FEDERAL UNIVERSITY OF PARÁ

INSTITUTE OF TECHNOLOGY

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

MACHINE LEARNING ALGORITHMS FOR DAMAGE DETECTION IN STRUCTURES UNDER CHANGING NORMAL CONDITIONS

MOISÉS FELIPE MELLO DA SILVA

DM 07/2017

UFPA / ITEC / PPGEE

Guamá University Campus

Belém-Pará-Brazil

2017

FEDERAL UNIVERSITY OF PARÁ

INSTITUTE OF TECHNOLOGY

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

MOISÉS FELIPE MELLO DA SILVA

MACHINE LEARNING ALGORITHMS FOR DAMAGE DETECTION IN STRUCTURES UNDER CHANGING NORMAL CONDITIONS

DM 07/2017

UFPA / ITEC / PPGEE

Guamá University Campus

Belém-Pará-Brazil

2017

FEDERAL UNIVERSITY OF PARÁ

INSTITUTE OF TECHNOLOGY

GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

MOISÉS FELIPE MELLO DA SILVA

MACHINE LEARNING ALGORITHMS FOR DAMAGE DETECTION IN STRUCTURES UNDER CHANGING NORMAL CONDITIONS

Master dissertation submitted to the Examining Board of the Graduate Program in Electrical Engineering from the Federal University of Pará to obtain the Master Degree in Electrical Engineering, Area of Concentration in Applied Computing.

UFPA / ITEC / PPGEE

Guamá University Campus

Belém-Pará-Brazil

2017

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da UFPA

Silva, Moisés Felipe Mello da, 1994-
Machine learning algorithms for damage detection in structures under changing normal conditions / Moisés Felipe Mello da Silva. - 2017.

Orientador: João Crisóstomo Weyl Albuquerque Costa;
Coorientador: Claudomiro de Souza de Sales Junior.

Dissertação (Mestrado) – Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém, 2017.

1. Inteligência artificial. 2. Reconhecimento de padrões. 3. Engenharia de estruturas – modelos matemáticos. I. Título.

CDD 23. Ed 006.3

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

MACHINE LEARNING ALGORITHMS FOR DAMAGE DETECTION IN STRUCTURES UNDER CHANGING NORMAL CONDITIONS

AUTOR: **MOISÉS FELIPE MELLO DA SILVA**

DISSERTAÇÃO DE MESTRADO SUBMETIDA À AVALIAÇÃO DA BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA DA UNIVERSIDADE FEDERAL DO PARÁ, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE COMPUTAÇÃO APLICADA.

APROVADO EM: 31 / 01 / 2017

BANCA EXAMINADORA:

Prof. Dr. João Crisóstomo Weyl Albuquerque Costa
(Orientador - PPGEE/UFPA)

Prof. Dr. Claudomiro de Souza de Sales Junior
(Coorientador - PPGCC/UFPA)

Prof. Dr. Diego Lisboa Cardoso
(Examinador Interno - ITEC/UFPA)

Prof. Dr. Eloi João Faria Figueiredo
(Examinador Externo - ULHT)

VISTO:

Prof. Dr. Evaldo Gonçalves Pelaes
(Coordenador do PPGEE/ITEC/UFPA)

Acknowledgements

I would like to dedicate this dissertation and my entire life's work to my mother, Cátia Regina Mello da Silva. She inspired me with superb dedication and gentle care. You'll be proud of me, wherever you are. I promise you that.

I thank to my father, José Carlos Lima da Silva, for the unconditional support provided me. I live to make you proud.

For my aunt, Lúcia Guerreiro, hugs and kisses. You're the one who can get close to fill the gap left by mommy. Love you.

To the professor Claudomiro de Souza de Sales Junior who accompany and drive my career since I got in the university, a thank you. Also, thank to professor João Crisóstomo Weyl Albuquerque Costa for friendship, companionship, experience, guided and knowledge provided in LEA.

