



This is a repository copy of *A Non-linear Manifold Strategy for SHM Approaches*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/87278/>

Version: Accepted Version

Article:

Dervilis, N., Antoniadou, I., Cross, E.J. et al. (1 more author) (2015) A Non-linear Manifold Strategy for SHM Approaches. *Strain*. ISSN 0039-2103

<https://doi.org/10.1111/str.12143>

This is the peer reviewed version of the following article: Dervilis, N., Antoniadou, I., Cross, E.J. and Worden, K. (2015) A Non-linear Manifold Strategy for SHM Approaches. *Strain*, which has been published in final form at <http://dx.doi.org/10.1111/str.12143>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

A Nonlinear Manifold Strategy for SHM Approaches

N.Dervilis*, I. Antoniadou, E.J.Cross, K.Worden

*Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield,
Mappin Street, Sheffield S1 3JD, England.*

Abstract

In the data-based approach to Structural Health Monitoring (SHM) when novelty detection is utilised as a means of diagnosis, benign operational and environmental variations of the structure can lead to false alarms and mask the presence of damage. The key element of this paper is to demonstrate a series of pattern recognition approaches which investigate complex correlations between the variables and thus potentially shed light on the variations within the data that are of interest for SHM. The nonlinear manifold learning techniques discussed here, like locally linear embedding (LLE) combined with robust discordance measures like the minimum covariance determinant (MCD) and regression techniques like Gaussian processes (GPs) offer a strategy that includes reliable novelty detection analysis but also a method of investigating the space where structural data clusters are lying.

Keywords: environmental and operational variations, manifold learning, pattern recognition, Structural Health Monitoring (SHM), Gaussian processes.

Email address: N.Dervilis@sheffield.ac.uk (N.Dervilis*)

1. Introduction

When SHM technology is adopted as a tool for monitoring a structure, then the system often has to run continuously and on-line. The effects of any environmental variations must be considered and identified before choosing and using a reliable feature for revealing any structural condition.

A catholic argument is that no sensor exists that can directly measure any type of novelty. For this reason, feature extraction is used to derive useful metrics from the raw data that can further be post-processed through advanced signal processing tools. The methods for feature extraction serve two purposes; a reduction in the dimensionality by mapping the data from high-dimensional spaces to lower-dimensional spaces and a revealing of hidden aspects of the data by learning the structure between the variables of interest.

One of the most challenging tasks in SHM methodology is to understand and eliminate the influence of temperature on structural response. Especially for bridges, which is the immediate concern of this study, temperature is generally understood to be an important environmental factor which affects the dynamic response of the structure, due to its influence on the stiffness of structural parameters and on the boundary conditions of the structure [1–5].

Various methods and algorithms have been proposed in order to counteract and remove the influence of external variations such as principal component analysis, auto-associative neural networks [6] or more recently cointegration [1, 7–13]. Although, these methods exhibit a series of advantages and disadvantages in terms of removing the influence of operational and environmental conditions, little effort has been carried out in terms of constructing a robust chain of methods that characterise the manifold that is constructed between the variables and distinguishing which of the outliers indicated in the data are

representing environmental/operational variations and which are representing damage or structural performance degradation. For more details regarding the different natures of outliers in multivariate statistics the reader can consult [14–17].

A recent paper by Dervilis et al.[17] explores an approach of robust regression and robust multivariate statistics as a means of characterising and distinguishing the influence of environmental and operational conditions on the structural response. Specifically, the outliers may arise in the data as the result of both benign and malign causes and it is important to understand their sources[17].

The layout of this paper is as follows. The discussion begins with the description of the Z24 bridge. In Section 3, the strategy that will be followed in this paper is presented. Section 4 gives some background analysis regarding the nonlinear manifold learning approach. The study concludes with the presentation of some key results.

2. A quick overview of the Z24 bridge

The Z24 Bridge (see Fig.1) was a concrete highway structure in Switzerland connecting Koppigen and Utzenstorf, and in the late 1990s, before its demolition procedure, it was used for SHM purposes under the “SIMCES” project [1, 18]. During a whole year of monitoring of the bridge, a series of sensor systems captured modal parameter measurements, as well as a family of environmental measurements such as air temperature, soil temperature, humidity, wind speed etc. The critical point in this benchmark project was the introduction of different types of real progressive damage scenarios towards the end of the monitoring year (Table.1).

For the purposes of this study, the four natural frequencies that were extracted over a period of year, including the period of structural failure of the bridge are used. Fig.2 shows the four natural frequencies with values between 0-12 Hz (vertical $y - axis$ is the natural frequencies in Hz). The beginning of the introduced failure occurs at observation 3476. The time instances between Fig.2 and Fig.3 are the same. It has to be mentioned that some values of failed measurements have been removed.

Sequence	Damage scenarios
1	Settlement of foundation
2	Tilt of foundation
3	Spalling of concrete at soffit
4	Landslide
5	Failure of concrete hinges
6	Failure of anchor heads
7	Number of post-tensioning tendon failures

Table 1: Progressive damage scenarios.

The Z24 bridge was recently extensively analysed using robust methods such as least trimmed squares (LTS) and minimum covariance determinant (MCD) techniques as means of exploring environmental variations for SHM purposes in a previous and ongoing work [16, 17]. It was found that environmental variations due to sub-zero temperatures manifest themselves differently in feature space compared with the damaged condition. This was a very vital information as it showed that the nature of outliers between operational/environmental variations and damage may have markedly different characteristics.



Figure 1: The Z24 bridge.

