased the discordancy of the data points from this regime, however, some structure still remains visible in the fluctuating temperature period and many points cross the threshold (indicated by the dashed line). In terms of damage detection this outlier analysis would still be very inappropriate.

8.2.1 Minor principal components for removing environmental sensitivity

The projection of damage sensitive features onto their minor components is explored here in attempt to remove temperature dependency. The method for doing this is simply to perform a PCA on the training data and the first two sets of testing data (from the uniform temperature period and the cycling temperature period) and discard the higher principal components which will account for the maximum variance in the data, which is expected to be due to the temperature variations. If and when these three sets of data from the unfaulted plate cluster together, the data

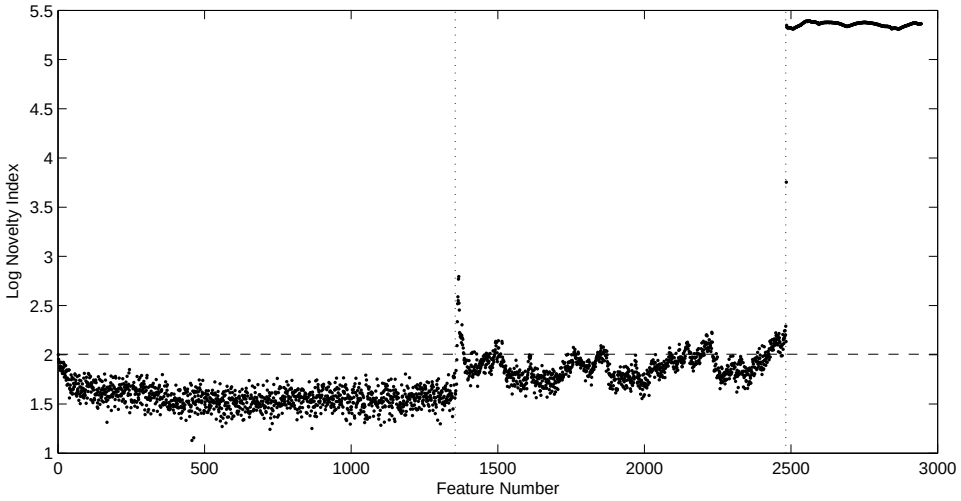


Figure 8.6: Results of outlier analysis using basic feature with extended training period

from the damage testing set may then be projected into the same minor component space.

For the basic features considered here, it was found that, by examining plots of principal component n versus principal component $n+1$, the vast majority of the variance due to temperature change was contained in about the first 10 components. However, in order to make sure that all three sets were overlapping, the last 10 principal components were used to form a new feature. The damage testing set was projected into the same space and an outlier analysis was performed using this new 10-point feature. The results shown in this section of the paper follow [63], where the principal components are calculated using every second sample of data recorded while the plate remained undamaged (both under stationary and cycling temperature), the outlier analysis uses a training data set, as described previously, of every second sample from the stationary temperature testing period. The results are shown in Figure 8.7, where it is obvious that this is an even more effective result than from the previous method. All of the temperature cycled, unfaulted data has been classified as unfaulted and there does not appear to be any cyclic behaviour to the novelty indices from this set. Also, all of the damage data is very clearly classified as such. This is a very encouraging result considering the complex nature of the data and also the temperature range considered. It should be noted, however, that the data has not been standardised prior to implementing the PCA, a fact that will be discussed further in the next section.

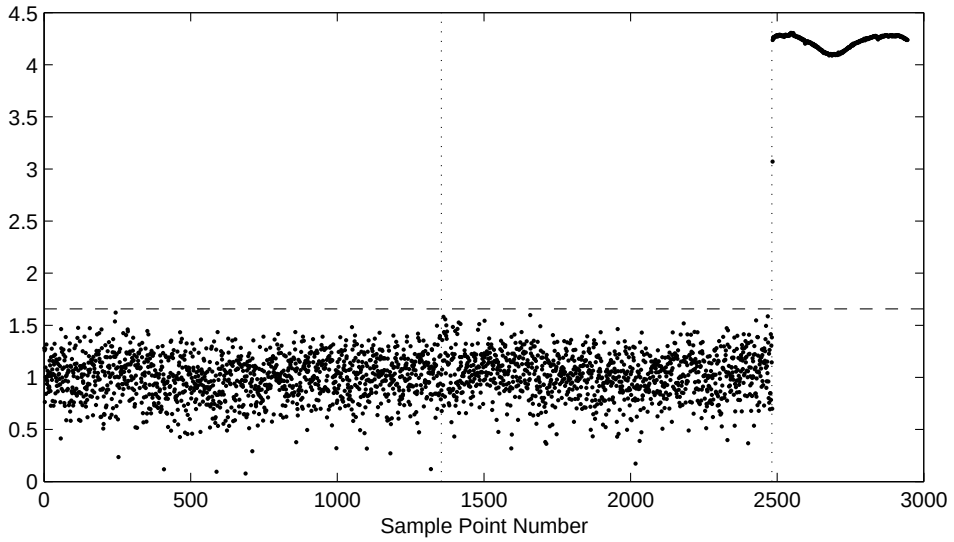


Figure 8.7: Results of outlier analysis using advanced feature calculated using minor principal components

8.2.2 Cointegration for the removal of environmental trends

The last method for creating environmentally insensitive damage detectors investigated here uses cointegration. As discussed in the previous chapter, cointegration has been developed in the context of nonstationary signals that can be described with a unit root process. In this study, the signals are clearly of a deterministic nature as they are driven by a temperature trend. As previously reasoned, the cointegration theory is applicable here as the signals of interest exhibit similar behaviour to a unit root process (i.e. stationarity of the original signals may be achieved through differencing).

Similarly to the previous methods, cointegration requires a training set of data from the normal condition of the undamaged structure. The Johansen procedure was used here to linearly combine the 50 features in question with the aim of creating a stationary residual. If a linear combination of the training data is stationary, the common trends shared by the 50 features (i.e. the temperature induced trends) will have been purged, any other abnormal change (such as the introduction of damage may cause) should then cause the combination residual to become nonstationary as long as each feature in question is not affected by the damage in a similar way. Figure 8.8 shows the linear combination of all 50 features for the training period chosen (data points 1000-2000, which includes 355 data points from the steady

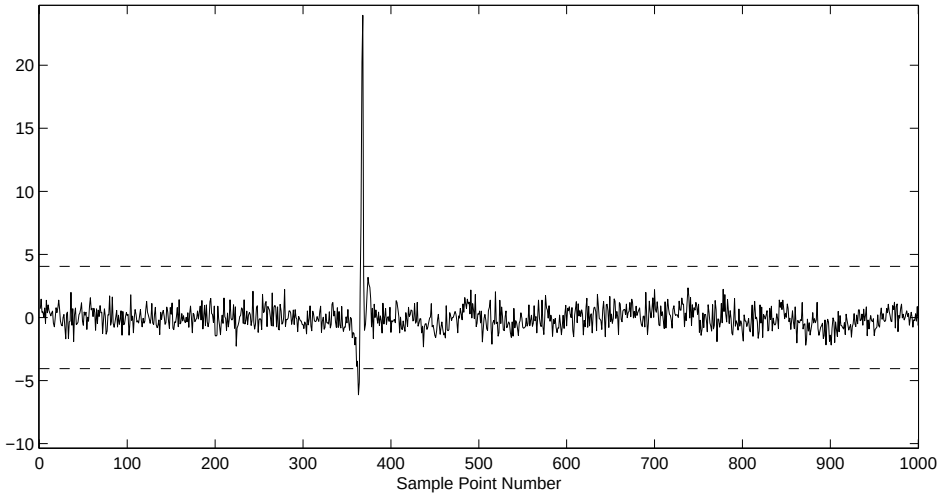


Figure 8.8: Cointegrated Signal over training period (linear combination of 50 spectral lines)

temperature regime and data from almost two temperature cycles). The dashed horizontal lines indicate plus and minus three standard deviations of the training data and are added to act essentially as a Statistical Process Control (SPC) X-chart ([86]), if a data point is outside of this threshold it can be considered as abnormal. Studying Figure 8.8, one can see that the Johansen procedure has successfully found a linear combination of the 50 features in question that is stationary over the training period, with the exception of a few points occurring around the time when the plate began to undergo its temperature cycles. This anomaly indicates that at the time of switching between the two test phases some more complex relationship between the environmental conditions and the recorded signals existed; happily after the transition period the features returned to an equilibrium quickly and are still valid as an anomaly detector.

As the Johansen procedure has successfully created a stationary combination of the variables from a training set it remains to project all of the rest of the data onto this combination and study what happens when damage is introduced. The results are shown in Figure 8.9, where the vertical lines indicate the beginning of the temperature cycling period and the point of the introduction of damage. A clear indication of damage is apparent when the residual becomes nonstationary and deviates significantly outside the control chart boundaries (at plus and minus three standard deviations of the training residual). Cointegration looks to be a very promising approach for the data normalisation problem.

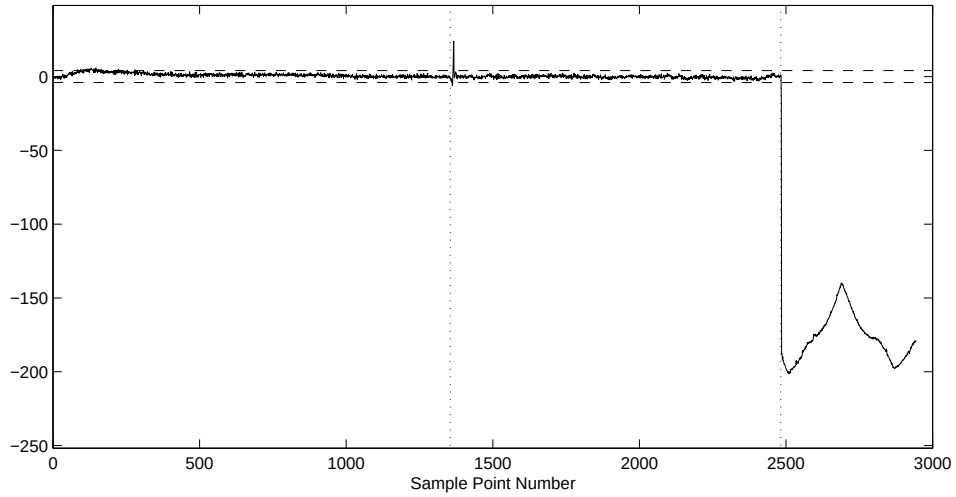


Figure 8.9: Cointegrated Signal over the whole duration of the test

Further to this result, it is interesting to note that the large anomaly visible in the combination of the training data (see Figure 8.8) is not present in the cointegrating combination when a subset of the 50 spectral lines is used. In this case the first twenty spectral lines from the feature set used previously was investigated. Using the same training period as before, the whole 20 feature data set projected on to the linear combination found by the Johansen procedure is shown in Figure 8.10, as before, the dotted horizontal lines indicate plus and minus three standard deviations of the training residual, and the two vertical lines indicate the introduction of the temperature gradient and the introduction of damage respectively. It seems that analysing a smaller sub-set of variables has eliminated the anomaly that previously occurred after the introduction of the temperature gradient, while the indication of damage is still very clear. Further discussion on this anomaly will follow in the next section.

8.3 A comparison of cointegration and principal component analysis

Cointegration and PCA have both been shown to be successful tools for the data normalisation problem in the previous section. As already alluded to in the introduction, they are in fact regarded in the field of econometrics as being from the same class of algorithms; both linearly combine multivariate data but by different

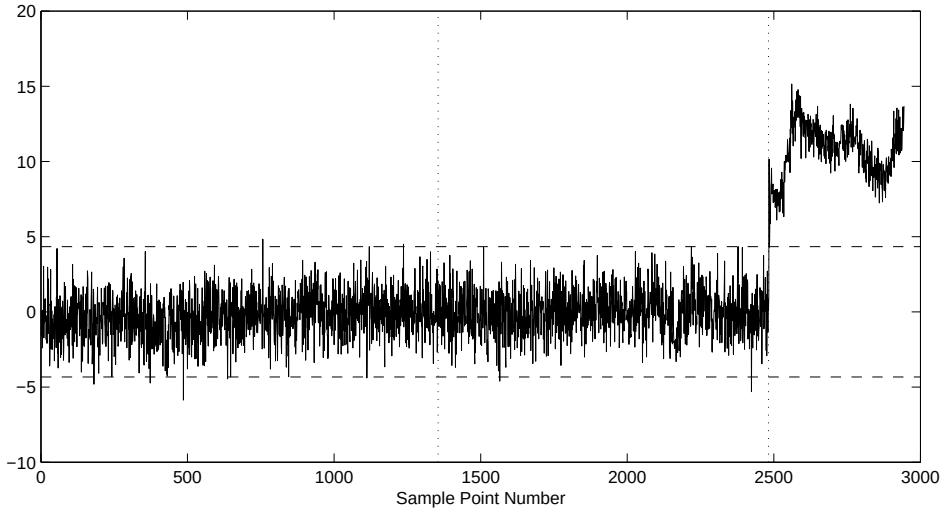


Figure 8.10: Cointegrated Signal (linear combination of 20 spectral lines)

means and for different objectives. PCA, using singular-value decomposition, creates and orders new variables according to amounts of variance each accounts for in the data, the Johansen procedure uses a maximum likelihood approach to evaluate the stationarity of a linear combination of variables and orders variables from the most stationary to the least (although only one cointegrating vector was used in the analysis above, the Johansen procedure will produce as many new variables, less one, as original variables included in the analysis). If one considers that the most stationary variable created by the Johansen procedure will most likely account for the least amount of variance in the data, loosely speaking, these two methods are doing roughly the same thing, only ordering the variables differently. In this way of thinking, the first n cointegrating vectors should be similar to the last n principal components for some multivariate data set.

To answer the question of how similar PCA and cointegration actually are in a mathematical way, a comparison between the set of principal components and the set of cointegrating vectors themselves should be made. The principal components in a PCA are (usually) computed using a singular value decomposition of the data matrix, and as such the principal components form an orthogonal set. To understand the properties of the cointegrating vectors produced by the Johansen procedure, one needs to dig a little deeper into how the theory works. Recall from Chapter 6 that the Johansen procedure calculates the cointegrating vectors through solving a generalised eigenvalue problem of the form;

$$(\lambda_i[N] - [M])\{v_i\} = 0 \tag{8.1}$$

where $\{v_i\}$ is an eigenvector with corresponding eigenvalue λ_i, and $[N]$ and $[M]$ are symmetric positive definite matrices. When applying the Johansen procedure, $[N]$ and $[M]$ are generated from the input data and the desired cointegrating vectors correspond to the eigenvectors $\{v_i\}$. The properties of a generalised eigenvalue problem dictate that, upon solving (8.1), the resulting eigenvectors (and therefore cointegrating vectors) have an orthogonality property dictated by $[N]$, which is that $\{v_j'\}[N]\{v_i\} = 1$ if $i = j$ and 0 otherwise [91].

In short, the orthogonality properties of principal components and the cointegrating vectors differ (unless the matrix $[N]$ is an identity matrix). This means that one can expect to see different results from each methodology, even though the goals of each could be viewed as being similar. To examine this, the current section provides a short comparison between results from PCA and cointegration analysis on the DAMASCOS data. How similar results from the two methodologies actually are, and which is more appropriate for the application will be explored.

Within this comparison, the issue of standardising data prior to the application of algorithms such as PCA and cointegration must be discussed. In the previous section, following [63], principal component analysis carried out for the projection of data onto the minor components was applied without first standardising the data. Although very good results have been produced, it is nowadays common practice to standardise data before attempting PCA, so as to not form principal components biased by the size of the variables under consideration. For a complete comparison, in the following, PCA on both non-standardised and standardised data will be investigated alongside the results from applying cointegration. Using cointegration on non-standardised data is not attempted as the Johansen procedure can easily become ill conditioned if variables of very different amplitudes are used.

In the following comparison of results, the same training period is utilised throughout which consists of the first 2000 sample points, this training set, therefore, covers the whole period of stationary temperature and just under two cycles of temperature fluctuation. While in the preceding sections of this work it has been common to utilise a training set made up of every other sample from the data when applying PCA and outlier analysis, this approach is less suitable where cointegration is concerned. In the Johansen procedure, the choice of the cointegrating vectors is in-

formed by the fitting of a vector error correction model, which is of similar construct to an AR model. It has been found that suboptimal cointegrating vectors are chosen by the Johansen procedure when non-consecutive samples are used for training.

The first comparison that will be made for the DAMASCOS data is between the 50th principal component score (with and without standardisation of data) and the cointegrated residual from the first cointegrating vector (most stationary). For the training period described above, the 50th PC score, without prior standardisation of data is plotted in Figure 8.11, the 50th PC score with prior standardisation of data is plotted in Figure 8.12, finally, the cointegrated residual is shown in Figure 8.13. One could expect that the 50th principal component score, which accounts for the smallest proportion of variance in the data would be similar to the cointegrated residual which is created using the 'most stationary' cointegrating vector.

Further comparison can be made by looking at an expanded number of principal components and cointegrating vectors. Below, multivariate outlier analyses on the first ten cointegrated residuals and the last ten principal component scores, trained on the same training data, will be plotted. Figure 8.14 shows the results of a multivariate outlier analysis on the last ten principal components from the PCA carried out on the non-standardised data, Figure 8.15 shows the same with the PCA applied to standardised data, lastly, Figure 8.16 shows the results of an outlier analysis on the residuals created from the first ten cointegrating vectors.