In special to my friends Adam Santos, Cindy Fernandes and Reginaldo Santos, whose friendship, I hope, can carry on for the life.

Thanks to my best friend, Thiago Araújo, who is closer than a brother. May our partnership last for life.

And, for you, Luena Canavieira, my dear, nothing that I could say is enough to represent your importance to me. In the absence of words, I prefer the simple ones: thank you, Lulu.

My acknowledgements for the financial support received from CAPES[1], CNPQ[2] and INESC P&D Brasil[3].

[1] `http://www.capes.gov.br/`
[2] `http://www.cnpq.br/`
[3] `http://www.inescbrasil.org.br/`

In memory of Cátia Silva, my beloved mother
You left fingerprints of grace on my life
You shan't be forgotten
Amare et sapere vix deo conceditu

"What man is a man who does not make the world better?"
Balian d'Ibelin.

"You can see a mountain as one of two ways: as an insurmountable barrier or as manner to grow further."
Unknown.

"The characteristic of genuine heroism is persistency. All men have wandering impulses, fits and starts of generosity and brilliance. But when you have resolved to be great, abide by yourself, and do not weakly try to reconcile yourself with the world. Cause the heroic cannot be the common, nor the common the heroic."
Ralph Waldo Emerson.

Summary

List of Figures

List of Tables

List of abbreviations and acronym

AANN	Auto-Associative Neural Network
ACH	Agglomerative Concentric Hypersphere
AIC	Akaike Information Criterion
BC	Baseline Condition
BIC	Bayesian Information Criterion
DC	Damage Condition
DI	Damage Indicator
DSA	Deep Stacked Autoencoder
EM	Expectation-Maximization
FA	Factor Analysis
GK	Gustafson-Kessel
GMM	Gaussian Mixture Model
KPCA	Kernel Principal Component Analysis
LogL	Log Likelihood
ML	Maximum Likelihood
MSD	Mahalanobis-Squared Distance
NLPCA	Nonlinear Principal Component Analysis
PCA	Principal Component Analysis
RBF	Radial Basis Function
SHM	Structural Health Monitoring
SMS	Structural Management System
SPR	Statistical Pattern Recognition
SVM	Support Vector Machine

Abstract

Engineering structures have played an important role into societies across the years. A suitable management of such structures requires automated structural health monitoring (SHM) approaches to derive the actual condition of the system. Unfortunately, normal variations in structure dynamics, caused by operational and environmental conditions, can mask the existence of damage. In SHM, data normalization is referred as the process of filtering normal effects to provide a proper evaluation of structural health condition. In this context, the approaches based on principal component analysis and clustering have been successfully employed to model the normal condition, even when severe effects of varying factors impose difficulties to the damage detection. However, these traditional approaches imposes serious limitations to deployment in real-world monitoring campaigns, mainly due to the constraints related to data distribution and model parameters, as well as data normalization problems. This work aims to apply deep neural networks and propose a novel agglomerative cluster-based approach for data normalization and damage detection in an effort to overcome the limitations imposed by traditional methods. Regarding deep networks, the employment of new training algorithms provide models with high generalization capabilities, able to learn, at same time, linear and nonlinear influences. On the other hand, the novel cluster-based approach does not require any input parameter, as well as none data distribution assumptions are made, allowing its enforcement on a wide range of applications. The superiority of the proposed approaches over state-of-the-art ones is attested on standard data sets from monitoring systems installed on two bridges: the Z-24 Bridge and the Tamar Bridge. Both techniques revealed to have better data normalization and classification performance than the alternative ones in terms of false-positive and false-negative indications of damage, suggesting their applicability for real-world structural health monitoring scenarios.