Furthermore, it was found that the Z24 bridge has a highly nonlinear behaviour.

It can be noted that there are some visible fluctuations between observations 1200-1500 (below -5 Celsius degrees). As one can see there are no visible fluctuations after the introduction of damage (3476) and is clear that the temperature fluctuation masks the dynamic presence of damage. This is the reason that advanced machine learning techniques are utilised as a means of revealing the hidden characteristics of the structural modal data.

The critical fluctuations are highly related to periods of very cold temperatures much under zero degrees Celsius and there is a direct connection with increased stiffness based on the freezing of the asphalt layer of the bridge deck. In turn, these large temperature fluctuations are suitable candidates to introduce nonlinear characteristics.

The motivation of this paper is to reveal the nonlinear manifold between the natural frequencies and then to try to remove these temperature variations

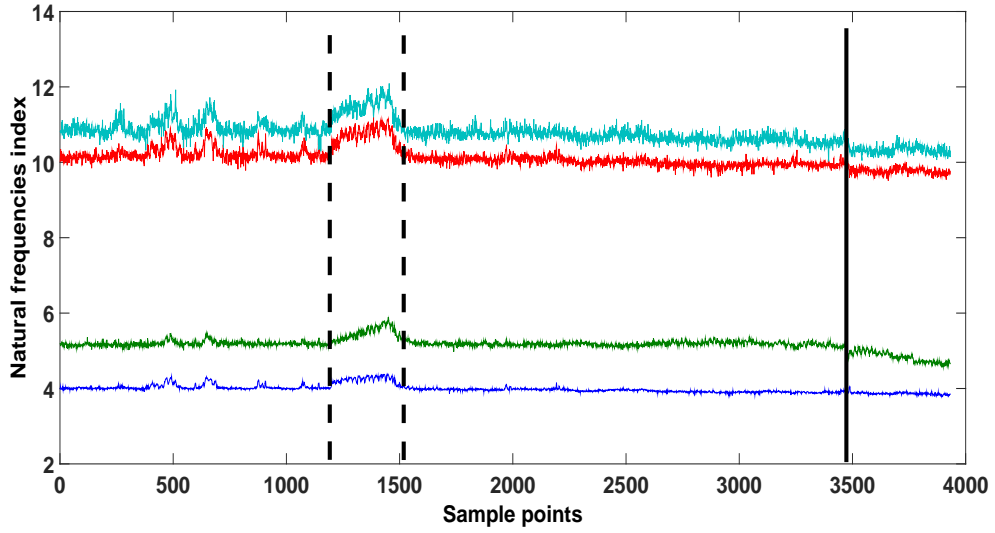


Figure 2: Time history of natural frequencies of the Z24 Bridge (The dotted lines represent the very cold temperatures fluctuation and the black solid line the introduction of damage).

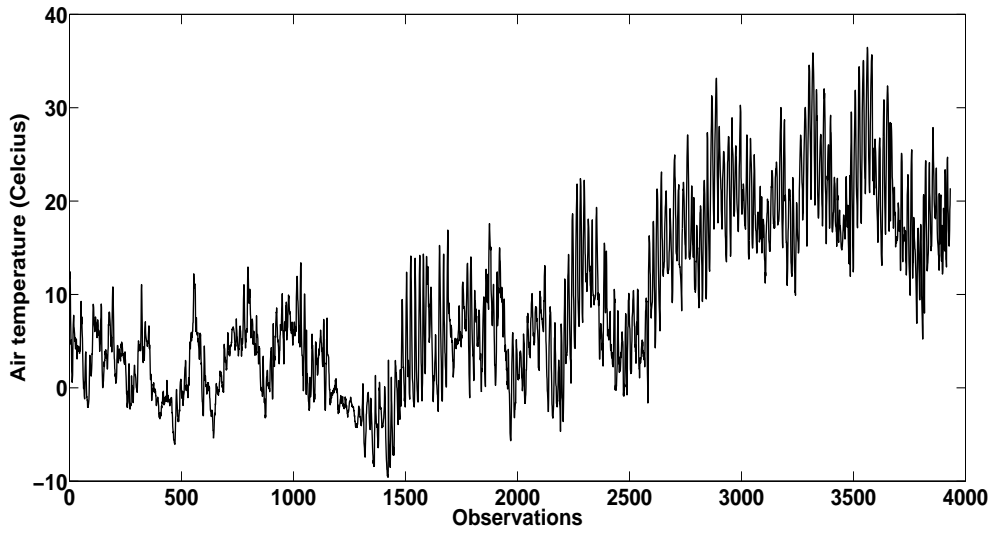


Figure 3: Time history of deck temperature.

and detect clearly the damage. The next section describes the strategy that will be followed.

3. Strategy

The chain of methods applied in this paper aims to investigate the appearance of benign fluctuations in data from the Z24 bridge. First, the whole data set of the four natural frequencies is reduced to two dimensions using a nonlinear manifold technique, in this case locally linear embedding (LLE) (nonlinear principal component analysis via the usage of auto-associative neural networks (AANN) [6] is another strong method). For the current purposes, LLE is used as it is a much simpler tool than AANN.

Next, the minimum covariance determinant estimator (MCD) index is used in order to reveal inclusive outliers without a priori knowledge of whether benign variations are present in the normal test data [14–16, 19–22]. The application of robust computation of location and covariance estimation of multivariate data is of significant interest in the investigation of inclusive outliers (for more details see also Appendix A).