From Figures 8.11, 8.12 and 8.13, one can see immediately that all three approaches have produced different results. Notably, standardising data prior to applying PCA has produced significantly different results to those where standardising has not been used. Where data has not been standardised, Figure 8.11, the score appears to be Gaussian before the introduction of damage, upon which the error bars of the control chart are exceeded. The 50th principal component score from the standardised data, Figure 8.12, also clearly indicates the introduction of damage, on close inspection, however, the score exceeds the control chart limits during the temperature fluctuation period before damage is introduced. Lastly, the cointegrated residual in Figure 8.13 remains within the control chart limits for the duration of the test until the introduction of damage (with one exception), where it clearly becomes nonstationary. As found previously, the cointegrated residual spikes at a time when the temperature regime was changed from stationary to cyclic.

Studying Figures 8.14, 8.15 and 8.16, one can see again that the three approaches

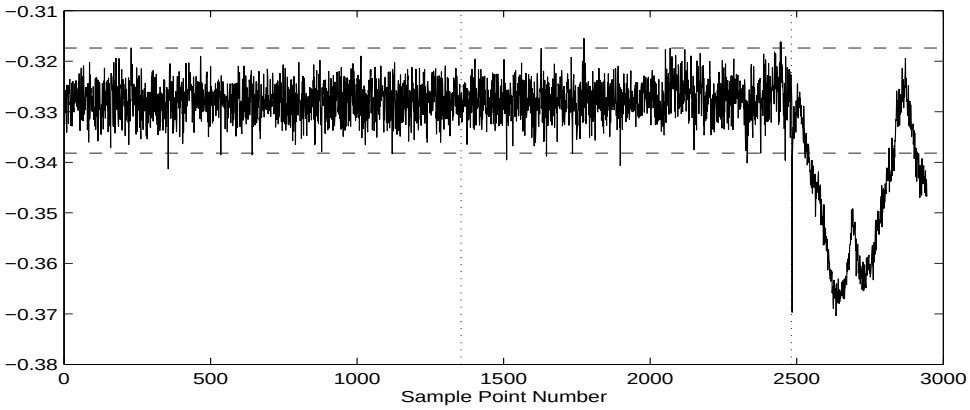


Figure 8.11: 50th Principal Component Score, PCA applied to non-standardised data

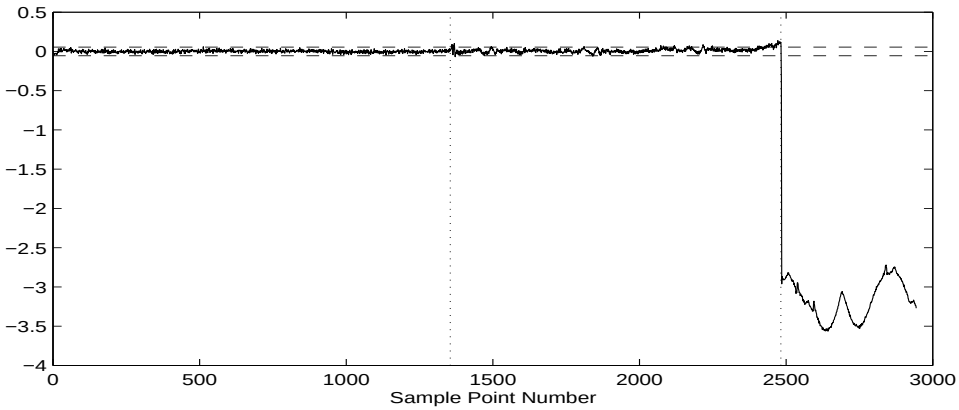


Figure 8.12: 50th Principal Component Score, PCA applied to standardised data

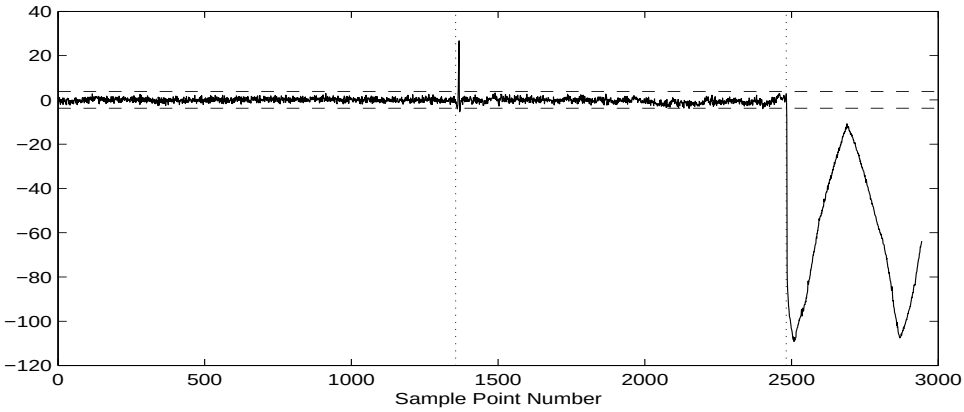


Figure 8.13: Cointegrated residual (corresponding to first cointegrating vector)

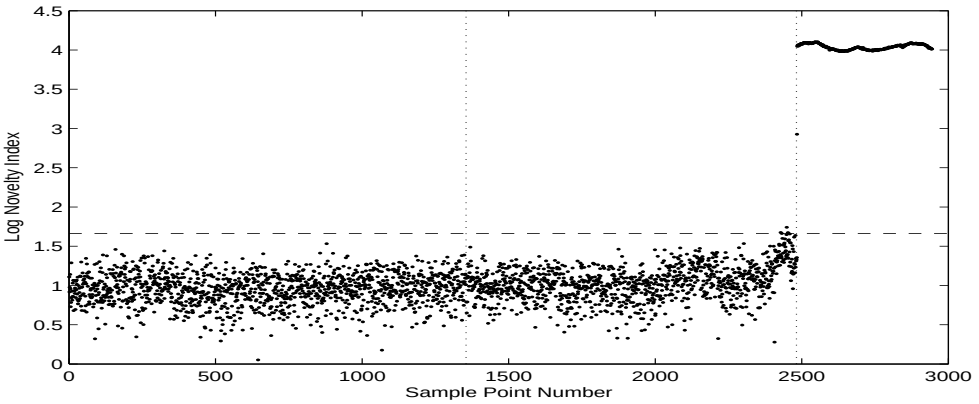


Figure 8.14: Multivariate outlier analyses of the last ten principal component scores, PCA applied to non-standardised data

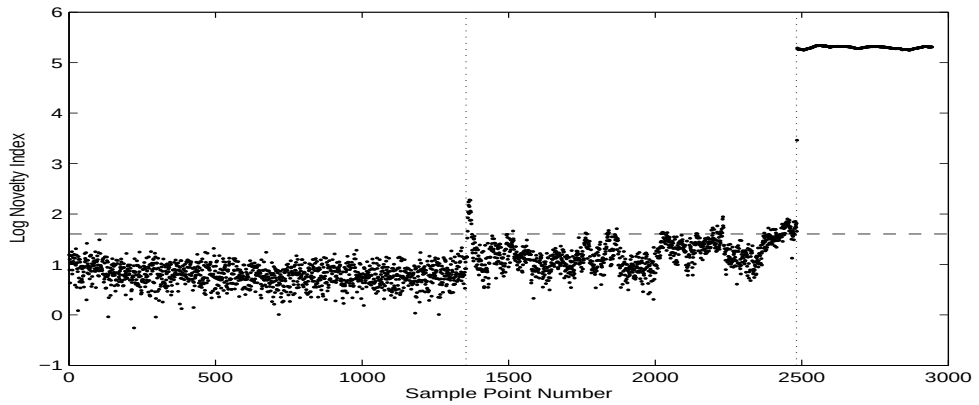


Figure 8.15: Multivariate outlier analyses of the last ten principal component scores, PCA applied to standardised data

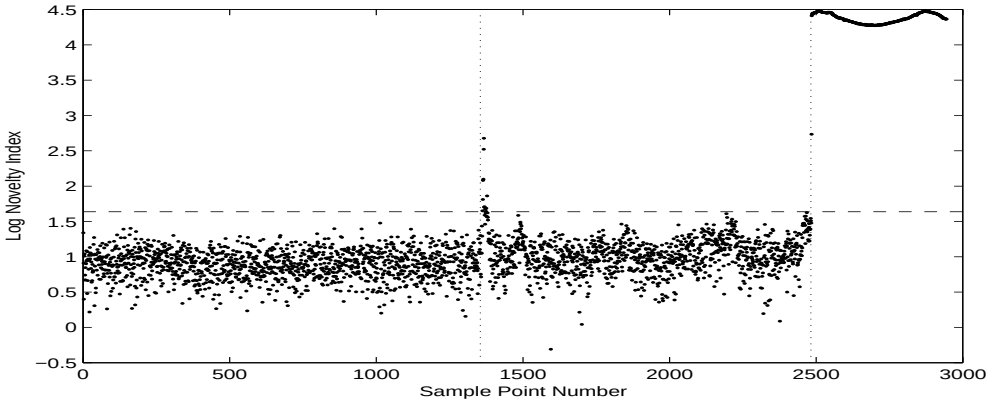


Figure 8.16: Multivariate outlier analyses of the first ten cointegrating vectors

have produced different results. All three plots show a clear detection of the damage introduced to the plate. The results from the PCA on standardised data, Figure 8.15, appear to be the least successful, as remaining structure from the temperature cycling period is still visible and the control chart limits are exceeded a number of times.

From the comparisons made above three direct observations are; firstly, that the methodologies are producing differing results, secondly that it is clear that the creation of features through projection onto minor components is more successful when the data is not standardised before PCA is applied, and, lastly, that a spike occurring at the time that the temperature cycling begins is visible in the cointegrated residuals but not in the non-standardised PCA results.

To firstly address the occurrence of the spike visible in Figure 8.13 (and indeed in Figure 8.8); it is interesting to note that upon inspection of the individual residuals created from the first ten cointegrating vectors in the above analysis, a number of them are free from the spike in question, and indeed may be more suitable for as damage sensitive features than the residual created by the first cointegrating vector. As an example the residual created from the second cointegrated vector is plotted in Figure 8.17. In this case, it seems that the 'most stationary' residual chosen by the Johansen procedure is not the most suitable for our cause. One should also recall that it was mentioned earlier that considering only the first 20 spectral lines of this 50 line set also produces a cointegrated residual from the first cointegrating vector that is free from the spike in question (Figure 8.10). It seems likely that the spike is an anomaly caused by one of the variables from around the peak area of the spectrum.

From the mathematical reasoning at the beginning of this section it is not unexpected that the results compared above are different for the two algorithms applied. Further insight into the applied algorithms can be gained by studying the specific linear combinations created by the two different approaches used to generate the results above, in doing this, light can also be shed on how standardising the data before applying PCA has an effect on the results. In Figure 8.18 the coefficients of the linear combinations that create the last ten principal components of PCA on the non-standardised data are plotted in a bar chart, those of the PCA applied to standardised data are plotted in Figure 8.19, similarly Figure 8.20 shows the coefficients of each of the linear combinations that make up the first ten cointegrating vectors.

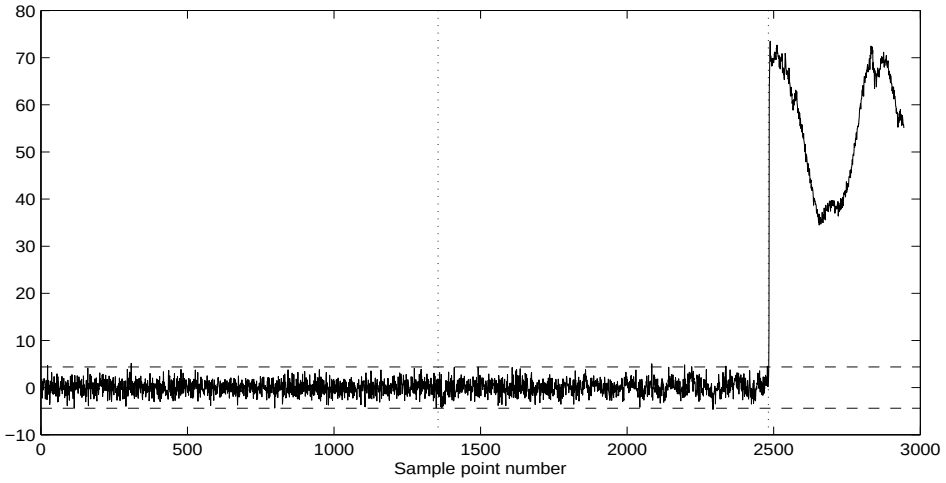


Figure 8.17: Cointegrated residual created from second cointegrating vector

Studying Figures 8.18, 8.19 and 8.20, one can see immediately that the constitution of the last ten principal components and the first ten cointegrating vectors are very different. The last ten principal components from the non-standardised data are dominated by contributions from those spectral lines away from the peak. This is easily explained; by not standardising the data when calculating the principal components, precedence has been given to the spectral lines displaying a larger response magnitude than others. Consequently the higher principal component scores will be dominated by the spectral lines from around the peak, and the lower ones the converse. If the data is standardised prior to the application of PCA, the higher principal component scores have equal contributions from each of the variables used, meaning that the contributions to the lower components are not dictated by the original amplitude of the spectral lines. Unlike the non-standardised PCA, looking at Figure 8.20, one can see that the higher cointegrating vectors have stronger contributions from around the peak of the spectrum than from anywhere else. An explanation for this could be that the spectral lines away from the peak, that vary less, contribute less to the nonstationarity of the linear combination and as such are assigned less dominant coefficients.

From these observations, the reason that the PCA on the standardised data does not perform as well as for non-standardised data becomes clearer. By not standardising the data the minor component scores are dominated by spectral lines not in the peak area, these variables show lower sensitivity to temperature, and as such the temperature trend has been more easily dispersed. Where standardisation has been

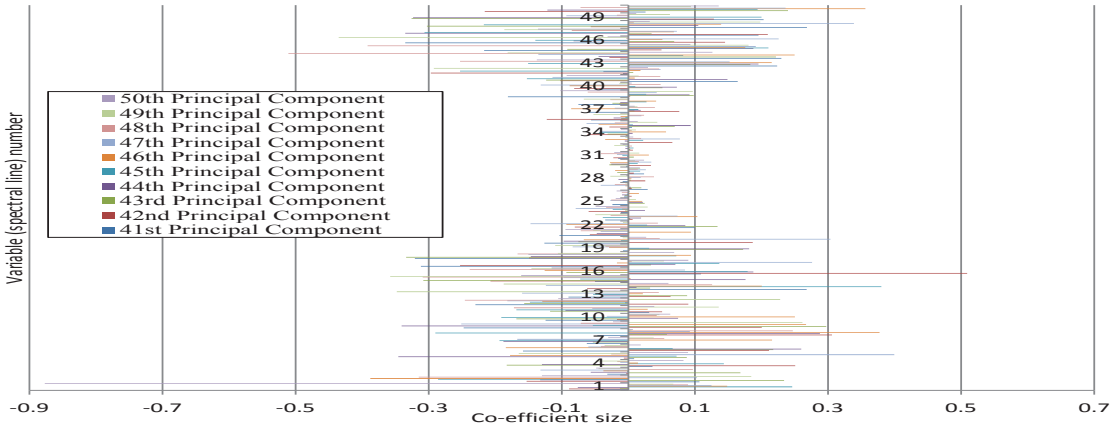


Figure 8.18: Coefficients of last ten principal components, nonstandardised data

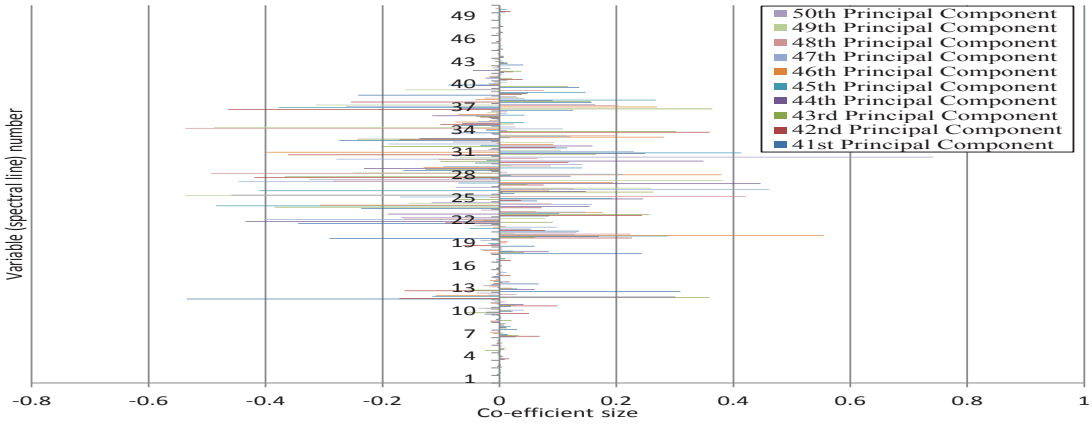


Figure 8.19: Coefficients of last ten principal components, standardised data

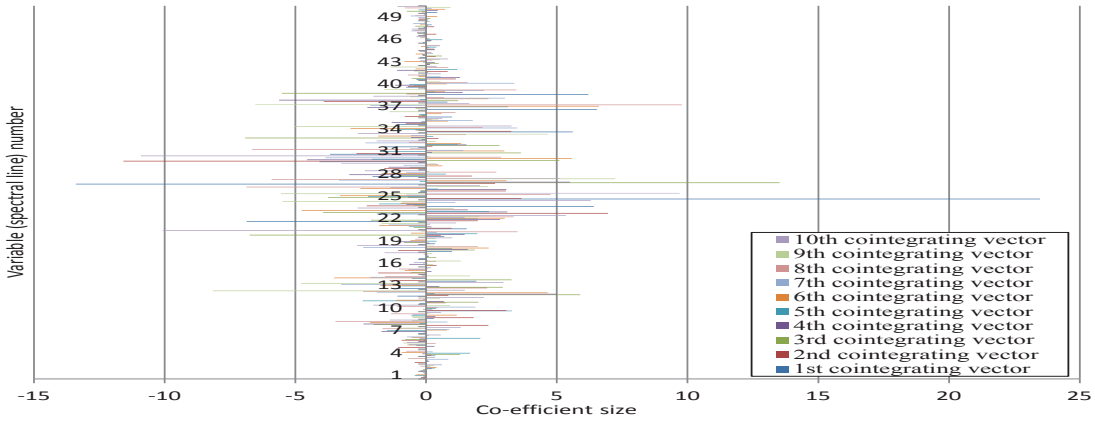


Figure 8.20: Coefficients of first ten cointegrating vectors

used this is not the case, instead each of the principal components is dictated by the direction of the most variance in the data set which is not longer biased by the amplitude of the spectral lines around the peak. This has, in fact, been detrimental to the performance of the minor components for the purposes of this work. That having been said, it could also be argued that the minor components of the non-standardised data are less satisfactory candidates for damage sensitive features due to the fact that the spectral lines around the peak, that are likely to display the greatest sensitivity to damage, have been assigned very low importance in the linear combinations. Here one can see an advantage to cointegration, where importance is assigned to the peak spectral lines.