Keywords: Structural health monitoring, Damage detection, Deep learning, Clustering, Operational conditions, Environmental conditions

Resumo

Estruturas de engenharia têm desempenhado um papel importante para o desenvolvimento das sociedades no decorrer dos anos. A adequada gerência e manutenção de tais estruturas requer abordagens automatizadas para o monitoramento de integridade estrutural (SHM) no intuito de analisar a real condição dessas estruturas. Infelizmente, variações normais na dinâmica estrutural, causadas por efeitos operacionais e ambientais, podem ocultar a existência de um dano. Em SHM, normalização de dados é frequentemente referido como o processo de filtragem dos efeitos normais com objetivo de permitir uma avaliação adequada da integridade estrutural. Neste contexto, as abordagens baseadas em análise de componentes principais e agrupamento de dados têm sido empregadas com sucesso na modelagem dessas condições variadas, ainda que efeitos normais severos imponham alto grau de dificuldade para a detecção de danos. Contudo, essas abordagens tradicionais possuem limitações sérias quanto ao seu emprego em campanhas reais de monitoramento, principalmente devido as restrições existentes quanto a distribuição dos dados e a definição de parâmetros, bem como os diversos problemas relacionados a normalização dos efeitos normais. Este trabalho objetiva aplicar redes neurais de aprendizado profundo e propor um novo método de agrupamento aglomerativo para a normalização de dados e detecção de danos com o objetivo de superar as limitações impostas pelos métodos tradicionais. No contexto das redes neurais profundas, o emprego de novos métodos de treinamento permite alcançar modelos com maior poder de generalização. Em contrapartida, o novo algoritmo de agrupamento não requer qualquer parâmetro de entrada e não realiza asserções quanto a distribuição dos dados, permitindo um amplo dominínio de aplicações. A superioridade das abordagens propostas sobre as disponíveis na literatura é atestada utilizando conjuntos de dados oriundos de dois sistemas de monitoramento instalados em duas pontes distintas: a ponte Z-24 e a ponte Tamar. Ambas as técnicas revelaram um melhor desempenho de normalização dos dados e classificação do que os métodos tradicionais, em termos de falsas-positivas e falsas-negativas indicações de dano, o que sugere a aplicabilidade dos métodos em cenários reais de monitoramento de integridade estrutural.

Palavras-chave: Monitoramento de integridade estrutural, Detecção de danos, Aprendizado profundo, Agrupamento, Condições ambientais, Condições operacionais.

1 Introduction

1.1 Context

Improved and more continuous condition assessment of civil structures has been demanded by our society to better face the challenges presented by aging civil infrastructure. Structural management systems (SMSs) plan to cover all activities performed during the service life of engineering structures, considering public safety, authorities' budgetary constraints, and transport network functionality. They possess mechanisms to ensure the structures are regularly inspected, evaluated, and maintained in a proper manner. Hence, a SMS is developed to analyze engineering and economic factors and to attend the authorities in determining how and when to make decisions regarding maintenance, repair, and rehabilitation of structures (FARRAR; WORDEN, 2013; FIGUEIREDO; MOLDOVAN; MARQUES, 2013).

However, the SMSs still depend on structural inspections, especially on the qualitative and not necessarily consistent visual inspections, which may impact the structural evaluation and, consequently, the maintenance decisions as well as the avoidance of structural collapses (WENZEL, 2009). In the last years, the structural health monitoring (SHM) discipline has emerged to aid the structural management with more reliable and quantitative information. Although the SMS has already been accepted by the structural managers around the world (CATBAS; GOKCE; GUL, 2012; WORDEN et al., 2015; HU et al., 2016), even though with inherent limitations imposed by the visual inspections, the SHM is becoming increasingly attractive due to its potential ability to detect damage at varying stages and near real-time, with the consequent life-safety and economical benefits (WORDEN et al., 2007; FARRAR; WORDEN, 2013).

The process involves the observation of a structural system over time using periodically sampled response measurements from an array of sensors, the extraction of damage-sensitive features from these measurements, and the statistical analysis of these features to discriminate the actual structural condition for short or long-time periods. Then, once the normal condition has been successfully learned, the model can be used for rapid condition assessment to provide, in nearly real time, reliable information regarding the integrity of the structure.