In order to make visible the influence of temperature on the measured natural frequencies, a powerful automatic clustering technique like affinity propagation algorithm (AP) can be applied [23, 24], as here. AP identifies exemplars among data points and forms clusters of data points around these exemplars. The specific algorithm operates by simultaneously considering all data points as potential exemplars and exchanging messages between data points until a good set of exemplars and clusters emerges. More detailed information about the exact procedure on how the AP algorithm is passing the messages between data points can be found in [23, 24].

The strategy finishes by trying to predict the components of the manifold in order to remove any doubt about which data is influenced by environmental fluctuations and which belongs to the damaged case. Use of Gaussian processes is a current research area of increasing interest, not only for regression but also for classification purposes (for more details readers are referred to Appendix B and [25]). Gaussian processes (GPs) are a stochastic nonparametric Bayesian approach to regression and classification problems. These Gaussian processes are computationally very efficient and the nonlinear learning is relatively easy. Gaussian process regression takes into account all possible functions that fit to the training data vector and gives a predictive distribution around a single prediction for a given input vector. A mean prediction and confidence intervals on this prediction can be calculated from the predictive distribution.

The initial and basic step in order to apply Gaussian process regression is to obtain a mean and covariance function. These functions are specified separately, and consist of a specification of a functional form and a set of parameters called hyperparameters. Here, a zero-mean function and a squared-exponential covariance function are applied (see Appendix B or [25]). When the mean and covariance functions are defined, then the inference method specifies the calculation of the exact model and in simple terms describes how to compute hyperparameters by determining the minimisation of the negative log marginal likelihood. The software used for the implementation of GP regression was provided by [25].

Setting an appropriate threshold in the absence of any damage-state data, as is the case in this study, is a non-trivial task. A Monte Carlo simulation based on extreme value statistics was used. The procedure that was conducted in order to calculate the threshold is as follows:

- A $p \times n$ (*number of dimensions* \times *number of observations*) matrix is constructed with each individual element a randomly generated number from a normal distribution with zero mean and unit standard deviation.
- The distance value is evaluated for all matrix values, where the robust mean and covariance matrix are inclusive. The largest (i.e. extreme), value recorded for each trial matrix is stored.
- The process is repeated for a large number of trials in order to generate an array of “extreme” distance calculations. Next, all the values are ordered in terms of magnitude. The critical values (alpha value, α) can take different values such as 5% or 1% for a test of discordancy. In this paper α is set equal to 1% giving a 99% confidence limit.

4. Nonlinear manifold learning via locally linear embedding

As mentioned in previous sections, the combination of strong nonlinearity and the influence of environmental fluctuations makes the damage detection performance very weak. This is the reason that a quick and effective method of nonlinear manifold learning such as locally linear embedding is introduced in this study [26, 27].

An extensive overview of the algorithm can be found in [26, 27]. Briefly and for the purposes of this paper a short description is given here.

The LLE method is based on simple geometric intuition. If the observations consist of N real-valued vectors $\{x_i\}$ with dimensions D and they are sampled from a smooth underlying nonlinear manifold, then each data point and its neighbours is expected to lie on or close to a locally formed patch of the manifold. The local geometries can be characterised by finding linear

coefficients that can reconstruct each data point with respect to each set of neighbours.

If one establishes K nearest neighbours per data point then the reconstruction error is given by the cost function:

$$error(W) = \sum_i \left| \{x_i\} - \sum_j [W_{ij}] \{x_j\} \right|^2 \quad (1)$$

where $[W_{ij}]$ is the weight contribution of the j^{th} data point to the i^{th} reconstruction. In order to compute these weights the cost function has to be minimised under the following constraints. The reconstruction errors that are subject to the constrained weights should be invariant to rotations and rescaling. In turn, in order that the LLE algorithm preserves this invariant manifold idea as a final step of the method, each measurement $\{x_i\}$ should be mapped to a lower dimensional vector $\{Y_i\}$ that minimises the cost function:

$$error(Y) = \sum_i \left| \{Y_i\} - \sum_j [W_{ij}] \{Y_j\} \right|^2 \quad (2)$$

The main difference with the previous cost function is that here the weights are fixed but the $\{Y_i\}$ co-ordinates are optimised.

5. Revealing the nonlinear manifold between the natural frequencies

As can be seen from Fig.4, the Z24 natural frequency data is projected into 2-D space using the LLE algorithm and in the Table.2 a description of the different data sets in relation to temperature is given.

The condition “undamaged” is given as a label to the data when the monitoring campaign started. The labels are just for characterising the “condition” just throughout the monitoring year when the project started (as of course, the bridge had been in operational service for some period before the monitoring campaign has started).

The manifold that is revealed is giving two distinctive directions of the data sets, one regarding the cold and very cold temperatures (green) and one regarding the damage observations (black). Furthermore, it is worth noting that the hot temperatures (red) lie in the same space as some early damage data. If one tries to identify outliers without setting beforehand a training set by using the MCD method, one can see that both the cold temperatures and the damage will appear as outliers.

Of course, this connection between data and temperature is known if measurements of the temperatures were obtained. If the measurements of the temperatures were not known beforehand then an automatic clustering method can be applied, in this case the AP algorithm. This automatic clustering method is presented as a novel future tool in structural data analysis as it offers the advanced characteristic of defining different categories within the manifold of the variables. As an example in this case study in order to make visible the dramatic influence of temperature on the measured natural frequencies the AP tool presents the high potential of applying it on modal data.