8.4 Conclusions

This chapter has applied the theory of cointegration introduced in this thesis to a benchmark study that used Lamb-wave propagation to detect damage in a composite plate. The results when using cointegration have been compared with Manson's approach to using PCA for the same task [63]. This chapter also has expanded Manson's work as presented in [63].

Both PCA and cointegration have produced very encouraging results when applied to the benchmark study. Both methods were able to create features that remained unchanged by temperature fluctuations but still were able to very clearly detect damage.

In the final section above, some comparisons were made between PCA and cointegration, which on the surface of things are similar methods, both creating linear combinations of original variables. It was found that cointegrating vectors and principal components are not necessarily similar, they are chosen on different criteria and have different orthogonality properties. On application to the DAMASCOS data investigated in this work, both approaches were successful for removing a temperature induced trend. Interestingly, however, it was found that the linear combinations of the minor principal components relied on variables (spectral lines) from different areas of the spectrum than those in the cointegrating linear combinations.

While in this work both methods performed well, the author believes that cointegration may prove more useful for the data normalisation problem. As principal

components are always orthogonal, after the first PC is chosen to account for the most variance in a data set, the directions of the remaining principal components are then constrained by this orthogonality condition. As such, the minor components may not provide optimal results for removing environmental trends. It is here that cointegration may have the advantage due to the fact that the first cointegrating vector is chosen to be the most stationary, and is not dictated by any other constraints.

Chapter 9

An exploration of nonlinear cointegration for Structural Health Monitoring

9.1 Introduction

The idea of using cointegration for SHM works well where responses are linearly related, however when nonlinearities are involved the cointegration theory applied in the previous chapters is no longer suitable. Instead a nonlinear approach to cointegration is needed, where a *nonlinear* combination of response variables is used to remove unwanted environmental and operational trends. In cases where the classical linear cointegration theory is applicable, SHM practitioners are able to draw on the large body of research on cointegration carried out in the field of econometrics (see for example [88, 89, 91, 94, 95, 99]), however, this is not the case as far as nonlinear cointegration is concerned. The idea of nonlinear cointegration has received considerably less attention from econometricians in the past than the equivalent linear theory, most probably due to the fact that an extension to include nonlinearities is often deemed unnecessary due to the nature of the economic variables under consideration. Interest, however, is currently growing and progress from an econometric point of view is well summarised in [105]. For monitored variables in SHM, it is not at all unlikely that a nonlinear approach to cointegration will be necessary, in fact, the motivation of this work comes from the analysis of data from the Z24 bridge

monitoring campaign [83] where nonlinear cointegration could prove very useful. After introducing the Z24 example as motivation, this chapter makes a start into exploration of nonlinear cointegration.

9.2 Nonlinear cointegration

Recall that cointegration is a property of nonstationary time series integrated of the same order that share common trends; more precisely, a number of nonstationary variables (integrated of the same order) are cointegrated if a linear combination of them can be found that is integrated to a lower order than the original variables. Intuitively, nonlinear cointegration relates to a set of nonstationary variables that require a nonlinear combination to reduce the nonstationarity of the resultant. In the following, a set of nonstationary variables y_i are nonlinearly cointegrated if z_i is integrated to a lower order than y_i (or ideally stationary), where f represents a nonlinear function.

$$z_i = f(\{y_i\}) \tag{9.1}$$

In this case $f(-)$ will be referred to as the *cointegrating function*.

When reviewing progress of research into nonlinear cointegration, it is useful to bear in mind that for SHM applications, the aims in using nonlinear cointegration are probably very different to those of the econometrician. The approach in this thesis seeks to exploit the cointegration property of variables and create a stationary combination of them purged of environmentally or operationally induced (nonlinear) trends, and then use this stationary residual as a diagnostic tool to determine whether a structure continues to respond in a normal way. In the field of econometrics, the behaviour and influences to a process/variable are more uncertain, and in general the aim of studying cointegration is to establish if potentially spuriously related variables are truly related, and to predict how they will move together in the long run even after shocks and unforeseen events. Really the concern is to very accurately model the structure of these stochastic variables in relation to each other. For these reasons, a lot of attention is given to the nature of the variables under consideration, and statistical tests are needed to determine whether a variable should be modelled as linear and stationary, linear and nonstationary, nonlinear

and stationary, or finally nonlinear and nonstationary [105]. Nonlinear cointegration analysis is necessary for the nonlinear and nonstationary variables (although nonlinear and stationary variables could also be included in this remit, these variables are normally addressed through nonlinear cotrending analysis, which is a subset of nonlinear cointegration [106]).

In econometrics, the property of nonlinear cointegration is often discussed and expressed through the ideas of stability, attractors and mixing; the property of cointegration generalises to whether or not variables have attractors with similar topological properties (these include entropies, Lyapunov exponents and topological dimension), or to whether the nonlinear combination of variables is mixing or not. The concept of mixing is quite an abstract one, but for the purposes of SHM, it can simply be regarded as a somewhat stronger condition than stationarity i.e. mixing in a time series implies ergodicity of that series, which in turn implies stationarity (see [105] for more details). For the purposes of this work a definition of nonlinear cointegration that will suffice applies to nonstationary variables that become stationary after some form of nonlinear combination, and it is finding this nonlinear combination that is most likely of interest to SHM practitioners more than anything else. The majority of this chapter will address the author's suggested methods for finding the most stationary nonlinear combination of a set of nonstationary nonlinear variables. It should be noted that very little attention is given to this issue in the econometrics literature, often the form of nonlinear function needed is presumed to be already known, or in other circumstances a neural network has been used as a nonparametric estimation approach [105]. In the next section, the well known Z24 highway bridge monitoring campaign is introduced as a motivation for the need of nonlinear cointegration.

9.3 Motivational example from the Z24 Bridge

The Z24, a pre-stressed concrete highway bridge in Switzerland, was subject to a comprehensive monitoring campaign under the 'SIMCES project' [107], prior to its demolishment in the late 1990s. It has since become a landmark benchmark study in SHM. The monitoring campaign, which spanned a whole year, tracked modal parameters and included extensive measurement of the environmental factors affecting the structure, such as air temperature, soil temperature, humidity etc. The

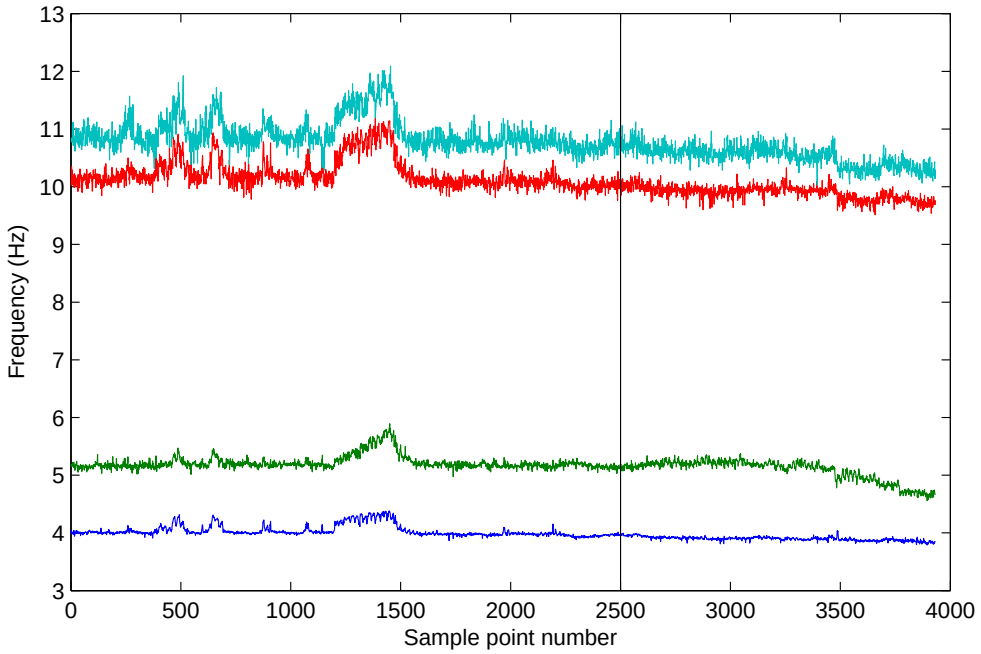


Figure 9.1: Time Histories of the extracted natural frequencies of the Z24 Bridge, monitored over one year including a period when damage was introduced

Z24 monitoring exercise was an important study in the history of SHM developments because towards the end of the monitoring campaign researchers were able to introduce a number of realistic damage scenarios to the structure. The progressive damaging scenarios initiated during this campaign, were, in order, [64]:

- Pier settlement
- Tilt of foundation followed by settlement removal
- Concrete spalling
- Landsliding
- Concrete hinge failure
- Anchor head failure
- Tendon rupture

Of interest here are the natural frequencies of the bridge which were tracked over the period of a year including the time where structural damage was introduced. Modal

properties of the bridge were extracted from acceleration data [83]. Figure 9.1 shows a time history of the four natural frequencies between 0-12Hz of the bridge. The solid vertical line marks the beginning of the period where the different damage scenarios were applied. Gaps where the monitoring system failed have been removed.

On inspection of Figure 9.1, the natural frequencies of the bridge are by no means stationary. There are some large fluctuations in the first half of the time history before the introduction of any damage. These fluctuations occurred during periods of very cold temperatures and have been associated with an increase in stiffness caused by freezing of the asphalt layer on the bridge deck. The natural frequency time histories are, therefore, a good illustrative example of damage sensitive parameters also sensitive to environmental variations, in this case temperature.

As the natural frequencies in their current form would not be suitable to monitor as a damage sensitive feature some action must be taken to remove the variable set's sensitivity to temperature. If each variable (natural frequency) is linearly related to temperature, cointegration would appear to be an ideal tool to remove the temperature induced trends. In this case, however, the modal properties of the bridge are nonlinearly dependent on temperature (as an example, Figure 9.2 plots how the first natural frequency changes with temperature), which means that the Z24 provides an excellent example with which to explore the ideas of cointegration in the presence of nonlinearity.

The first sensible step when exploring the ideas of cointegration in this nonlinear context is to look at the results of using the linear Johansen procedure. Figure 9.3 shows the four natural frequencies of the Z24 bridge projected onto the 'best' cointegrating vector found by the Johansen procedure, when trained on data points 1-500 visible in Figure 9.1. In Figure 9.3, the dotted horizontal lines indicate confidence intervals at 3σ of the residual from the training period, while the vertical solid line indicates the beginning of the period where damaging scenarios were introduced. Studying Figure 9.3, the cointegrating vector has successfully de-trended the data set and furthermore some indication of damage is visible towards the end of the data. It is perhaps unexpected that the linear Johansen procedure is able to remove nonlinear trends; looking at the relationships between the natural frequencies and how the Johansen procedure has combined them, however, sheds light on how this has been achieved. Although each frequency is related nonlinearly to temperature, which drives the frequency fluctuations, some of the frequencies (although not all) are linearly related to each other, which means that the Johansen procedure can

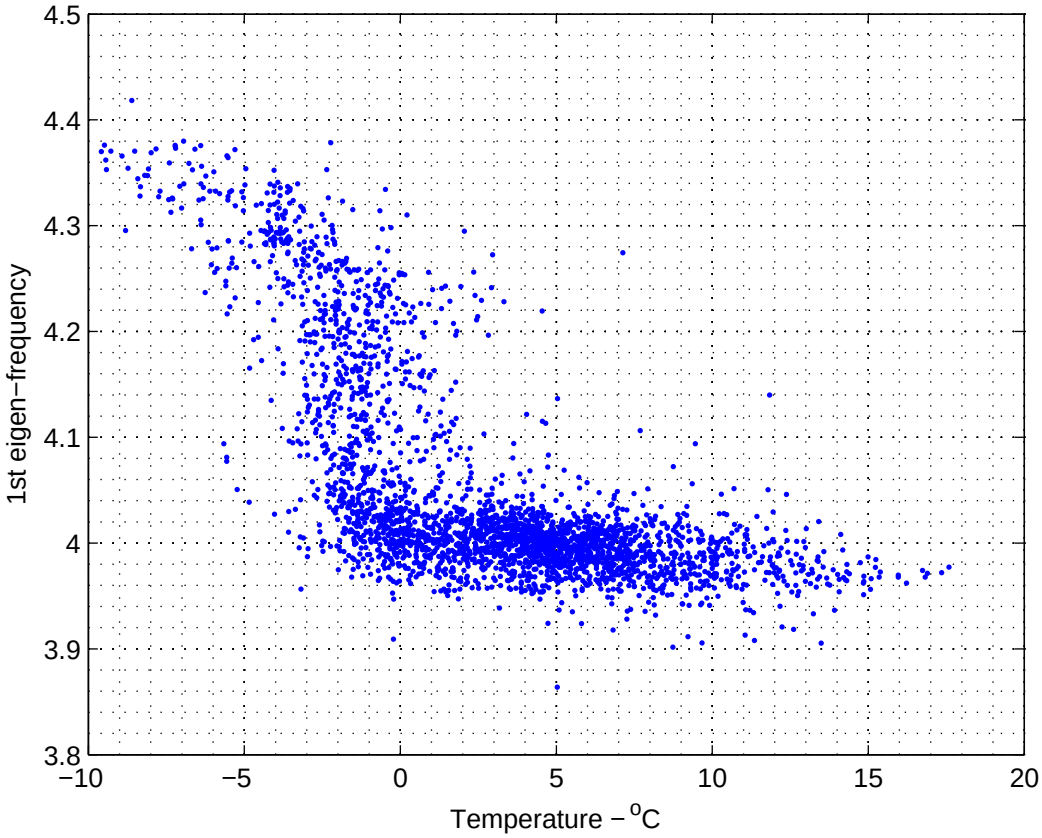


Figure 9.2: First natural frequency nonlinearly dependent on temperature

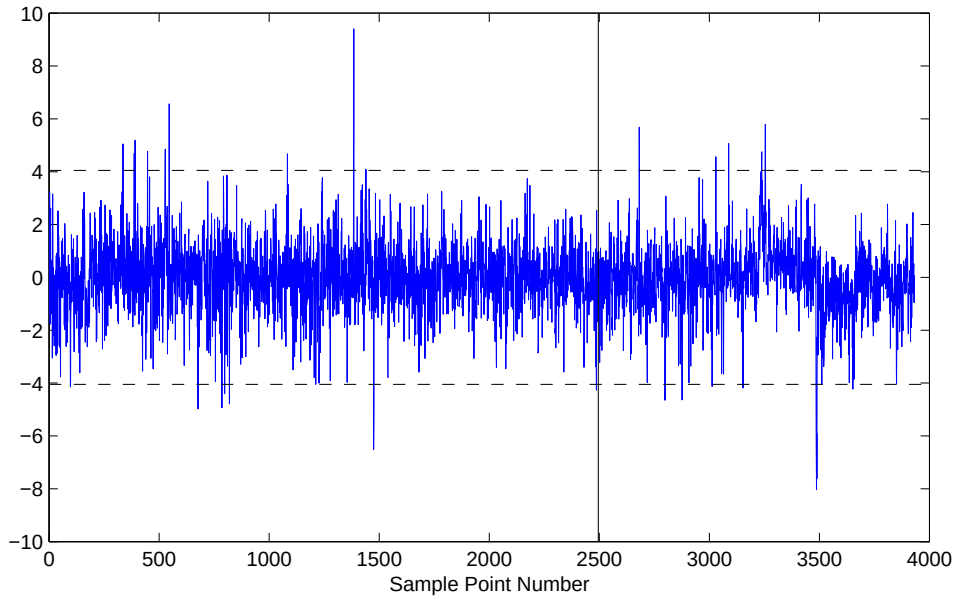


Figure 9.3: Linear combination of the Z24 data set; data projected onto the first cointegrating vector found by the Johansen procedure

successfully combine them to remove their common trends. The way in which the natural frequencies from the Z24 relate to each other is noted in Table 9.1. On studying the cointegrating vector in question, the residual in Figure 9.3 predominantly results from a combination of the first and third frequencies, which effectively removes the temperature dependent trends, the combination only contains very small contributions from the second and fourth frequencies. Although this is a successful removal of the temperature dependent trends, two of the variables have effectively not been included in the analysis. This is not an ideal situation as the loss of two variables reduces the chances of being able to successfully detect damage.