It can be seen in Fig.6 (free to find maximum number of clusters) and especially Fig.7 (restrained), that the AP algorithm finds 5 classes which have a very good agreement with the separation of data presented in Table.2. The AP algorithm finds the cold, normal and hot temperature (and some damage

data) influenced points as well as three stages of damage observations which, in comparison with the MCD index, are showing a progressive direction.

As a last step after the MCD and AP tools are applied one can use GP regression in order to predict the 1st component based on the 2nd one and vice versa. As a training set for the GPs only the first 500 points were used and the rest as testing set.

GP regression model error is used as an index of abnormal response. Furthermore, as it will be seen later using this regression error (residual error, which is the difference between the algorithm predictions and the actual data) a strong visualisation that indicates when faults occur will be presented. The calculation of the threshold is explained in [6, 16].

As can be seen from Figs.8-9 the GPs predict/classify correctly most of temperatures (especially Fig. 9 and mainly the very cold temperatures which were found as outliers in Fig.5) and the residual error exceeds the thresholds when damage is present.

This is a strongly encouraging result as the strategy followed, manages to a great extent, to minimise the novelty due to temperature variations by learning their nonlinear characteristics and by applying a strong nonlinear regression tool like GPs to detect novelty that is directly connected with the damaged state of the bridge.

It has to be pointed out that the training involved the first 500 points which means no freezing temperatures below -5 Celsius degrees (as between points 1201-1500) are utilised in the training data set. This key comment is a vital point in validating that the chain of strategic steps that are followed here can offer a useful tool in the robust investigation of benign variations during a monitoring campaign.

Observation	Condition	Colour
1-400	undamaged	blue
401-700	undamaged (with some cold temperature variations)	yellow
701-1200	undamaged (with some cold temperature variations)	pink
1201-1500	undamaged (very cold temperature)	green
1501-2200	undamaged	brown
2201-3475	undamaged (with hot temperature variations)	red
3476-3932	damaged	black

Table 2: Description of data sets as they appear in the Figs.4-5 and 8-9.

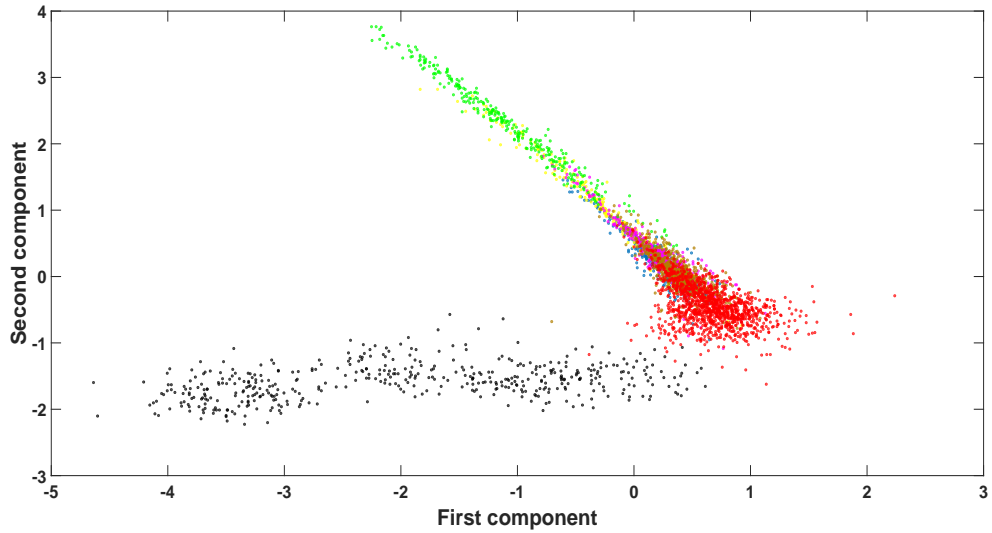


Figure 4: Data projection of the four natural frequencies onto 2-D space using LLE.

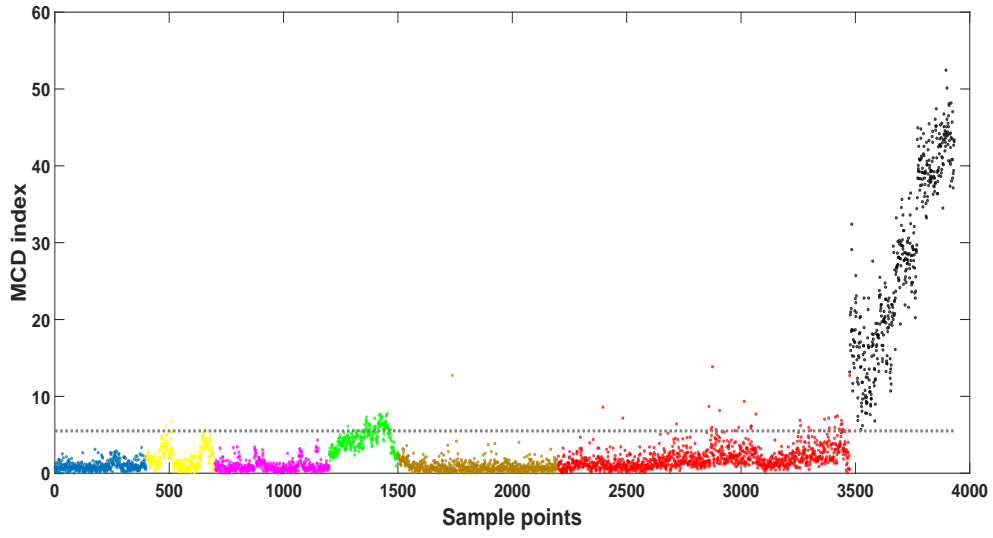


Figure 5: MCD index for LLE.

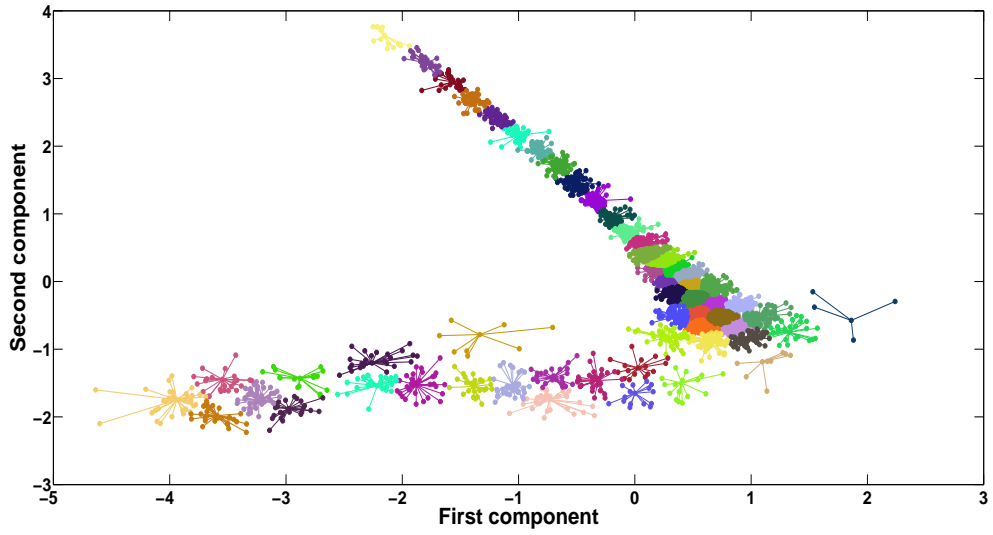


Figure 6: AP algorithm identifies 55 different classes on the LLE manifold.

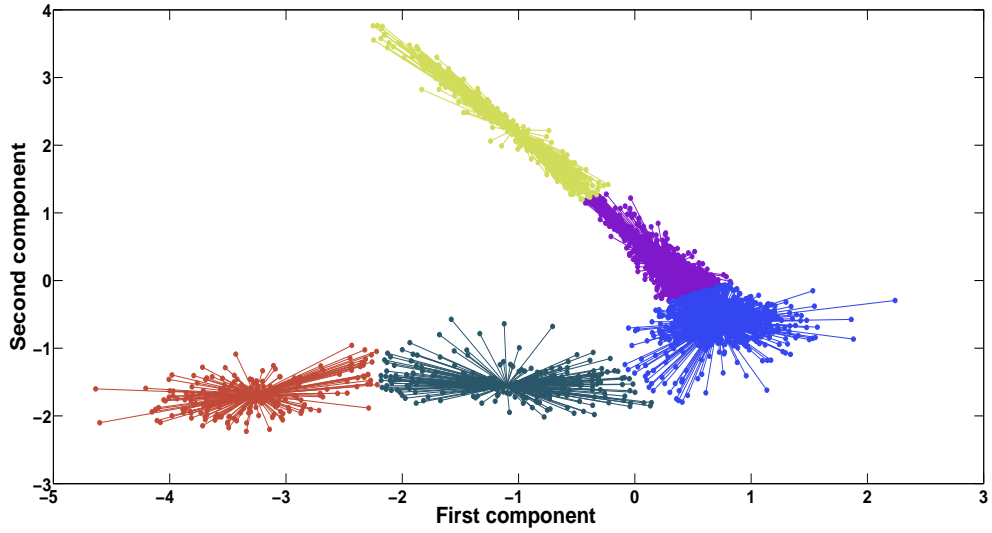


Figure 7: AP algorithm finds 5 different classes on the LLE manifold by limiting the desired number of classes.

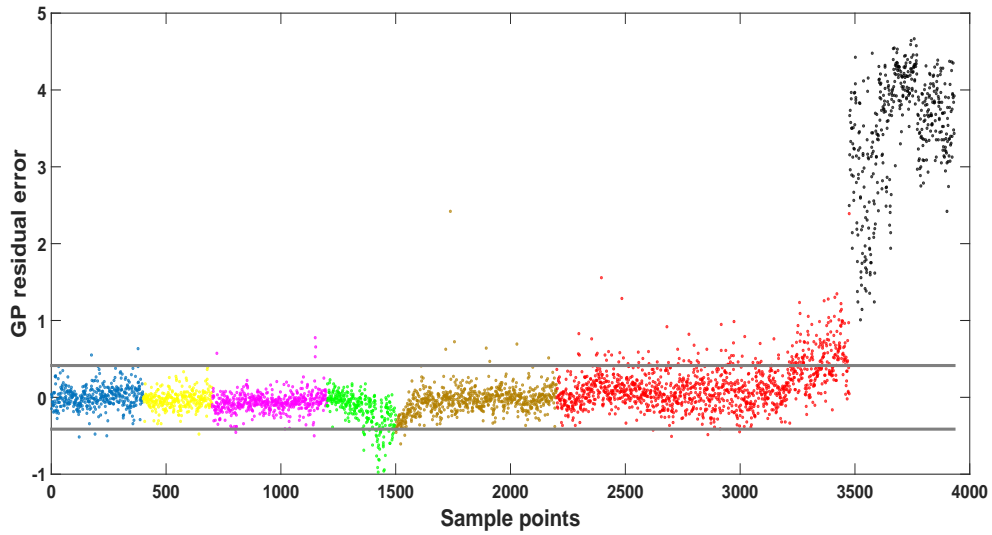


Figure 8: GP prediction error of the 2nd component based on 1st component.