This point is well illustrated by studying the residual when the data are projected onto the second best cointegrating vector from the Johansen procedure, as illustrated in Figure 9.4. This combination does not penalise any of the variables and cannot therefore remove the large blip occurring when very low temperatures were recorded that induced the nonlinearity. However, the second cointegrating vector does show a much clearer indication of damage. To have the best of both worlds, it seems necessary to find a combination purged of temperature dependency that has a meaningful contribution from each frequency. For this, nonlinear cointegration is necessary.

Frequency	**1**	**2**	**3**	**4**
1	-	Nonlinear	Linear	Linear
2	Nonlinear	-	Nonlinear	Nonlinear
3	Linear	Nonlinear	-	Linear
4	Linear	Nonlinear	Linear	-

Table 9.1: Relationships between modal frequencies of the Z24 Bridge

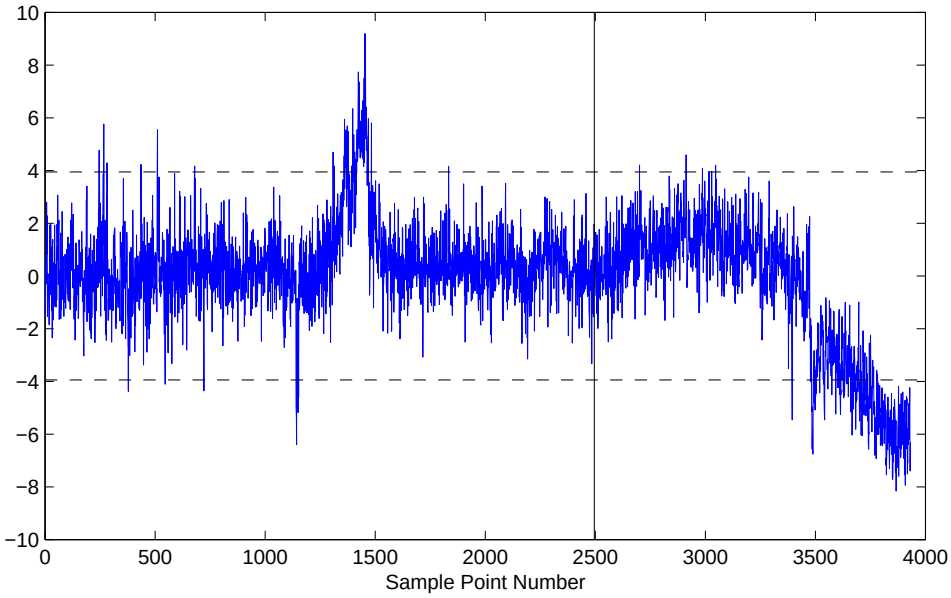


Figure 9.4: Linear combination of the Z24 data set; data projected onto the second cointegrating vector found by the Johansen procedure

This having been said, without using nonlinear cointegration, one solution in this circumstance could be to use 'locally linear models'. Recall that for the data from the Z24, when temperatures are above zero degrees, each variable is linearly related to temperature. By only looking at data from the locality of the linear regime, cointegration principles once again become valid. To illustrate this, the Johansen procedure was re-implemented using the four variables, this time with all data points removed if they occurred at temperatures below $1^{\circ}C$. Figure 9.5 shows the residuals of the first two cointegrating vectors, the first of which, similarly to before, has lost much of its sensitivity to damage. The second cointegrating vector has been much more successful, however, in that the effects of temperature have been removed and yet the residual still becomes clearly nonstationary after the introduction of damage. This would now be a good candidate feature for damage detection. Of course, this damage indicator could only ever be used to infer structural condition at times above freezing temperatures. Once again, nonlinear cointegration would provide a more ideal solution.

9.4 A simple approach to nonlinear cointegration

From the above it is clear that in some circumstances cointegration is limited by its linear nature. In this chapter a simple approach to how nonlinear cointegration may be achieved is explored through use of a simple synthetic scenario. One situation that may commonly arise is where two (or more) different variables from the same system exhibit nonlinear dependencies on some external disturbance, such as temperature fluctuation, for example. To demonstrate this idea theoretically, suppose there are two different variables x_i, y_i from the same system, one which reacts linearly with respect to some external disturbance, t, and one which reacts nonlinearly, in a quadratic way, say, to that same external disturbance. Suppose these variables take the form

$$x_i = \alpha t_i + \varepsilon_i \tag{9.2}$$

$$y_i = \beta {t_i}^2 + \epsilon_i \tag{9.3}$$

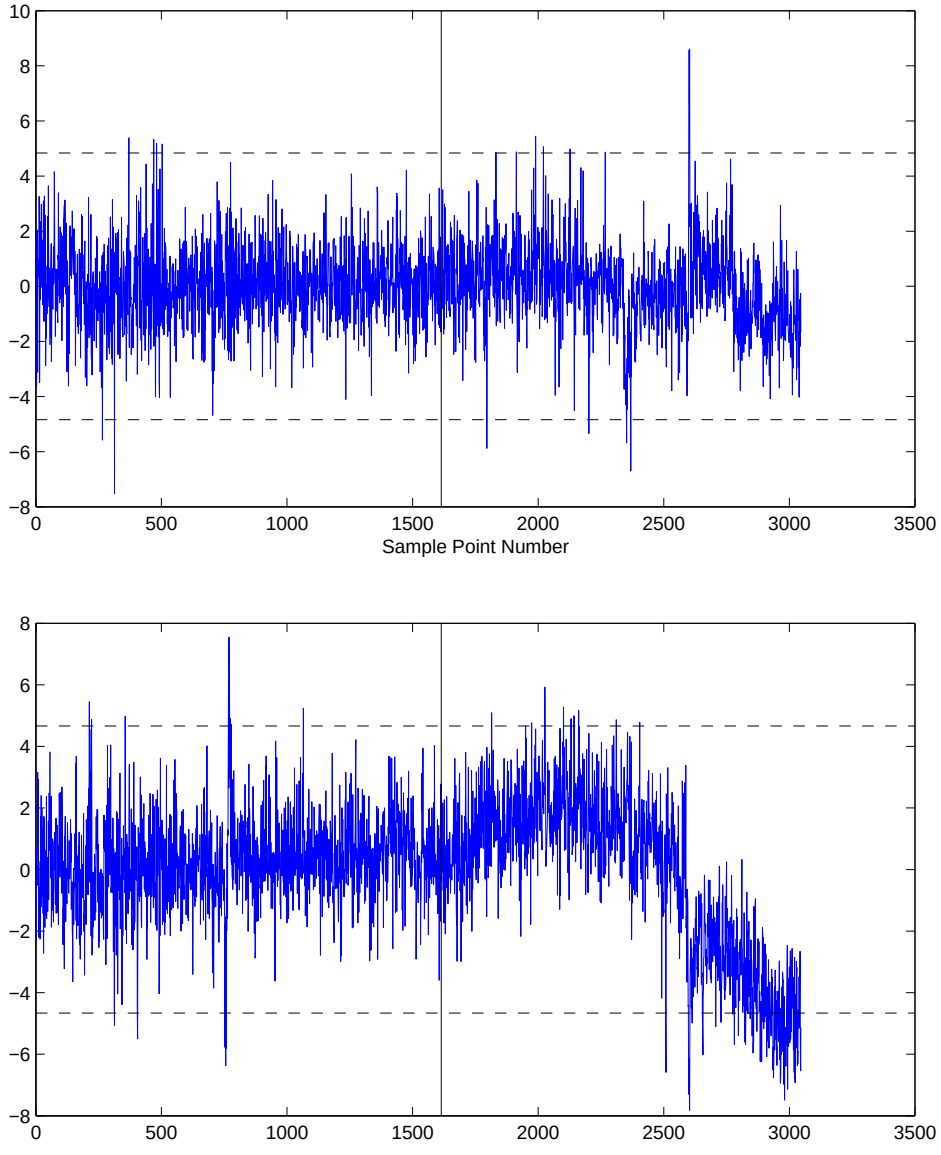


Figure 9.5: First and second cointegrating vectors using locally linear models

where α, β are constants, t_i is some deterministic trend caused by the external disturbance and ε_i and ϵ_i are random normally-distributed processes. It is clear that a linear combination of x and y could not result in a stationary sequence. However, some combination of y and the square of x should produce a comparatively stationary signal:

$$z_i = a_1 {x_i}^2 + a_2 y_i \tag{9.4}$$

If the parameters a_1 and a_2 can be found so that z_i is stationary, the vector $[a_1 a_2]$ will be analogous to the linear cointegrating vector, and x and y will be nonlinearly cointegrated. It is then a matter of finding the parameters a_1 and a_2.

One way to do this would be to compute ${x_i}^2$ and to include it as variable in its own right in the Johansen procedure, i.e. the input to the Johansen procedure would be $\{{x_i}^2, y_i\}$. In this way, the only difference in the approach to the Johansen procedure is a manipulation of the form of the variables that are linearly combined.

A different approach again is to treat (9.4) in terms of an optimisation problem, where the aim is to choose parameters $[a_1 a_2]$ such that z_i is as stationary as possible. For this purpose, a nonlinear optimisation routine based on differential evolution will be utilised here. The following subsection will briefly describe differential evolution but, as these are not new constructs, readers are referred to [108] and [109] for more details. A section will follow with results using the techniques on data simulated to represent the theoretical situation above.

9.4.1 Differential Evolution

Differential evolution, first introduced by Storn and Price in 1997 [109], is an evolutionary algorithm that begins with an initial population of trial solutions to some problem and reaches an optimal set of solutions through successive cycles of mutation, crossover and selection. The suitability of trial solutions are determined by some objective function, set according to the individual problem in hand. For this application, the trial solutions take the form of a vector of parameter guesses $[a_1 a_2]$ that satisfy (9.4) with z_i stationary.

The optimisation routine is summarised in Figure 9.6. To begin with an initial

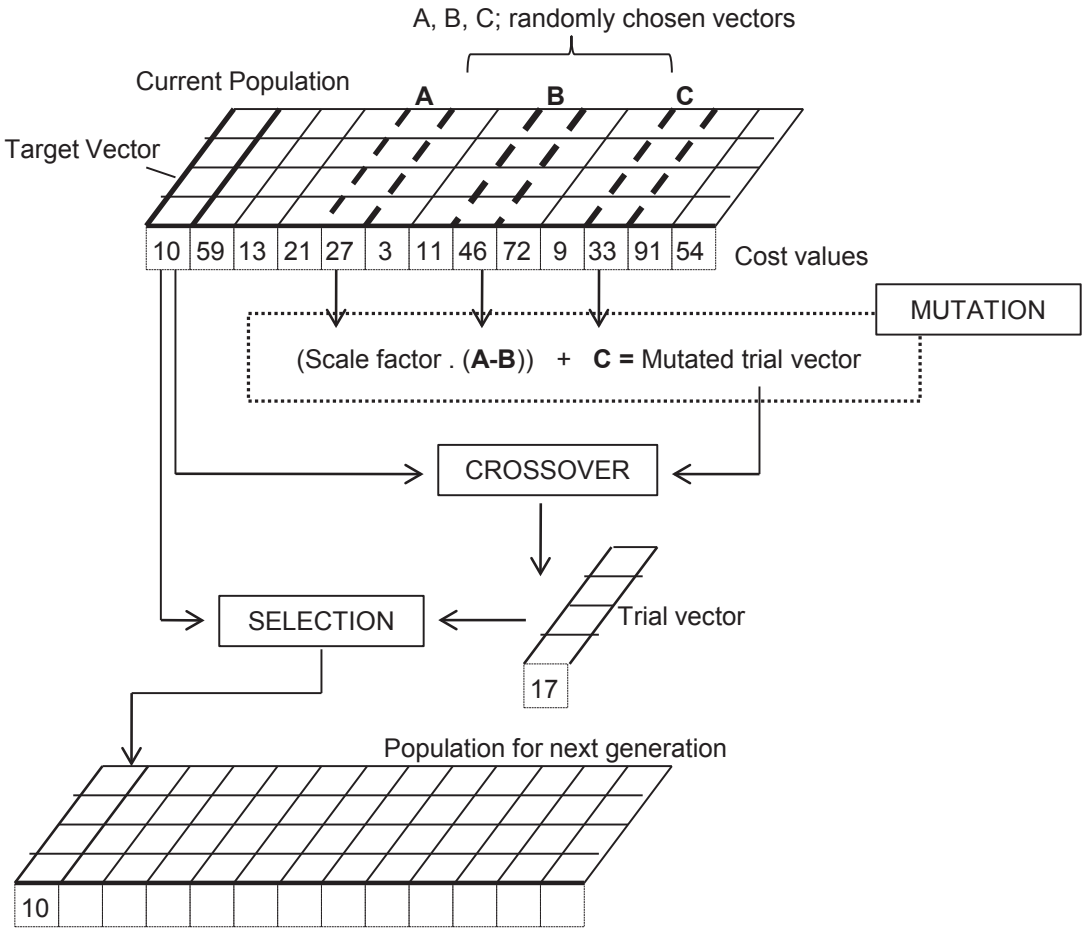


Figure 9.6: A schematic of differential evolution

population of parameter vectors are randomly generated. To each parameter vector in the initial population, a cost value is specified according to the objective function chosen. A new generation of solutions is created from this initial population as follows. Firstly, a target vector is chosen from the initial population. Next, a trial vector is created by 'mutation'; from the initial population, two parameter vectors are randomly chosen (A and B in Figure 9.6), their difference (A-B) is multiplied by some scaling factor, to which finally a third randomly chosen parameter vector (C) from the initial population is added. The resultant is called the *mutated trial vector*.

A new parameter vector is now created through 'crossover' of the mutated trial vector and the target vector. Crossover creates a new vector by choosing individual elements from the mutated trial vector and the target vector by a series of binomial

experiments (see [109] for details). This newly created vector will then be selected for the next generation if its cost value is lower than that of the target vector. If it is higher, the target vector will be placed in the next generation population. The process is repeated for each vector in the initial population. As the process evolves through the generations the population will eventually become full of suitable parameter vectors with low cost values.

For the purposes of nonlinear cointegration, a suitable objective function must be chosen on the basis of the stationarity of the cointegrated signal (such as (9.4)). Several options are available, the simplest being to choose the objective function to minimise the variance of the cointegrated signal. Another suitable option would be to use the ADF statistic from econometrics, introduced earlier in the thesis, which has a larger negative value the more stationary the time series is. Both cost functions will be trialled for the theoretical example above, and the results will be compared in the next section along with the results of using an altered basis for the Johansen procedure.

9.4.2 Results using Differential Evolution for synthetic example

The nonlinear cointegration procedures suggested here are trialled for combining time series of type (9.2) and (9.3). Namely, differential evolution using the two different cost functions discussed and a slightly modified version of the Johansen procedure are used to choose the parameters in (9.4) that produce the most stationary combination of (9.2) and (9.3). For simplicity and for ease of visualising the results, when simulating these two time series the driving trend t is, for the present, set to be linearly increasing with time. For engineering applications, this is highly unrealistic, as any driving trend, such as temperature will fluctuate, it also has the consequence that the simulated time series are trend stationary, rather than difference stationary for which the Johansen procedure has been developed for. At this initial stage, however, where the interest is in testing the concept and visualising the nonlinearity of the time series easily, it seems sensible to stick with a simplistic deterministic trend.

For the simulated time series, the results are shown in Figure 9.7. Figure 9.7(a) and 9.7(b) show the results using differential evolution with the variance-based cost

function and the ADF statistic-based cost function respectively. For this particular trial a scaling factor of 0.9 and a crossover ratio of 0.5 were used in the differential evolution step, for more details on the choice of such parameters readers, are referred to [108] and [109]. Figure 9.7(c) shows the results when using $\{{x_i}^2, y_i\}$ as a basis for the Johansen procedure.

On inspection of Figure 9.7, all three methods have found coefficients for the combination (9.4) that successfully remove the nonlinear trend; the three residuals are mean-stationary. In all cases as the nonstationarity of the residual is less that the original time series, the methods are providing successful nonlinear cointegrating vectors for the times series.