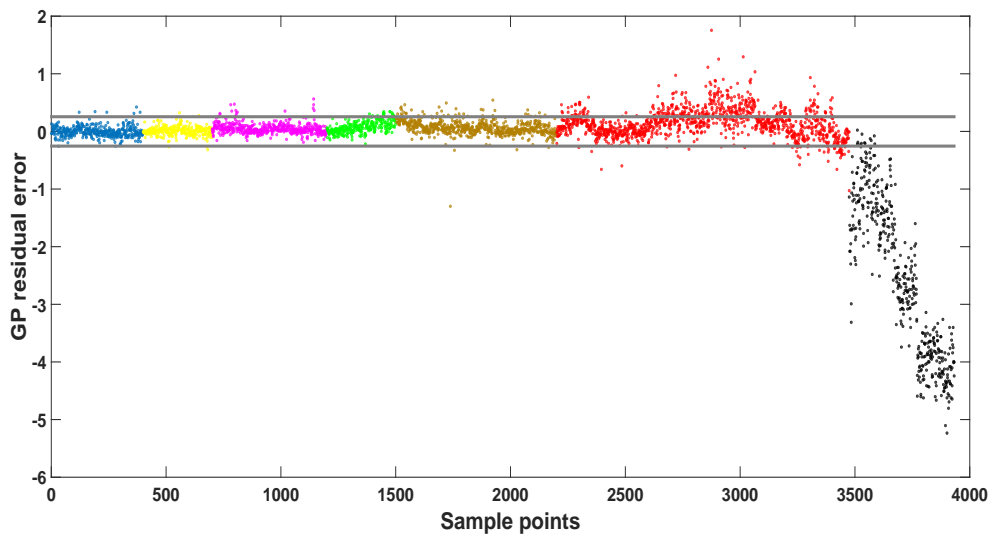


Figure 9: GP prediction error of the 1st component based on 2nd component.

6. Conclusion

The purpose of this paper is to highlight the key utility of some specific machine learning methods, not only for novelty detection analysis but also as a method of investigating the space where data clusters are lying. It also gives a chain of tools for revealing the influence of benign variations like temperature when modal parameters are extracted. The main benefit of the approach taken here is that complicated algebraic analysis is not necessary. Furthermore, in this paper, robust outlier statistics and unsupervised learning techniques are used, focussed mainly on a high level estimation of the “masking effect” of inclusive outliers, not only for determining the presence or absence of novelty - something that is of fundamental interest - but also to examine the normal condition set under the suspicion that it may already include multiple abnormalities.

Acknowledgements

The support of the UK Engineering and Physical Sciences Research Council (EPSRC) through grant reference number EP/J016942/1 and EP/K003836/2 is gratefully acknowledged.

References

- [1] E. J. Cross, On structural health monitoring in changing environmental and operational conditions, Ph.D. thesis (University of Sheffield, 2012).
- [2] E. J. Cross, K. Y. Koo, J. M. W. Brownjohn, K. Worden, Long-term monitoring and data analysis of the Tamar bridge, *Mechanical Systems and Signal Processing* 35 (1) (2013) 16–34.

- [3] B. Peeters, J. Maeck, G. De Roeck, Vibration-based damage detection in civil engineering: excitation sources and temperature effects, *Smart Materials and Structures* 10 (3) (2001) 518.
- [4] S. Alampalli, Effects of testing, analysis, damage, and environment on modal parameters, *Mechanical Systems and Signal Processing* 14 (1) (2000) 63–74.
- [5] P. Cornwell, C. R. Farrar, S. W. Doebling, H. Sohn, Environmental variability of modal properties, *Experimental Techniques* 23 (6) (1999) 45–48.
- [6] N. Dervilis, M. Choi, S. Taylor, R. Barthorpe, G. Park, C. Farrar, K. Worden, On damage diagnosis for a wind turbine blade using pattern recognition, *Journal of Sound and Vibration* 333 (6) (2014) 1833–1850.
- [7] E. J. Cross, K. Worden, Q. Chen, Cointegration: a novel approach for the removal of environmental trends in structural health monitoring data, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 467 (2133) (2011) 2712–2732.
- [8] E. Figueiredo, G. Park, C. R. Farrar, K. Worden, J. Figueiras, Machine learning algorithms for damage detection under operational and environmental variability, *Structural Health Monitoring* 10 (6) (2011) 559–572.
- [9] E. Cross, G. Manson, K. Worden, S. Pierce, Features for damage detection with insensitivity to environmental and operational variations, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 468 (2148) (2012) 4098–4122.

- [10] T.-H. Yi, H.-N. Li, H.-M. Sun, Multi-stage structural damage diagnosis method based on "energy-damage" theory, *Smart Structures and Systems* 12 (3-4) (2013) 345–361.
- [11] K. Worden, E. Cross, I. Antoniadou, A. Kyprianou, A multiresolution approach to cointegration for enhanced shm of structures under varying conditions—an exploratory study, *Mechanical Systems and Signal Processing* 47 (1) (2014) 243–262.
- [12] W. Zheng, W. Yu, Probabilistic approach to assessing scoured bridge performance and associated uncertainties based on vibration measurements, *Journal of Bridge Engineering*.
- [13] E. Reynders, G. Wursten, G. De Roeck, Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification, *Structural Health Monitoring* 13 (1) (2014) 82–93.
- [14] A. M. Leroy, P. J. Rousseeuw, *Robust regression and outlier detection*, Wiley Series in Probability and Mathematical Statistics, New York: Wiley, 1987 1.
- [15] P. J. Rousseeuw, B. C. Van Zomeren, Unmasking multivariate outliers and leverage points, *Journal of the American Statistical Association* 85 (411) (1990) 633–639.
- [16] N. Dervilis, E. J. Cross, R. J. Barthorpe, K. Worden, Robust methods of inclusive outlier analysis for structural health monitoring, *Journal of Sound and Vibration* (2014) 333 (20) (2014) 5181 – 5195.
- [17] N. Dervilis, K. Worden, E. Cross, On robust regression analysis as a means of exploring environmental and operational conditions for SHM data, *Journal of Sound and Vibration* 347 (0) (2015) 279 – 296.

- [18] G. D. Roeck, The state-of-the-art of damage detection by vibration monitoring: the SIMCES experience, *Journal of Structural Control* 10 (2) (2003) 127–134.
- [19] P. J. Rousseeuw, K. Van Driessen, A fast algorithm for the minimum covariance determinant estimator, *Technometrics* 41 (3) (1999) 212–223.
- [20] S. Verboven, M. Hubert, LIBRA: a Matlab library for robust analysis, *Chemometrics and Intelligent Laboratory Systems* 75 (2) (2005) 127–136.
- [21] M. Hubert, P. J. Rousseeuw, S. Van Aelst, High-breakdown robust multivariate methods, *Statistical Science* (2008) 92–119.
- [22] D. A. Jackson, Y. Chen, Robust principal component analysis and outlier detection with ecological data, *Environmetrics* 15 (2) (2004) 129–139.
- [23] B. J. Frey, D. Dueck, Clustering by passing messages between data points, *science* 315 (5814) (2007) 972–976.
- [24] I. E. Givoni, B. J. Frey, A binary variable model for affinity propagation, *Neural computation* 21 (6) (2009) 1589–1600.
- [25] C. E. Rasmussen, C. Williams, *Gaussian processes for machine learning*. 2006, The MIT Press, Cambridge, MA, USA.
- [26] L. K. Saul, S. T. Roweis, An introduction to locally linear embedding, unpublished. Available at: <http://www.cs.toronto.edu/~roweis/lle/publications.html>.
- [27] S. T. Roweis, L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, *Science* 290 (5500) (2000) 2323–2326.

- [28] C. M. Bishop, et al., Neural Networks for Pattern Recognition, Clarendon press Oxford, 1995.
- [29] I. T. Nabney, NETLAB: Algorithms for Pattern Recognition, Springer, 2004.

Appendix A

7. Robust MCD index

The application of robust computation of location and covariance estimation for multivariate data is of significant interest in the investigation of inclusive outliers. The method discussed here is called the Minimum Covariance Determinant estimator (MCD). The computation of the MCD estimator is not a trivial procedure and requires an extensive calculation. In the current study, the FAST-MCD algorithm is implemented [14–16, 19–22]. The algorithm is given in detail in the references [14–16, 19–22], and the code was provided via a statistical Matlab library called LIBRA [20]. A brief description of the algorithm is provided by explaining the basic steps of the FAST-MCD technique.

A multivariate data matrix $[X] = (\{x_1\}, \dots, \{x_m\})^T$ is assumed of m points in n -dimensional space ($n \times m$) where $\{x_i\} = (x_{i1}, \dots, x_{in})^T$ is an observation. Robust estimates of the centre μ and the scatter matrix σ of X can be calculated by the MCD estimator. The MCD tool looks for the $h(> \frac{m}{2})$ observations out of m whose classical covariance matrix has the lowest possible determinant. The raw MCD estimate of location (arithmetic mean) is then computed from the average of these h points and the raw MCD estimation of scatter is the covariance matrix multiplied by a consistency factor.

The calculation of the lowest determinant is critical, as one moves from one approximation of MCD to another one with lower determinant. This tool and the proof that follows it are not obvious and can be found in the appendix of [19].

Based on the raw MCD estimates, a reweighting step can be added in order to increase the finite sampling efficiency. The advantage is that MCD estimates can resist up to $(m - h)$ outliers and in turn, the number h (or equally $a = \frac{h}{m}$) controls the robustness of the estimator. The highest resistance compared to contamination is achieved by calculating $h = \frac{(n+m+1)}{2}$. It is proposed that when a large proportion of contamination is assumed then $h = an$ with $a = 0.5$. Detecting outliers can be challenging when m/n is small because some data points can become coplanar. This is a general problem in the machine learning community called the “curse of dimensionality”. It is recommended [20] that when $\frac{m}{n} > 5$, a should be 0.5. Generally, the MCD estimates of location and scatter are affine equivariant which means that they are invariant under affine transformation behaviour. This is crucial as the underlying model is then immune to different variable scales and data rotations. Rousseeuw and Van Driessen (1999) [19] developed the FAST-MCD algorithm based on a Concentration step (C-step). C-steps select the h observations with the smallest distances and the scatter matrix with the lower determinant [19] and the main details are given.