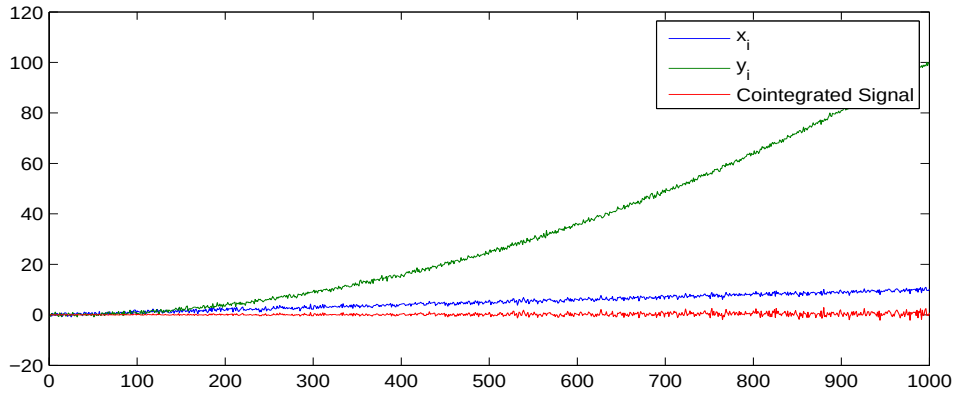
An interesting property and a possible drawback, however, of the kinds of combinations used to cointegrate these nonlinear trends 9.4, is that on closer inspection, the variance of each of the combined signals is increasing with time, although each cointegrated signal is mean-stationary. This growing variance is small in Figure 9.7(a) where variance was used a cost function, it grows at a faster rate, however, in Figures 9.7(b) and (c) where the ADF statistic and the Johansen procedure were used. To understand this one expands (9.4):

$$\begin{aligned} z_i &= a_1 {x_i}^2 + a_2 y_i \\ &= a_1(\alpha^2 {t_i}^2 + t_i\varepsilon_i + \varepsilon_i^2) + a_2(\beta t_i^2 + \varepsilon_i) \end{aligned} \tag{9.5}$$

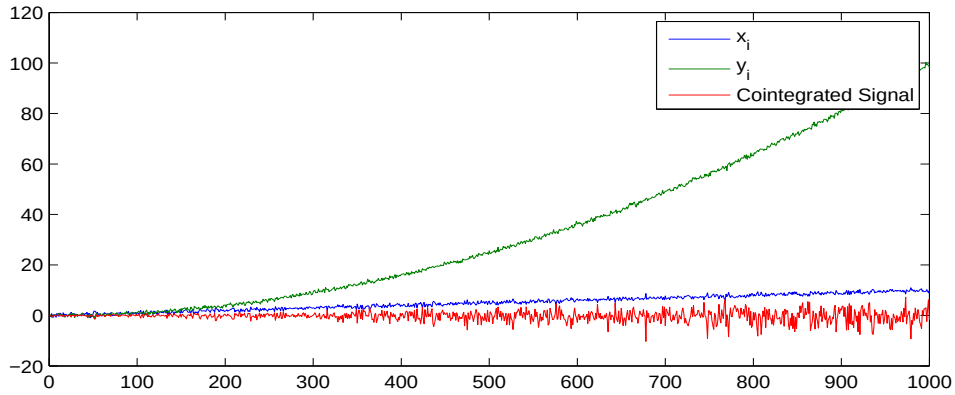
The differential evolution, and indeed the Johansen procedure, have chosen the parameters a_1 and a_2 so that the quadratic deterministic trends cancel each other out. The residual from the nonlinear combination in (9.5) will then take the form,

$$z_i = a_1 t_i \varepsilon_i + a_2 \epsilon_i + a_1 \varepsilon_i^2 \tag{9.6}$$

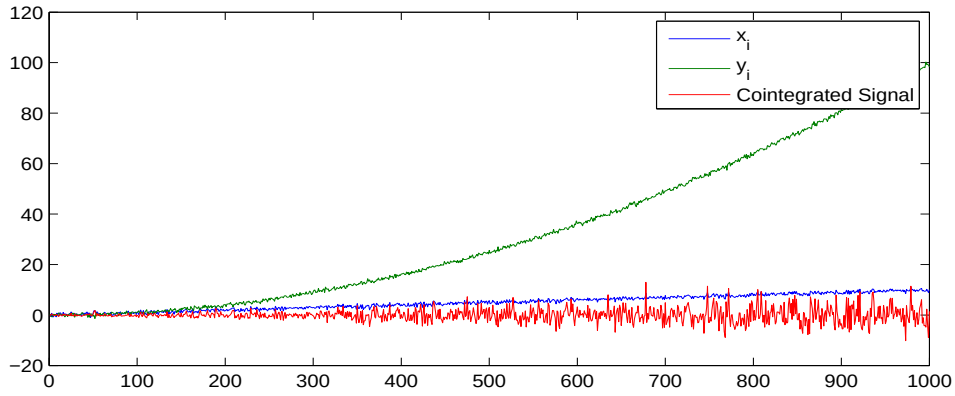
The remaining terms will include randomly distributed noise but also a term depending on $t_i\varepsilon_i$. This term is responsible for the increasing variance, as the driving trend of the simulated time series in this case was chosen to linearly increase with time. Although the residuals created shown in Figure 9.7 are 'more stationary' in comparison to the original signals (and, therefore, nonlinear cointegration holds), with the initial trends removed, they are not truly stationary on closer inspection.



(a)



(b)



(c)

Figure 9.7: Combination of signals with linear and quadratic deterministic trends, combination found using (a) differential evolution with a variance based cost function, (b) differential evolution with an ADF statistic-based cost function, (c) the Johansen procedure with modified basis.

Regardless of the nature of the driving trend t, a combination of the type (9.4), will always result in a residual that has variance dependent on that trend. If one's aim in applying the ideas of nonlinear cointegration is to remove a nonlinear trend, then the approach under discussion is fit for purpose. However, if stationarity of the residual is required (in the variance as well as the mean), then further steps must be taken after applying a combination such as (9.4). One potential means of arriving at a stationary residual could be to split the group of feature variables of interest into smaller subsets and find a suitable nonlinear combination for each. The residuals, with trend dependent variances, could then be linearly combined with each other using the Johansen procedure, in order to create a stationary residual independent of the driving trends of the original time series.

9.5 Nonlinear cointegration as combinatorial optimisation problem

In the previous section, a simple approach to nonlinear cointegration was discussed with use of a simulated data set. Inherent to this approach was knowledge of the correct structure of the nonlinear combination of variables that would reduce the nonstationarity of the residual. For real data, the underlying structure of each variable with respect to its driving trends will be unknown, and therefore too, the optimal form of combination such as (9.4).

To extend the simple example in the previous section of how nonlinear cointegration may be achieved to something that might work for real data, a sum of the form (9.4) could be extended, with all possible multinomials of the variates considered for inclusion. However, in situations with many features, such an approach is likely to run into difficulties due to the fact that the number of candidate terms grows explosively with the number of features and the allowed order of nonlinearity. An alternative approach is proposed herein whereby another optimisation procedure (Genetic Algorithm) is used to select an optimal subset of the candidate terms, with the parameters for combination then determined by exploiting linear cointegration theory. The assumption made is that, once the multinomial candidate terms for the combination are established, the Johansen procedure will be used in the same way as in the previous section to ascertain the correct coefficients. The application of the method will be again be demonstrated via simulated data.

9.5.1 Genetic Algorithms for candidate term selection

The Genetic Algorithm (GA) is the most fundamental of the evolutionary optimisation schemes based on the Darwinian principal of natural selection or 'survival of the fittest' [110]. In its simplest form, the algorithm uses a binary encoding to express possible solutions as individuals in a population which evolves in a manner analogous to natural selection. The GA is the algorithm of choice for the current work as it has proved extremely powerful for combinatorial optimisation problems and the subset selection problem posed here is one of this type. The main difference between the simple GA and the one adopted here is that the individuals for the problem are encoded as integer vectors rather than bit strings and this means that slightly modified versions of the genetic operators are needed. Extensions of the GA of this type are often called Evolutionary Programs [111].

The implementation of the GA/EP used here is conducted as follows:

- All variables to be included are standardised. This step is necessary as the nonlinear combinations of variables can differ considerably from each other in terms of scale and this can result in severe ill-conditioning for the coefficient estimation step.

- The initial population for the GA is generated. The user specifies the number of terms to be used in each sum N, and the highest nonlinear order allowed, n. A candidate term is generated by creating a random string of integers of the same length as the number of variables included in the analysis. This string is then used in the following way, suppose three variables are included in an analysis x_i, y_i, z_i, a random string (abc), where each bit is an integer from 0 to n, would be generated, and the candidate term chosen as $x^a y^b z^c$. This action is repeated N times and a candidate linear combination arises which is a single individual of the population. The whole process is repeated until the initial population is established.

- Evolution begins. In order to drive the process towards an optimum, the algorithm requires a cost or fitness function in order to evaluate how good a solution to the problem each individual represents. As the objective of the current problem is to generate the 'most stationary' combination of variables, the fitness function adopted here needs to be a measure of stationarity. For each individual in the population, coefficients for a linear combination are

calculated by the Johansen procedure and the (hopefully stationary) residual is computed. The cost or fitness of the individual is then expressed as either the variance of the residual, or its ADF statistic (user defined); both of these quantities are measures of stationarity.

- The operation of selection is performed whereby two parent individuals are selected from the population for mating. Roulette wheel selection is used in the implementation here. The individuals are more likely to be selected the fitter they are. Once two parents are established, two child individuals are generated by a crossover operation which mixes their genetic material. One-point crossover is used in the implementation here. The two children are carried forward into the next generation and the selection operation is repeated until a full population is established. The crossover operation applied here is carried out carefully to avoid introducing two identical candidate terms into the same sum.

- The mutation operation is applied to the child individuals. Each individual is mutated with probability p_m (user specified). If an individual is to be mutated, one of the integers which forms the genetic material is randomly picked and randomly regenerated. This operation encourages diversity and can prevent stagnation of the population; however, it should be used sparingly as too much mutation converts the algorithm into random search. Again mutation is monitored to prevent the generation of duplicate terms.

- To further reduce the chances of stagnation a 'new blood' stage is included at this time, where a defined number of 'un-fit' candidates are replaced new with randomly generated candidates for the next generation.

- In order to prevent the loss of the fittest solutions, an elite was used here whereby the N_e fittest individuals were written directly into the next generation (necessarily overwriting some of the children of the selection process).

- The whole process is iterated until a stopping criterion is met. The usual stopping criteria are that one reaches a pre-established number of generations or that the fitness function attains some desired value; in this case a pre-established number of generations was used.

Results using a genetic algorithm

The functionality of the genetic algorithm created for this work will be explored here using a simulated test case. Specifically data simulated to reproduce the theoretical example used above to explain the approach taken in the previous section is utilised. The input time series generated are 1000 data points long and the deterministic trend value t in (9.2) and (9.3) has a maximum value of 10.

The GA implementation is used to select the candidate terms for a linear combination that should result in the most stationary combination possible. The GA may choose candidate terms from any multinomial of the original variables x and y up to a given order. For the first trials, the mulitnomial order was limited to three, hence the GA could choose any candidate terms from the set $\{x^a y^b, a, b = 0, 1, ...3\}$, the number of candidate terms to be chosen for the final combination was set to two. The ADF test statistic was used as a cost function for the GA in this case.

Given these inputs, the genetic algorithm successfully chose the combination of form (9.4) from an initial randomly generated population of 30 possibilities. The GA was run ten times with a different initial population each time and the maximum number of generations was also set to ten. Figure 9.8 is a plot of the nonlinear combination chosen by the GA along with the original time series that make up the combination. As one can see the results are analogous to those in Figure 9.7(c).

During various trials of the GA, it was noted that on occasion the GA made unexpected choices for the nonlinear combination. One example of this is given below. The same structure of inputs as above were trialled but with the maximum value of t in (9.2) and (9.3) set to four. With this input the GA chose a combination, shown in Figure 9.9, of the form,

$$z_i = a_1 x_i y_i + a_2 {x_i}^3 \tag{9.7}$$

This is not the expected choice of nonlinear combination and the reason for it appears to be concerned with the cost function used in the algorithm. The cost of the chosen combination (i.e. the value of the ADF statistic) was -22.5174 (more stationary signals are indicated by larger negative values of the ADF statistic), the cost of the correct expected function as expressed by equation (9.4) was -22.1052, hence the combination of form (9.7) was chosen by the GA. If one expands (9.7) and assumes

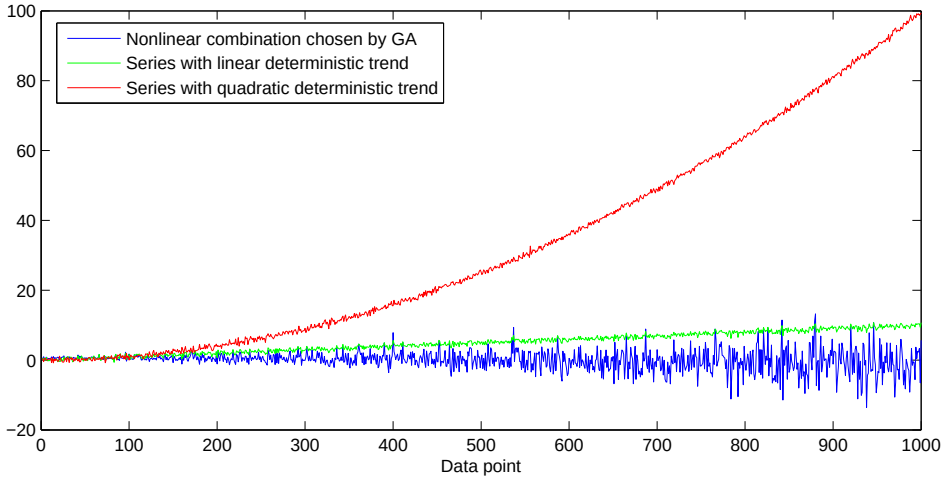


Figure 9.8: GA choice for nonlinear combination of linear and quadratic trends

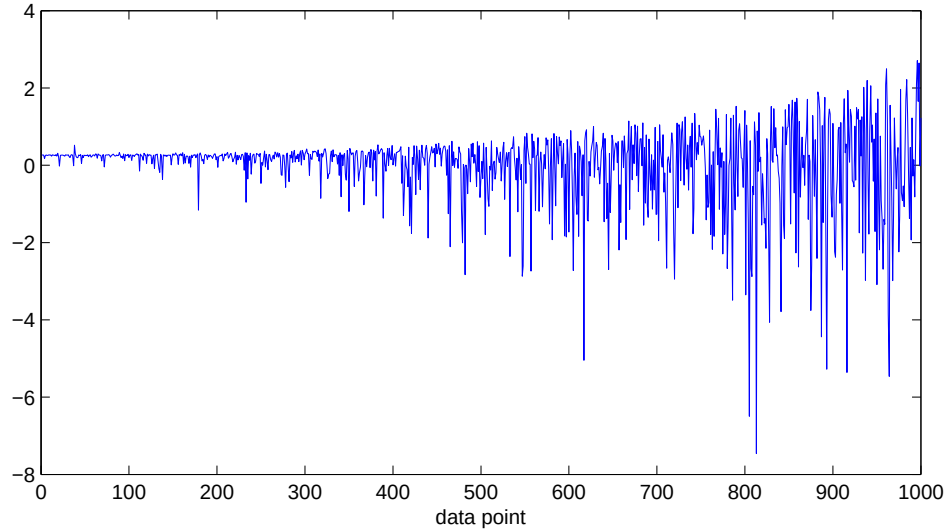


Figure 9.9: Unsuccessful GA choice for nonlinear combination of linear and quadratic trends

that the Johansen procedure eliminates the leading order trends, the residual will take the form,

$$z_i = (a_1\beta + 3a_2\alpha^2)t_i^2\varepsilon_i + (a_1\epsilon_i + 3a_2\varepsilon_i)\alpha t_i + (a_1\epsilon_i + a_2\varepsilon_i^2)\varepsilon_i. \tag{9.8}$$

Now, although this residual appears more complicated, it would appear that the combination in equation (9.7), in this instance, with these data, allows the Johansen procedure to select coefficients which make the residual in (9.8) a little more stationary that the one in (9.4). This is undesirable as the residual in (9.8) now retains a quadratic trend (visible in Figure 9.9) which will eventually dominate at later times even if the combination (9.7) appears fortuitous on the training set. A possible means of circumnavigating this problem is to introduce a penalty term in the optimisation which weights against higher order residual trends. The ADF statistic, for this type of simulation trial at least, will bear further investigation.

An alternative to the ADF statistic is, as discussed above, to use a cost function that is dictated by the variance of the candidate nonlinear combination, or potentially one which factors in both the variance and the ADF statistic. Unfortunately neither of these options are immediately viable due to the fact that the variance also introduces subtleties in the combination choice. This is due to the fact that the Johansen procedure, which is used to determine the coefficients of the nonlinear combination, often chooses large coefficients for the combination that would normally be the most successful, which drives the variance of the residual up. It is likely that the solution to this issue will involve a combination of adding penalty terms and regularisers.