Appendix B

8. Gaussian Process Regression algorithm

Rasmussen and Williams [25] define a Gaussian process (GP) as “a collection of random variables, any finite number of which have a joint Gaussian distribu-

tion”. In recent years GPs are gaining a lot of attention as a machine learning approach in the area of regression (or classification) analysis as they offer fast and simple computations. Gaussian process regression is a robust tool which takes into account all possible functions that fit to the training data set and gives a predictive distribution of a single prediction for a given input vector. As result, a mean prediction and confidence intervals on this prediction can be calculated from this predictive distribution. The basic details of the algorithm are presented following the steps in [25]. The algorithm that was used in the previous Sections is also coming from Rasmussen and Williams [25].

8.1. Algorithm theory

The initial and basic step in order to apply Gaussian process regression is to obtain a mean $m(\{x\})$ and covariance function $k(\{x\}, \{x'\})$ as GPs are completely specified by them, $\{x\}$ represents the input vector. So for any real process $f(\{x\})$ one can define:

$$m(\{x\}) = E[f(\{x\})] \tag{3}$$

$$k(\{x\}, \{x'\}) = E[(f(\{x\}) - m(\{x\}))(f(\{x'\}) - m(\{x'\}))] \tag{4}$$

where E represents the expectation. Often, for practical reasons because of notation purposes (simplicity) and little knowledge about the data at the initial stage the prior mean function is set to zero. The Gaussian processes can then be defined as:

$$f(\{x\}) \sim GP(0, k(\{x\}, \{x'\})) \quad (5)$$

Assuming a zero-mean function, the covariance function can be described as:

$$\begin{aligned} \text{cov}(f(\{x\}_p), f(\{x\}_q)) &= k(\{x\}_p, \{x\}_q) \\ &= \exp\left(-\frac{1}{2}|\{x\}_p - \{x\}_q|^2 - \frac{1}{2}|\{x\}_p - \{x\}_q|^2\right) \end{aligned} \quad (6)$$

This is the squared-exponential covariance function (although not the only option). It is very important to mention an advantage of the previous equation as the covariance is written as a function only of the inputs. For the squared-exponential covariance, it can be noted that it takes unit values between variables where their inputs are very close and starts to decrease as the variable distance in the input space increases.

Assuming now that one has a set of training outputs $\{f\}$ and a set of test outputs $\{f\}_*$ one has the prior:

$$\begin{bmatrix} \{f\} \\ \{f\}_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(X, X) & K(X, X_*) \\ K(X_*, X) & K(X_*, X_*) \end{bmatrix}\right) \quad (7)$$

where the capital letters represent matrices. As can be seen, the covariance matrix must be symmetrical about the main diagonal.

As the prior has been generated by the mean and covariance functions, in order to specify the posterior distribution over the functions, one needs to limit the prior distribution in a such a way that includes only these functions that agree with actual data points. An obvious way to do that is by generating

functions from the prior and select only the ones that agree with the actual points. Of course, this is not a realistic way of doing it as it would consume a lot of computational power. In a probabilistic manner this can be done easily via conditioning the joint prior on the observations and this will give (for more details see [25, 28, 29]):

$$\begin{aligned} & \{f\}_* | [X]_*, [X], \{f\} \\ & \sim N \left(\begin{array}{c} K([X_*], [X])K([X], [X])^{-1}\{f\}, K([X_*], [X_*]) \\ -K([X_*], [X])K([X], [X])^{-1}K([X], [X_*]) \end{array} \right) \end{aligned} \quad (8)$$

Function values $\{f\}_*$ can be generated by sampling from the joint posterior distribution and at the same time evaluating the mean and covariance matrices from (8).

The covariance functions used in this study are usually controlled by some parameters in order to obtain a better control over the types of functions that are considered for the inference. As an example, the squared-exponential covariance function can take the form (1-dimensional):

$$k_y(k_p, k_q) = \sigma_f^2 \exp\left(-\frac{1}{2l^2}(k_p - k_q)^2\right) + \sigma_n^2 \delta_{pq} \quad (9)$$

where k_y is the covariance for the noise target set y . The length-scale l (determines how far one needs to move in input space for the function values to become uncorrelated), the variance σ_f^2 of the signal and the noise variance σ_n^2 are free parameters that can be varied. These free parameters are called *hyperparameters*.

The tool that has to be applied for selecting the model for choosing the optimal

hyperparameters for GP regression, is the maximum marginal likelihood of the predictions $p(\{y\}||[X], \{\theta\})$ with respect to the hyperparameters θ :

$$\begin{aligned} \log p(\{y\}||[X], \{\theta\}) = \\ -\frac{1}{2}\{y\}^T [K_y]^{-1} \{y\} - \frac{1}{2} \log |[K_y]| - \frac{n}{2} \log 2\pi \end{aligned} \quad (10)$$

where $[K_y] = [K_f] + \sigma_n^2 I$ is the covariance matrix of the noise test set $\{y\}$ and $[K_f]$ is the noise noise-free covariance matrix. In order to optimise these hyperparameters through maximising the marginal log likelihood the partial derivatives give the solution, via gradient descent:

$$\begin{aligned} \frac{\partial}{\partial \theta_j} \log p(\{y\}||[X], \{\theta\}) \\ = \frac{1}{2} \{y\}^T [K]^{-1} \frac{\partial [K]}{\partial \theta_j} [K]^{-1} \{y\} - \frac{1}{2} \text{tr} \left([K]^{-1} \frac{\partial [K]}{\partial \theta_j} \right) \\ = \frac{1}{2} \text{tr} \left((\alpha \alpha^T - [K]^{-1}) \frac{\partial [K]}{\partial \theta_j} \right) \end{aligned} \quad (11)$$

where $\alpha = [K]^{-1} \{y\}$. Of course this solution is not a trivial procedure and for specific details readers are referred to [25].