The genetic algorithm discussed here is, perhaps, a cumbersome way to go about finding a suitable form of nonlinear combination for a set of variables. The implementation relies on using the Johansen procedure within the evolution, and even on simulated data, problems arise with its application, as evidenced by the example above. It is the opinion of the author that more suitable ways of finding useful nonlinear combinations of variables could be found.

Neural networks or Gaussian process regression could be suitable candidates for finding useful nonlinear combinations of variables. Indeed, some research has already been carried out on using GPs for such a task [62], although the concept of cointegration is not directly mentioned. Such methods are well worth looking into, and will be the topic of further work by the author. One potential drawback of

applying such methods, however, could be the fact that one variable out of a set would have to be chosen as a target to regress all other variables on; this is how tests for cointegration are achieved in the Engle-Granger two step method [95]. If a single variable must be chosen as a target, the damage detecting capabilities of the residual may be determined by which variable is chosen.

9.6 Conclusions

The ideas of nonlinear cointegration have been explored in this chapter. Nonlinear cointegration is necessary where feature variables are nonlinearly related. This chapter has begun to explore how one might attempt to nonlinearly combine feature variables in order to create useful diagnostic tools in the face of changing environmental and operational conditions. The ideas followed are based on multinomial combinations of feature variables, and optimal combinations are attempted using genetic algorithms. At this early stage of research the developed approaches are only applicable to data in simulation where the generating function is known. Other suggestions of using neural networks or Gaussian processes to form suitable nonlinear combinations have been made but not yet attempted. This will be the focus of future work.

Chapter 10

Summary and Conclusions

This thesis has focused on the issue of the confounding influence of changing environmental and operational conditions on technology developed for SHM. This issue, often referred to as the data normalisation problem, is widely considered as one of the largest stumbling blocks preventing the practical application of SHM to real world structures. The main case study of the thesis is the monitoring campaign of the Tamar Suspension Bridge in the UK, which is led by the Vibration Engineering Section at the University of Sheffield. Data collected from this structure has demonstrated that features that may be of interest for inference on structural condition or performance are highly influenced by the changing environment.

Chapter 3 introduced and described the substantial monitoring campaign being carried out on the Tamar Suspension Bridge. Three comprehensive monitoring systems currently in place have provided a wealth of data detailing the static and dynamic behaviour of the bridge deck and cables, as well as the operational and environmental factors affecting them. In Chapter 4 this monitoring data was analysed in order to understand the normal response of the bridge under the influence of the changing environmental and operational conditions. Various analysis techniques were used to better understand which environmental/operational conditions drive the fluctuations observed in the modal frequencies of the bridge deck. Traffic loading was found to be a dominant driver of daily frequency fluctuation, whilst temperature was found to have more of a seasonal effect than daily. The acceleration of the deck was also found to have a significant effect on the modal frequencies at times when the wind speed was higher than 25mph and hitting the bridge side-on. As an investigative

tool, simple response surface models were fitted in an attempt to predict the change in the lower modal frequencies of the bridge deck. The models including input parameters reliant on traffic loading, temperature and deck acceleration were able to predict the lower modal frequencies to a good degree of accuracy. A suggestion was made that the errors of the successful models could be incorporated into an SHM system for the bridge as an indicator of structural condition, this idea was further investigated in Chapter 5.

Chapter 5 aimed to build on the knowledge gained in Chapter 4 in order to make tentative steps towards the development of diagnostic tools for the Tamar bridge that would function in varying environmental and operational conditions. The use of novelty detection in such circumstances was investigated. Two different approaches for the implementation of a novelty detection scheme were discussed. The first approach, more readily applicable for features that do not exhibit seasonal behaviour, was to incorporate responses varying under a changing environment into the definition of the 'normal condition'. Specifically a novelty detector is trained on data including the fluctuating behaviour. Although this approach is very easy to implement given a reasonably large bank of historic data from an undamaged structure, it was concluded that the method may create features with low sensitivity to potential anomalies or structural change. The alternative approach suggested involved using the error of predictive models as a novelty indicator. The predictive models implemented in this chapter were the parametric response surface models used earlier when attempting to understand the normal condition of the bridge and models created using Gaussian process (GP) regression. Both the parametric response surface models and GP regression were found to be provide good candidates for generating accurate predictive models for deck displacement and the lower deck natural frequencies. A new suggestion was made that the prediction confidence intervals available when using GP regression can be used to directly detect novelty, in the place of a more classical control chart. GP regression naturally provides confidence intervals on predictions which widen if new circumstances not present in the training set occur, this enables the construction of a more conservative novelty detector than a traditional control chart may provide.

In Chapter 6, the idea of cointegration has been put forward as a new way to attempt to deal with the problem of environmentally-induced variation in measured structural response. The idea, which originates from econometrics, is to linearly combine response variables that are cointegrated to create a stationary residual

whose stationarity represents the structure's normal condition. When monitoring this residual, a departure from stationarity will indicate that the structure is no longer operating under its normal condition. Chapter 6 introduced the, sometimes complex, mathematics of the Johansen procedure which finds the most stationary linear combination of a set of variables under scrutiny. The ADF test has also been introduced as a stationarity test. Although not used for its intended purpose, cointegration has provided a useful tool for inclusion in an SHM analysis/system. Its advantages lay in the simplicity of the idea, the huge background of sophisticated research already carried out in the field of econometrics available for use, and the fact that in essence no information is being lost as features are created through combination of monitored variables. The suggested cointegration procedure can also be implemented where no measurement of the environment/operational conditions are available, the only stipulation being that the residual should be trained on data coming from the normal condition of a structure.

The implications of applying cointegration theory developed for econometric time series to SHM data are discussed in Chapter 7. The Johansen procedure used in this work to establish stationary linear combinations of feature variables assumes that each variable is a difference stationary process - in other words that the generating process of the time series has a unit root. It was argued in this chapter that, although the idea of a unit root generating process does not fit exactly with the time series of interest to SHM, in general these variables will behave similarly to unit root processes. This assertion allows one to apply the cointegration theory without further worry that the processes may not be valid for SHM time series.

In the latter half of Chapter 7, the cointegration process is applied to data from the Tamar monitoring campaign. The process was able to successfully remove the temperature dependent trend from deck displacement measurements. This case study has highlighted a further advantage to using cointegration, which is the fact that the training data needed to establish the stationary linear combination does not necessarily have to span a long period of time in order to be able to remove the trends in the data. Other data normalisation approaches rely on having training data available from a whole year for example, or from every operational condition [62]. The stationary linear combination of the deck displacement measurements was further tested in this chapter through use of a simulated damage scenario. The displacement data was doctored to simulate what might happen if one of the bridge stay cables suffered a loss of tension. When projecting this doctored data

onto the established cointegrating relation the linear combination became clearly nonstationary at the time that the artificial damage was introduced to the data set. These were very encouraging results.

Chapter 8 compared the application of cointegration and principal component analysis (PCA) for the data normalisation problem. Results of the two approaches applied to data from a benchmark study are compared. The benchmark study focused on in this chapter was carried out as part of the EU DAMASCOS consortium and used Lamb-wave propagation to attempt to detect damage introduced to a composite plate under fluctuating temperature conditions. Both methods were able to successfully create features that remained unchanged by the temperature fluctuations but still were able to very clearly detect damage. Some comparisons were made between PCA and cointegration, which on the surface of things are similar methods, both creating linear combinations of original variables. It was found that cointegrating vectors and principal components are not necessarily similar, they are chosen on different criteria and have different orthogonality properties. Although both methods could successfully remove the temperature induced trend, interestingly, it was found that the linear combinations of the minor principal components relied on variables (spectral lines) from different areas of the spectrum than those in the cointegrating linear combinations. Finally an argument was made that cointegration provides a more suitable means of trend removal than using PCA, this is because the focus of the PCA algorithm is not exactly suited to the needs of a data normalisation procedure.

Chapter 9 explores how nonlinear cointegration may be useful for SHM. The benchmark Z24 monitoring campaign is used as a motivational example for why an extension to the cointegration theory used in this thesis would be of benefit. For nonlinear cointegration, a nonlinear combination of variables is required for stationarity. This chapter makes some initial headway into how such a nonlinear combination may be achieved. The ideas explored include using a genetic algorithm for selection of multinomial candidate terms in a nonlinear combination. The use of differential evolution has also been explored for parameter estimation in a fixed form combination. Although the methods applied have been largely successful on synthetic data sets, much more work is needed in order to be able to apply nonlinear cointegration in this form to real data.

10.1 Limitations and further work

SHM for the Tamar Bridge

This is one of the first works to use multiple data from a comprehensive monitoring campaign of an in use bridge for SHM. A main aim, as far as data from the Tamar bridge is concerned, was to simply learn how the structure typically responds to the changing environmental and operational conditions. The drivers of lower modal frequencies are now known and understood well enough to be able to predict their fluctuation to a reasonable degree of accuracy. Conceivably these models could be used in an SHM system in the way described in Chapter 5, in that model error could be used to flag anomalous behaviour. In a similar way cointegration of the modal frequencies and deck displacements could be used to detect anomalous response. At this initial stage of development the signs are encouraging that the construction of an SHM system that works in the face of environmental and operational variations is possible. At this point in time it would be feasible to implement online novelty detectors for deck displacement and the lower deck modal frequencies. Several novelty detectors for each measurement type could be applied, for example using response surface model errors and classic control charts, using GP regression with confidence interval type controls charts and finally using classic control charts with cointegrated residuals. Multiple novelty detectors on the same measurements would enhance confidence in any anomaly detection and may help to avoid issues with false positive identifications of anomalous behaviour.

Although the detection of anomalous behaviour is a positive step in the right direction, it is only an initial step as far as the aims of SHM or performance monitoring are concerned. A detection of novelty alone is uninformative, novelty may occur due to a sensor failure, an extreme weather condition, a performance anomaly or because damage has been introduced into the system. The next stage in the development of an SHM system for the Tamar bridge is to be able to identify the causes of anomaly, or to be able to detect specific changes in structural condition. Generally such challenges require a supervised learning approach and a step away from novelty detection. As discussed in the introductory chapter, this is a challenge in the face of a lack of data available for supervised learning for the kind of events one may be interested in identifying. The author envisages that the ability to identify the causes of anomalous behaviour or structural response, or the ability to identify

particular structural behaviour, will take considerable effort and time to develop. One way to progress to the next step could be to use a high fidelity model of the bridge to simulate specific damage or performance related scenarios and provide data for supervised learning. Such a model is currently under creation [112]. Another possibility that may prove useful is to develop specific novelty detectors whose construction ensures that a positive detection is informative. For example, if the strain measurements of two stay cables were cointegrated and at some stage their stationary residual became nonstationary, the detection of novelty from a control chart would imply that a change had occurred in one of the two stay cables. It is anticipated that of most use would be to implement many such detectors for one structure. If each of the eight stay cables of the Tamar Bridge were to be cointegrated with each other and each residual saved, a loss of tension, for example in one stay cable could be pinpointed by studying which of the residuals become nonstationary. Needless to say, the scope for further work on the development of a working SHM system for the Tamar Bridge is large.

It must also be noted that as far as damage detection is concerned, an obvious limitation when employing the modal data available here, arises from the fact that natural frequencies are well known to be insensitive to localised damage scenarios. This situation is further limited by the fact that only estimates of the five lowest natural frequencies are available, as greater sensitivity to damage scenarios may be gained from higher modes.

Novelty detection with model errors and cointegrated residuals

All of the novelty detection approaches in this work have required a control chart in some form or another. In Chapter 5 it was suggested that GP confidence intervals may be used as a control chart in order to provide a more conservative novelty detector that may be less susceptible to false positive identifications if new environmental and operational conditions were to occur. This idea requires further investigation.

Control charts have also been used as a visual aid to assess the stationarity of cointegrated residuals. A control chart seems a natural way to detect a mean change of a residual and has been successful for use in this thesis, however this is only one constraint on the stationarity of a residual. It is likely that additional and different control chart types may be necessary, for example, one which tracks the variance of a residual. This issue will also bear further investigation.

Cointegration for SHM

This thesis has introduced econometric theory on cointegration for the use of SHM. As discussed in Chapters 6 and 7 the application of cointegration has relied upon a large background of econometric developments. The Johansen procedure has been used to find the cointegrating vectors for feature variables. This procedure was developed for unit root processes which are commonly occurring in econometrics. As discussed in Chapter 7, one is able to apply this to engineering data because the variables of interest display similar features to unit root (difference stationary) processes. Future work that may prove insightful and useful would be the development of an analogue to the Johansen procedure that acts on SHM features when modelled in a more classical way (i.e. as a physics based model), or is applicable when using, for example, a time-dependent autoregressive moving average (TARMA) model [32]. Further investigation into the presence of complex roots and seasonal dummies in AR type models is also planned. Other avenues of research into the application of cointegration for SHM could also include further investigation into tests for stationarity. Although stationarity tests have been studied in this thesis, none of the analysis carried out for the linear application has relied upon them. For the purposes of SHM the author believes that stationarity tests may be exploited and used in place of control charts. This requires further attention, however, there is some suggestion that the ADF test may be low powered in some circumstances [94], which would also require some investigation.

Chapter 9 provided motivation for how nonlinear cointegration may be useful to SHM. Currently, investigation into nonlinear cointegration for SHM is in the beginning stages, and the ideas explored only applied to simulated data. A number of avenues for further research are immediately evident following on from Chapter 9. Firstly, further work is needed to address the noise-variance dependency of the cointegrated residuals that originate from multiplying signals. The suggestion made was that two nonlinearly cointegrated residuals displaying growing or changing variance could then be linear cointegrated to remove this behaviour. This idea needs further investigation and trial. Secondly, a convenient means of finding a suitable nonlinear combination for feature variables is desirable. A suggestion was made that neural networks or Gaussian processes may be of use here. The development of an analogue to the Johansen procedure for nonlinear cointegration would be a very desirable progression. This shall be the focus of further work.

10.2 Continuing Challenges in SHM

10.2.1 Civil infrastructure and SHM

As this thesis has focused for the large part on applications of SHM to civil infrastructure it seems appropriate in this concluding chapter to discuss some of the challenges facing SHM in this context. It is often thought that aerospace applications monopolise, or are the most common concerns of, research carried out in SHM. Looking back through the SHM literature, however, shows that many studies are based on civil infrastructure, or have used it as a base for development and validation of algorithms [9]. This view, therefore, perhaps originates from the fact that many of the considered SHM successes have come from aerospace applications. Some technologies developed for aerospace have, in fact, made the jump from a research interest to applied technology in industry, the most notable example of which is the Health and Usage Monitoring System (HUMS) [113] that has been accepted into the rotorcraft industry as a matter of legislation (in some countries at least). There are a number of clear reasons why this success has, so far, not been mirrored in civil applications.

One potential reason may be attributed to a lack of general consensus within the civil community on how SHM should be approached, which can in turn be attributed to the fact that civil infrastructure is often privately owned, where owners answer to no single regulatory body. In comparison, in the aerospace industry a concerted effort is currently being made by industry and regulatory partners to coordinate and shape the development of SHM for aircraft. An Aerospace Industry Steering Committee (AISC) has been formed for this express purpose [114], where the aim is to develop guidelines and certification requirements and to encourage the use of SHM. Without a parallel effort for civil infrastructure dissemination of good SHM practice is difficult.

A separate issue that will also play a large role in the uptake of SHM for civil infrastructure, is the fact that most civil structures are one-offs, and as such are completely unique [24]. Consequently, an SHM system developed for one structure will not be directly applicable to another, as may be possible with fleets in aerospace. Finally, an additional complication comes from the challenge that the sheer size and complexity of some civil infrastructure provides. Detecting damage in a single

component of a long-span bridge, for example, becomes a very difficult task, on top of the complications that arise from instrumenting a structure of that size.

Due to these challenges, a slightly different outlook on SHM seems to be emerging within the civil community. For some the belief seems to be, that for now certainly, automated condition assessment on a local level, e.g. detection of cracks on structural elements, is unobtainable. The focus, has therefore become, for some, more on monitoring practices, studying global response, and monitoring for performance. Many papers which are published in the name of SHM, solely focus on the monitoring effort of civil infrastructure, which is, of course, not inconsiderable. This is reflected by the fact that the state of the art in civil infrastructure SHM is commonly thought to be the large monitoring campaigns occurring in the East [115], where long-span bridges are instrumented with thousands of sensors. The author would argue that, despite the sophistication, until processes are developed to analyse the measurements obtained in order to make inferences on structural condition, so far only structural monitoring has occurred. The danger in this current trend lies, not in the development of instrumentation, which will always be useful, but in the loss of sight of the fundamental aims of SHM, a monitoring campaign cannot be useful without a process in place that utilises the information obtained. This thesis has gone some way to address this issue, being one of the first works to utilise data from a comprehensive monitoring campaign to begin to develop an understanding of how a healthy structure responds and start on the development of diagnostic tools.

This having been said, the idea of what constitutes SHM for civil applications is perhaps more blurred than in other areas due to the fact that monitoring is often undertaken for slightly different purposes than for other SHM applications. For large scale civil infrastructure, for example, monitoring is often carried out at the start of a structure's life and during construction to ensure that the structure responds in an expected manner to its environmental and operational conditions (especially to wind conditions). This is motivated by the problems (and disasters no less) caused by the phenomena of self-excited oscillations, the most famous example of which being the collapse of the Tacoma Narrows bridge [116]. Although many of these examples motivated the beginnings of research into SHM, the current practice used to safeguard against these events doesn't align exactly with the view of SHM presented in this thesis, as the monitoring is generally not intended to extend into the general assessment of the structure's health throughout its life.

10.2.2 Challenges for SHM in general

Reflecting the difficulty of the challenge, it is an unavoidable fact that, despite the maturity of the research field, little of the SHM practice developed has made it into application in the real world. There are a number of diverse reasons why this might be (some touched on previously in this thesis), and a number of challenges still to face before SHM can become a reality.

On a practical level, some challenges still remain for the instrumentation of a comprehensive monitoring system, such as how to maintain continual data collection on a large scale (this includes consideration of power sources, data transferral and storage, sensor failures, etc.). In practice, it is a considerable challenge to keep an in-place monitoring system working continuously. Other questions arising concern the capability of any sensors now available to detect smaller scale damage on large structures.

One general reason why SHM uptake has been slow could be put down to the fact that aerospace and civil operators have yet to be convinced by any technology developed so far. The most obvious explanation for this is that the developed technology is not yet up to the challenge and must be developed further. This explanation, is perhaps, however not the only contributing factor to why SHM uptake has not been fast. Other explanations concern general negative attitudes towards SHM; while some believe that SHM is an unobtainable vision, others may believe that change is unnecessary. Understandably, operators will be reluctant to accept an automated unproven system that may be expensive and unreliable (issues which must be addressed within the SHM community), but a general negativity towards SHM can be very unhelpful as this attitude will make it difficult in fact to even put test systems in place on commercial structures. Thankfully, however, this attitude is becoming less common; sceptics who where critical of the HUMS system mentioned earlier for rotorcraft, now find that they are reliant on it, and the aerospace industry are now making a concerted effort to shape the future of SHM.

Closely related to attitudes towards SHM is the issue of validation for developed systems. Before any of the benefits of SHM outlined in the first chapter of this thesis can be realised, a proposed system for inference on structural condition must be rigorously proven or validated. A system that fails to detect serious or dangerous faults (referred to as a *false negative* detection of damage), or conversely detects

faults where there are none (*false positive* detection of damage), would have serious life-safety or negative economic consequences respectively, which would completely undermine the reasons for embarking on SHM in the first place.

This question of validation is, however, one of the hardest faced by those working in the field of SHM, especially due to the fact that data from the damaged condition of a structure are hard to come by, which would be one natural way to validate a decision making tool. The problem of validation in the civil infrastructure context is one of the reasons why benchmark structures such as the Z24 and I-40 highway bridges assume such importance. It seems to the author that the question of validation is not commonly addressed in current research in the field, perhaps because it is a premature one in relation to current progress towards the aims of SHM. Research that does address the issue of validation tends to focus on, for example, how scenarios arising from sensor failures can be dealt with. The lack of a real solution to the issue of validation has the consequence that, currently, an SHM system will not be able to completely replace timed visual inspections or routine maintenance of safety critical components, which is perhaps disappointing in view of the fundamental aims of SHM.

Appendix A

Gaussian Process Regression

Gaussian process regression is a powerful Bayesian machine learning tool where predictions and their distributions can be obtained without having to specify a particular parametric model/functional form. Instead, all possible functions that fit the training data (within reason) are considered. This is achieved by defining a distribution that describes the whole set of feasible functions that may fit the data as a Gaussian process. From Rasmussen and Williams [84] a Gaussian process is defined as *"a collection of random variables, any finite number of which have a joint Gaussian distribution."*

If one considers the values of each function of interest at a specific input or time to be a random variable, then the values of the functions at any number of input points or times can be described by a multivariate Gaussian distribution under the Gaussian process assumption. In any practical application, only sampled values of functions over a finite time or set of inputs will be of interest, and so the Gaussian process assumption provides a way to describe all possible functions in that time (or input) frame.

A Gaussian process is completely defined by a mean $m(x)$ and covariance function $k(x, x')$, this means that any finite subset of function values will be distributed according to this mean and covariance function.

Gaussian process regression works by choosing a mean and covariance function that defines the distribution of a suitable set of candidate functions (this is the same as choosing a prior in other Bayesian approaches). This distribution is then con-

ditioned on training data, to give an updated mean and covariance consistent with the training data (conditioning effectively discounts all functions that do not match the training data). This process is outlined in more detail below.

A.1 Gaussian Process Regression

As mentioned above, the starting point of Gaussian process regression is to choose a prior mean and covariance function for the process. In doing this, all possible functions that, over any finite area, have a multivariate Gaussian distribution with this mean and this covariance are considered. As with other Bayesian approaches one must choose the prior carefully as it limits the functions that are considered, for example, specification of a particular covariance function (including hyperparameters) can define the smoothness of all functions considered. To define them more formally, for a real process f, dependent on inputs $\mathbf{x}$, the mean and covariance functions are

$$m(\mathbf{x}) = E[f(\mathbf{x})] \tag{A.1}$$

$$k(\mathbf{x}, \mathbf{x}') = cov(f(\mathbf{x}), f(\mathbf{x}')) = E[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))] \tag{A.2}$$

respectively, where E represents expectation.

Commonly, because little is known about the data at the beginning stage, and for simplification purposes, the prior mean function is set to zero. Choice of the covariance function is therefore critical.

Some notes on the covariance function: The covariance function is always set as a function of the inputs $\mathbf{x}$, hence the notation $k(\mathbf{x}, \mathbf{x}')$, which helps simplify the process later on. A valid covariance function $k(\mathbf{x}_i, \mathbf{x}_j)$ defines a covariance matrix K_{ij}, whose elements are defined by the covariance function evaluated at the points $\mathbf{x}_i$ and $\mathbf{x}_j$, because of this a complete covariance matrix will always be symmetrical about the main diagonal. If a number of points arranged in a design matrix X is considered the covariance matrix is denoted $K(X, X)$. Additionally, if one considers the covariance of function values corresponding to the design matrix X, and another

set of inputs X_*, the covariance matrix will have the following structure:

$$\begin{bmatrix} K(X,X) & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix} \tag{A.3}$$

The fact that a complete covariance matrix must be symmetrical about the main diagonal constrains the possible choices available for covariance functions - only a small number of functions will be valid covariance functions. A common choice of covariance function when using Gaussian processes is the squared exponential, which has a general form [84];

$$cov(f(\mathbf{x}_p), f(\mathbf{x}_q)) = k(\mathbf{x}_p, \mathbf{x}_q) = \exp\left(-1/2|\mathbf{x}_p - \mathbf{x}_q|^2\right) \tag{A.4}$$

This implies that function values at similar inputs will be highly correlated.

Once the prior has been specified with a mean and covariance function any number of functions could be generated from it, should it be desired. However, the real interest lies only in the functions that fit the training data. One way to obtain these functions could be to generate candidate functions from the prior and reject any that don't agree with the training data. As this approach could be very time consuming, an equivalent probabilistic approach which is to condition the prior on the training data (targets) is used instead.

In using a GP, it has been specified that the distribution of the targets in the training data will have a multivariate Gaussian distribution, it has also been specified that any new output data, such as any predictions to be made, will have a multivariate Gaussian distribution. When conditioning the prior on the training data, what is being calculated is the probability of the new targets given the training data. As all targets are multivariate Gaussian, this conditional probability has a known form as follows: for a set of training target values $\mathbf{y}$, and a set of unknown function values $\mathbf{y}_*$ to be predicted with a distribution as follows,

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}_* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} A & C \\ C^T & B \end{bmatrix}\right) \tag{A.5}$$

the conditional distribution $\mathbf{y}_*|\mathbf{y}$ will be

$$\mathbf{y}_*|\mathbf{y} \sim \mathcal{N}(C^T A^{-1} y, B - C^T A^{-1} C). \tag{A.6}$$

This is non-trivial and requires proof, for more details on this see [117] or Rasmussen and Williams Appendix A.2 [84].

Given a set of training data, with inputs arranged in a design matrix X and target values $\mathbf{y}$, a set of testing data with inputs arranged in a design matrix X_* and unknown target values $\mathbf{y}_*$, under the Gaussian process assumption (with a zero mean function), according to the prior, the targets for the training and test set will have a joint Gaussian distribution as follows

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}_* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X,X) & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix}\right) \tag{A.7}$$

The unknown targets, conditioned on the training data, according to (A.6) will then be distributed as follows;

$$\mathbf{y}_*|X,\mathbf{y},X_* \sim \mathcal{N}(K(X_*,X)K(X,X)^{-1}\mathbf{y}, K(X_*,X_*) - K(X_*,X)K(X,X)^{-1}K(X,X_*)) \tag{A.8}$$

From which the mean predicted value $\mathbf{m}_*$ and the variance of that prediction $\mathbf{k}_*$ can be directly lifted as follows:

$$\mathbf{m}_* = K(X_*,X)K(X,X)^{-1}\mathbf{y} \tag{A.9}$$

$$\mathbf{k}_* = K(X_*,X_*) - K(X_*,X)K(X,X)^{-1}K(X,X_*) \tag{A.10}$$

Hyperparameters Equation (A.4) specified the general form of a squared exponential covariance function, in practice additional parameters are added to this form to gain a greater control over the types of functions that are considered for the inference. The squared exponential function that will be used in this work will have the form

$$k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_y^2 \exp\left(-\frac{1}{2l^2}|\mathbf{x}_p - \mathbf{x}_q|^2\right) + \sigma_n^2 \delta_{pq} \tag{A.11}$$

where σ_y^2 is the signal variance (limits the vertical scale of the process), l is the length scale of the process, which defines the smoothness (determines the length between inputs before function values can change significantly), and σ_n^2 is the variance from the noise on the measurements.

As with all Bayesian approaches, the choice of these hyperparameters is very important. The Bayesian way to deal with the uncertainty that comes from choosing specific hyperparameters is to remove their influence from any of the calculations through marginalisation (integrating the hyperparameters out). Through marginalisation one can avoid the problem of choosing specific hyperparameters by specifying a probability distribution for the hyperparameters (using Bayes' rule) and using a double integral. Unfortunately, this integral is usually intractable given any reasonable choice of prior for the hyperparameters. Alternatively, the problem of selecting hyperparameters can be viewed as an optimisation problem.

Within the machine learning community, the most common approach to the problem of hyperparameter choice for practical applications is to use a maximum likelihood approach to optimise the hyperparameters, which avoids the difficulty of direct marginalisation. To do this, the optimal hyperparameters are chosen by maximising the marginal likelihood of the predictions $p(\mathbf{y}|X, \boldsymbol{\theta})$ with respect to the hyperparameters $\boldsymbol{\theta}$. In log form this can be expressed as:

$$\log p(\mathbf{y}|X, \boldsymbol{\theta}) = -\frac{1}{2}\mathbf{y}^T K_y^{-1}\mathbf{y} - \frac{1}{2}\log|K_y| - \frac{n}{2}\log 2\pi \tag{A.12}$$

with K_y the covariance matrix. The likelihood $p(\mathbf{y}|X, \boldsymbol{\theta})$ describes the probability of observing the targets (in the training data set) $\mathbf{y}$ given the input data X and the hyperparameters $\boldsymbol{\theta}$ in the squared exponential covariance function.

Following Rasmussen and Williams [84], when maximising this likelihood one seeks its partial derivative with respect to the hyperparameters:

$$\frac{\partial}{\partial \theta_j}\log\left(p(\mathbf{y}|X, \boldsymbol{\theta}\right) = -\frac{1}{2}\mathbf{y}^T \frac{\partial K}{\partial \theta_j} K_y^{-1}\mathbf{y} - \frac{1}{2}\operatorname{tr}\left(K^{-1}\frac{\partial K}{\partial \theta_j}\right) \tag{A.13}$$

Appendix B

Stationarity of AR and VAR models

This appendix serves as a short introduction to auto-regressive and vector auto-regressive models and their properties relating to stationarity.

B.1 Auto-regressive (AR) Models

The first step of the cointegration procedure involves the generation of an autoregressive model for each nonstationary variable. An auto-regressive model is one that describes the evolution of a time series by a combination of its previous values. Once each variable is represented in AR form it should be determined if that AR model is stationary or nonstationary. Consider the auto-regressive model of order p (AR(p));

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + \cdots + a_p y_{i-p} + \varepsilon_i \tag{B.1}$$

where ε_i can be considered to be a Gaussian white noise process driving the model. $\varepsilon_i \sim N(0, 1)$. Equation (B.1) can be considered to be a randomly forced difference equation. With this is mind, the general solution of (B.1) can be considered in the normal way one would a difference equation:

$$y_i = y_i^c + y_i^p \tag{B.2}$$

where y_i^p is the 'particular integral' and is a solution of (B.1), and y_i^c is the 'complementary function' and is a solution of the equation

$$y_i = a_1 y_{i-1} + a_2 y_{i-2} + \cdots + a_p y_{i-p} \tag{B.3}$$

Assuming a solution to (B.3) of the form $y_i = A\lambda^i$, (B.3) becomes

$$A\lambda^i \left(1 - a_1\lambda^{-1} - \cdots - a_p\lambda^{-p}\right) = 0 \Rightarrow 1 - a_1 z - a_2 z^2 - \cdots - a_p z^p = 0 \tag{B.4}$$

where $z = \frac{1}{\lambda}$. This form is called the characteristic equation of the process, there are p possible λ satisfying (B.4), which in turn leads to a general solution of

$$y_i = A_1\lambda_1^i + A_2\lambda_2^i + \cdots + A_p\lambda_p^i \tag{B.5}$$

where $A_1, \ldots,\ A_p$ are fixed by p initial conditions. The general solution (B.5) can be used to indicate the stability/stationarity of the time series. Looking at (B.5), this solution will remain stable as long as $|\lambda_i| < 1$ for all i, or alternatively as long as $|z_i| > 1$. If any $|z_i| < 1$ the process will behave explosively. Now if any $\lambda_i = 1$, which is called a *unit root*, then $y_i^c \to A_1 + \cdots + A_n$ and the process has marginal stability. In terms of statistics then, the following properties hold true

1. If $|z_i| > 1$ the time series will be stationary
2. If $|z_i| = 1$ (unit root) the time series will be nonstationary
3. If $|z_i| < 1$ the time series will be nonstationary and explosive

B.2 Vector Auto-regressive (VAR) Models

Vector auto-regressive models are an extension of auto-regressive models to include multiple time series. Now, the evolution of two or more time series is described by combinations of past outputs from each series. As before the stability conditions for a VAR model are of interest. A general VAR process of order p takes the form

$$\{y_i\} = [A_1]\{y_{i-1}\} + [A_2]\{y_{i-2}\} + \cdots + [A_p]\{y_{i-p}\} + \{\varepsilon_i\} \tag{B.6}$$

where each ε_i can be considered to be a Gaussian white noise process driving the model; $\varepsilon_i \sim N(0,1)$. Here, the $[A_i]$ are $n \times n$ matrices and $\{y_i\}$ are n-vectors. Again the complementary function $\{y_i^c\}$ is studied, which is a solution of

$$\{y_i\} - [A_1]\{y_{i-1}\} - [A_2]\{y_{i-2}\} - \cdots - [A_p]\{y_{i-p}\} = 0 \tag{B.7}$$

Considering a trial solution of $\{y_i\} = \{\alpha\}\lambda^i$, then as before the characteristic equation of the process can be obtained as

$$(1 - [A_1]z - [A_2]z^2 - \cdots - [A_p]z^p)\{\alpha\} = 0 \tag{B.8}$$

with $z = \frac{1}{\lambda}$. From this form, it is clear that the stability conditions of the process are the same as for the general AR process described previously.

Bibliography

[1] 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 Laboratories, 1996.

[2] A. Rytter. *Vibration-based inspection of civil engineering structures.* PhD thesis, Department of Building Technology and Structural Engineering, University of Aalborg, Denmark, 1993.

[3] K. Worden and J. Dulieu-Barton. Damage identification in systems and structures. *International Journal of Structural Health Monitoring*, 3:85–98, 2004.

[4] J.P. Lynch. An overview of wireless structural health monitoring for civil structures. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):345–372, 2007.

[5] J. Mohammadpour, M. Franchek, and K. Grigoriadis. A survey on diagnostic methods for automotive engines. *International Journal of Engine Research*, 13(1):41–64, 2012.

[6] R.B. Randall. *Vibration-Based Condition Monitoring: Industrial, Aerospace and Automotive Applications.* Wiley-Blackwell, 2011.

[7] J.E. Anderson and M. Fustioni. *Structural Health Monitoring Systems.* COWI-Futurtec, 2006.

[8] K. Worden, C.R. Farrar, G. Manson, and G. Park. The fundamental axioms of structural health monitoring. *Proceedings of the Royal Society A: Mathe-*

matical, Physical and Engineering Science, 463(2082):1639, 2007.

[9] 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 Laboratories, 2002.

[10] E.P. Carden and P. Fanning. Vibration based condition monitoring: a review. *Structural Health Monitoring*, 3(4):355, 2004.

[11] W. Staszewski, G. Tomlinson, C. Boller, and G. Tomlinson. *Health monitoring of aerospace structures*. Wiley Online Library, 2004.

[12] S.R. Hunt and I.G. Hebden. Validation of the Eurofighter Typhoon structural health and usage monitoring system. *Smart materials and structures*, 10:497, 2001.

[13] R. Brincker, C. Ventura, and P. Andersen. Why output-only modal testing is a desirable tool for a wide range of practical applications. In *Proceedings of the International Modal Analysis Conference (IMAC XXI)*, volume 265, 2003.

[14] C. Rainieri and G. Fabbrocino. Automated output–only dynamic identification of civil engineering structures. *Mechanical Systems and Signal Processing*, 24 (3):678–695, 2010.

[15] B. Peeters, J. Maeck, and G. De Roeck. Vibration-based damage detection in civil engineering: Excitation sources and temperature effects. *Smart Materials and Structures*, 10(3):518–527, 2001.

[16] S.J. Lee and H. Sohn. Active self-sensing scheme development for structural health monitoring. *Smart Materials and Structures*, 15(6):1734–1746, 2006.

[17] A. Raghaven and C.E.S. Cesnik. Review of guided-wave structural health monitoring. *Shock and Vibration Digest*, 39:91–114, 2007.

[18] C.C. Ciang, J.R. Lee, and H.J. Bang. Structural health monitoring for a wind turbine system: a review of damage detection methods. *Measurement Science and Technology*, 19:122001, 2008.

[19] P. Rizzo and F.L. di Scalea. Acoustic emission monitoring of carbon-fiber-reinforced-polymer bridge stay cables in large-scale testing. *Experimental mechanics*, 41(3):282–290, 2001.

[20] G. Park and D.J. Inman. Structural health monitoring using piezoelectric impedance measurements. *Philosophical Transactions of the Royal Society, Series A*, 365:373–392, 2007.

[21] J.P. Lynch and K.J. Loh. A summary review of wireless sensors and sensor networks for structural health monitoring. *Shock and Vibration Digest*, 38(2): 91–130, 2006.

[22] G. Park, T. Rosing, M.D. Todd, C.R. Farrar, and W. Hodgkiss. Energy harvesting for structural health monitoring sensor networks. *Journal of Infrastructure Systems*, 14:64, 2008.

[23] D.L. Mascarenas, E.B. Flynn, M.D. Todd, T.G. Overly, K.M. Farinholt, G. Park, and C.R. Farrar. Experimental studies of using wireless energy transmission for powering embedded sensor nodes. *Journal of Sound and Vibration*, 329(12):2421–2433, 2010.

[24] J.M.W. Brownjohn. Structural health monitoring of civil infrastructure. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):589, 2007.

[25] C.R. Farrar and K. Worden. An introduction to structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):303, 2007.

[26] K. Worden, W.J. Staszewski, and J.J. Hensman. Natural computing for mechanical systems research: A tutorial overview. *Mechanical Systems and Signal Processing*, 25(1):4–111, 2011.

[27] H. Wenzel. *Health Monitoring of Bridges*. John Wiley and Sons, Ltd, 2009.

[28] H. Sohn, C.R. Farrar, N.F. Hunter, and K. Worden. Structural health monitoring using statistical pattern recognition techniques. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, 123(4):706–711, 2001.

[29] T. Söderström and P. Stoica. *System identification*. Prentice-Hall, Inc., 1988.

[30] B. Peeters and G. De Roeck. Reference-based stochastic subspace identification for output-only modal analysis. *Mechanical Systems and Signal Processing*, 13(6):855–878, 1999.

[31] S.D. Fassois and J.S. Sakellariou. Time-series methods for fault detection and identification in vibrating structures. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):411–448, 2007.

[32] A.G. Poulimenos and S.D. Fassois. Parametric time-domain methods for non-stationary random vibration modelling and analysisa critical survey and comparison. *Mechanical Systems and Signal Processing*, 20(4):763–816, 2006.

[33] E. Papatheou, G. Manson, R.J. Barthorpe, and K. Worden. The use of pseudo-faults for novelty detection in SHM. *Journal of Sound and Vibration*, 329: 2349–2366, 2010.

[34] M.I. Friswell and J.E. Mottershead. *Finite element model updating in structural dynamics*. Springer, 1995.

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

[36] N.N. Taleb. *The Black Swan: The Impact of the Highly Improbable*. Penguin, 2011.

[37] H. Sohn. Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):539, 2007. ISSN 1364-503X.

[38] T.H. Yi, H.N. Li, and M. Gu. Full-scale measurements of dynamic response of suspension bridge subjected to environmental loads using GPS technology. *SCIENCE CHINA Technological Sciences*, 53(2):469–479, 2010.

[39] J.M.W. Brownjohn and P. Carden. Real-time operation modal analysis of Tamar Bridge. In *26th International Modal Analysis Conference (IMAC XXVI), Orlando, Florida, USA*, pages 4–7, 2008.

[40] C.R. Farrar, H. Sohn, and K. Worden. Data normalization: a key for structural health monitoring. Technical report, Los Alamos National Laboratory, 2001.

[41] S. Alampalli. Effects of testing, analysis, damage, and environment on modal parameters. *Mechanical Systems and Signal Processing*, 14(1):63–74, 2000.

[42] P. Cornwell, C.R. Farrar, S.W. Doebling, and H. Sohn. Environmental variability of modal properties. *Experimental Techniques*, 23(6):45–48, 1999.

[43] H. Sohn, M. Dzwonczyk, E.G. Straser, A.S. Kiremidjian, K.H. Law, and T. Meng. An experimental study of temperature effect on modal parameters of the Alamosa Canyon Bridge. *Earthquake engineering & structural dynamics*, 28(8):879–897, 1999.

[44] P. Moser and B. Moaveni. Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge. *Mechanical Systems and Signal Processing*, 25:23362357, 2011.

[45] Y.J. Ge and H. Tanaka. Aerodynamic flutter analysis of cable-supported bridges by multi-mode and full-mode approaches. *Journal of Wind Engineering and Industrial Aerodynamics*, 86(2):123–153, 2000.

[46] S.R. Chen and J. Wu. Performance enhancement of bridge infrastructure systems: Long-span bridge, moving trucks and wind with tuned mass dampers. *Engineering structures*, 30(11):3316–3324, 2008.

[47] D.M. Siringoringo and Y. Fujino. System identification of suspension bridge from ambient vibration response. *Engineering Structures*, 30(2):462–477, 2008.

[48] Y. Xia, H. Hao, G. Zanardo, and A. Deeks. Long term vibration monitoring of an RC slab: Temperature and humidity effect. *Engineering structures*, 28 (3):441–452, 2006.

[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] Q.W. Zhang, L.C. Fan, and W.C. Yuan. Traffic-induced variability in dynamic properties of cable-stayed bridge. *Earthquake engineering & structural dynamics*, 31(11):2015–2021, 2002.

[51] F. Lanata and F. Schoefs. Multi-algorithm approach for identification of structural behavior of complex structures under cyclic environmental loading. *Structural Health Monitoring*, 2011.

[52] H. Sohn, M. Dzwonczyk, E.G. Straser, K.H. Law, T. Meng, and A.S. Kiremid-

jian. Adaptive modeling of environmental effects in modal parameters for damage detection in civil structures. *Smart Systems for Bridges, Structures, and Highways*, pages 127–138, 1998.

[53] P. Andersen, P.H. Kirkegaard, and R. Brincker. Filtering out environmental effects in damage detection of civil engineering structures. In *Proceedings of SPIE, the International Society for Optical Engineering*, volume 3089, pages 905–911. Society of Photo-Optical Instrumentation Engineers, 1997.

[54] B. Moaveni, X. He, J.P. Conte, M. Fraser, and A. Elgamal. Uncertainty analysis of Voigt Bridge modal parameters due to changing environmental condition. In *Proceedings of International Conference on Modal Analysis (IMAC-XXVII)*, 2009.

[55] Y.Q. Ni, X.G. Hua, K.Q. Fan, and J.M. Ko. Correlating modal properties with temperature using long-term monitoring data and support vector machine technique. *Engineering structures*, 27(12):1762–1773, 2005.

[56] X.G. Hua, Y.Q. Ni, J.M. Ko, and K.Y. Wong. Modeling of temperature–frequency correlation using combined principal component analysis and support vector regression technique. *Journal of computing in civil engineering*, 21:122, 2007.

[57] Y.Q. Ni, H.F. Zhou, and J.M. Ko. Generalization capability of neural network models for temperature- frequency correlation using monitoring data. *Journal of Structural Engineering*, 135(10):1290–1300, 2009.

[58] K. Worden, H. Sohn, and C.R. Farrar. Novelty detection in a changing environment: regression and interpolation approaches. *Journal of Sound and Vibration*, 258(4):741–761, 2002. ISSN 0022-460X.

[59] E. Balmes, M. Basseville, L. Mevel, and H. Nasser. Handling the temperature effect in vibration monitoring of civil structures: A combined subspace-based and nuisance rejection approach. *Control Engineering Practice*, 17(1):80–87, 2009.

[60] L.D. Avendaño-Valencia, M.D. Spiridonakos, and S.D. Fassois. In-operation identification of a wind turbine structure via non-stationary parametric models. In *Proceedings of the International Workshop on Structural Health Monitoring IWSHM*, page 2611, 2011.

[61] C. Surace and K. Worden. Some aspects of novelty detection methods. In *Proceedings of the third international conference on Modern Practice in Stress and Vibration Analysis*, Dublin, 1997.

[62] J. Kullaa. Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring. *Mechanical Systems and Signal Processing*, 25(8):1–8, 2011.

[63] G. Manson. Identifying damage sensitive, environment insensitive features for damage detection. In *Proceedings of the IES conference*, 2002.

[64] A.M. Yan, G. Kerschen, P. De 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.

[65] J. Kullaa. Structural health monitoring under variable environmental or operational conditions. In *Proceedings of the 2nd European Workshop on Structural Health Monitoring*, 2004.

[66] A. Deraemaeker, E. Reynders, G. De Roeck, and J. Kullaa. Vibration-based structural health monitoring using output-only measurements under changing environment. *Mechanical Systems and Signal Processing*, 22(1):34–56, 2008.

[67] A.-M. Yan, G. Kerschen, P. De Boe, and J.-C. Golinval. Structural damage diagnosis under varying environmental conditions part ii: local PCA for nonlinear cases. *Mechanical Systems and Signal Processing*, 19:865–880, 2005.

[68] 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, 2002.

[69] C.K. Oh, H. Sohn, and I.H. Bae. Statistical novelty detection within the Yeongjong suspension bridge under environmental and operational variations. *Smart Materials and Structures*, 18:125022, 2009.

[70] C.K. Oh and H. Sohn. Damage diagnosis under environmental and operational variations using unsupervised support vector machine. *Journal of Sound and Vibration*, 325(1-2):224–239, 2009.

[71] 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.

[72] S. Vanlanduit, E. Parloo, B. Cauberghe, P. Guillaume, and P. Verboven. A robust singular value decomposition for damage detection under changing operating conditions and structural uncertainties. *Journal of sound and vibration*, 284(3-5):1033–1050, 2005.

[73] E. Figueiredo, M.D. Todd, C.R. Farrar, and E. Flynn. Autoregressive modeling with state-space embedding vectors for damage detection under operational variability. *International Journal of Engineering Science*, 48(10):822–834, 2010.

[74] E.J. Figueiredo. *Damage Identification in Civil Engineering Infrastructure under Operational and Environmental Conditions*. PhD thesis, University of Porto, 2010.

[75] B. Peeters and G. De Roeck. Stochastic system identification for operational modal analysis: a review. *Journal of Dynamic Systems, Measurement, and Control*, 123:659, 2001.

[76] R. Brincker and P. Andersen. Understanding stochastic subspace identification. In *Proceedings of International Conference on Modal Analysis (IMAC-XXV)*, 2006.

[77] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Kluwer Academic Publishers, 1996.

[78] UK Department for Transport. Transport statistics bulletin, road statistics 2007. Technical report, 2007.

[79] S. Sharma. *Applied multivariate techniques*. John Wiley & Sons, Inc., 1995. ISBN 0471310646.

[80] G. Box and N. Draper. *Empirical model-building and response surface*. John Wiley & Sons Inc., New York, 1986.

[81] G.E.P. Box and K.B. Wilson. On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B (methodological)*, 13(1):1–45, 1951.

[82] D.G. Kleinbaum, L.L. Kupper, and K.E. Muller. *Applied regression analysis*

and other multivariable methods. Duxbury Pr, 2007.

[83] B. Peeters and G. De Roeck. One-year monitoring of the Z24-Bridge: environmental effects versus damage events. *Earthquake engineering & structural dynamics*, 30(2):149–171, 2001.

[84] C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*, volume 38. The MIT Press, Cambridge, MA, USA, 2006.

[85] S. Mohanty, A. Chattopadhyay, P. Peralta, and S. Das. Bayesian statistic based multivariate Gaussian process approach for offline/online fatigue crack growth prediction. *Experimental mechanics*, 51:833–843, 2011.

[86] D.C. Montgomery. *Introduction to statistical quality control*. John Wiley & Sons, 2009. ISBN 812651471X.

[87] J. Kullaa. Sensor validation using minimum mean square error estimation. *Mechanical Systems and Signal Processing*, 24(5):1444–1457, 2010.

[88] J.H. Stock and M.W. Watson. Testing for common trends. *Journal of the American statistical Association*, 83(404):1097–1107, 1988. ISSN 0162-1459.

[89] J.H. Stock and M.W. Watson. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, 61(4):783–820, 1993. ISSN 0012-9682.

[90] Q. Chen, U. Kruger, and A. Leung. Cointegration testing method for monitoring nonstationary processes. *Industrial & Engineering Chemistry Research*, 48(7):3533–3543, 2009.

[91] S. Johansen. *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press, USA, 1995. ISBN 0198774508.

[92] W. A. Fuller. *Introduction to Statistical Time Series*. Wiley Series in Probability and Statistics. Wiley - Interscience, second edition edition, 1996.

[93] G.S. Maddala and I.M. Kim. *Unit roots, cointegration, and structural change*. Number 4. Cambridge University Press, 1998.

[94] K. Juselius. *The cointegrated VAR model: methodology and applications*. Oxford University Press, Oxford, UK, 2006.

[95] R.F. Engle and C.W. Granger. Cointegration and error correction: representation, estimation and testing. *Econometrica*, 55(2):251–276.

[96] S. Johansen. Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2-3):231–254, 1988.

[97] 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*, pages 427–431, 1979. ISSN 0162-1459.

[98] 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 (4):1057–1072, 1981. ISSN 0012-9682.

[99] R. Perman. Cointegration: an introduction to the literature. *Journal of Economic Studies*, 18(3), 1993. ISSN 0144-3585.

[100] T.W. Anderson. *The statistical analysis of time series*. Wiley New York, 1971. ISBN 0471029009.

[101] G.U. Yule. Why do we sometimes get nonsense-correlations between time-series?–a study in sampling and the nature of time-series. *Journal of the royal statistical society*, 89(1):1–63, 1926.

[102] R.K. Kaufmann and D.I. Stern. Cointegration analysis of hemispheric temperature relations. *Journal of Geophysical Research*, 107:4012, 2002.

[103] T. Verbeke and M. De Clercq. The ekc: some really disturbing monte carlo evidence. *Environmental Modelling & Software*, 21(10):1447–1454, 2006.

[104] D.I. Stern and R.K. Kaufmann. Detecting a global warming signal in hemispheric temperature series: Astructural time series analysis. *Climatic Change*, 47(4):411–438, 2000.

[105] G. Dufrénot and V. Mignon. *Recent developments in nonlinear cointegration with applications to macroeconomics and finance*. Springer Netherlands, 2002.

[106] H.J. Bierens. Nonparametric nonlinear co-trending analysis, with an application to inflation and interest in the us. *Journal of Business and Economic Statistics*, 18:323–337, 2000.

[107] G. De Roeck. The state-of-the-art of damage detection by vibration moni-

toring: the SIMCES experience. *Journal of Structural Control*, 10:127–134, 2003.

[108] K. Worden, G. Manson, H. Sohn, and C. Farrar. Extreme value statistics from differential evolution for damage detection. In *Proceedings of the 23rd International Conference on Modal Analysis, Florida*, 2005.

[109] R. Storn and K. Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359, 1997.

[110] D.E. Goldberg. *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Inc., 1989.

[111] Z. Michalewicz. *Genetic Algoritms + Data Structures = Evolution Programs*. Springer, 3rd edition, 1996.

[112] R.J. Westgate and J.M.W. Brownjohn. Development of a Tamar Bridge finite element model. *Dynamics of Bridges, Volume 5*, pages 13–20, 2011.

[113] J.D. Chronkite. Practical application of health and usage monitoring (HUMS) to helicopter rotor, engine and drive system. In *Proceedings of the 49th Forum of the American Helicopter Society.*, 1993.

[114] P. Foote. The aerospace industry steering committee on structual health monitoring management (AISC SHM): Progress on SHM guidelines for aerospace. In *Proceedings of the 8th International Workshop on Structural Health Monitoring*, 2011.

[115] J.M. Ko and Y.Q. Ni. Technology developments in structural health monitoring of large-scale bridges. *Engineering structures*, 27(12):1715–1725, 2005.

[116] A. Larsen. Aerodynamics of the Tacoma narrows bridge - 60 years later. *Structural Engineering International*, 10(4):243–248, 2000.

[117] C.M. Bishop. *Pattern recognition and machine learning*. springer New York, 2006.