The author believe that all approaches to SHM, as well as all traditional nondestructive evaluation techniques, can be posed in the context of a statistical pattern

recognition (SPR) problem. Thus, the SPR paradigm for the development of SHM solutions can be described as a four-phase process (FARRAR; DOEBLING; NIX, 2001): (1) operational evaluation, (2) data acquisition, (3) feature extraction, and (4) statistical modeling for feature classification.

Particularly, in the feature extraction phase, damage-sensitive features (e.g., natural frequencies, autorregresive model parameters) are derived from the raw data, being correlated with the severity of damage present in the monitored structure. Data compression is also an inherent part of most feature extraction procedures. Unfortunately, operational and environmental variations (e.g. temperature, operational loading, humidity and wind speed) often arise as undesired effects in the damage-sensitive features and usually mask changes caused by damage, which might negatively influence the proper identification of damage (SOHN, 2007).

In that regard, data normalization procedures are required to surpass the effects of operational and environmental variability, as an effort to improve the damage assessment (CATBAS; GOKCE; GUL, 2012). This procedure is fully connected to the data acquisition, feature extraction, and statistical modeling phases of the SHM process, including a wide range of steps for mitigating (or even removing) the effects of normal variations on the extracted features as well as for separating changes in damage-sensitive features caused by damage from those caused by varying operational and environmental conditions (SOHN; WORDEN; FARRAR, 2002; KULLAA, 2011). Without such data normalization procedures, varying operational and environmental conditions will produce false-positive indications of damage and quickly erode confidence in the SHM system. In general, the treatment of such influences starts in data collection (by choosing less sensitive physical parameters to varying normal condition), appears in feature extraction (by selection of features with high sensitivity to damage and insensitive to normal variations), and finishes in the statistical modeling (remaining effects are accounted by automated procedures inspired in machine learning field) (FARRAR; SOHN; WORDEN, 2001).

Therefore, for statistical modeling phase, several machine learning algorithms with different working principles have been proposed (WORDEN; MANSON, 2007; FIGUEIREDO et al., 2011). These machine learning approaches are often characterized as unsupervised and output-only because they are trained only with damage-sensitive features related to undamaged condition without any measurement directly related to operational and environmental parameters. One of the reasons for this choice is the limited applicability of the supervised learning, which carries out the training phase using data from both conditions (undamaged and damaged), and the input-output approaches should know in advance all parameters to be measured, respectively (PEETERS; ROECK, 2001; ZHOU et al., 2008; KULLAA, 2009; HOSSEINABADI et al., 2014; VAN; KANG, 2015; D'ANGELO; RAMPONE, 2016).

The most traditional unsupervised approaches used in SHM field are, no doubt, the ones based on Mahalanobis squared distance (MSD) and principal component analysis (PCA) (WORDEN; MANSON; FIELLER, 2000; MALHI; GAO, 2004; YAN et al., 2005c; PURARJOMANDLANGRUDI; GHAPANCHI; ESMALIFALAK, 2014; XIANG; ZHONG; GAO, 2015). They are linear algorithms adapted to act as data normalization and damage detection techniques used to model, mainly, effects of linear variations. However, the linear behavior imposed for these techniques has limited their applicability in SHM. If nonlinearities are present in the monitoring data, the MSD and PCA might fail in modeling the normal condition of a structure because the former assumes the baseline data follow a multivariate Gaussian distribution (or only one data cluster) and the principal components in the latter are independent only if the baseline data is jointly normally distributed.

To extend the capabilities of the traditional methods, improved approaches based on the auto-associative neural network (AANN), kernel PCA and Gaussian mixture models (GMMs) were proposed to deal with real-world structures and more complex SHM applications such that the nonlinear influences on the damage-sensitive features could be accounted for (HSU; LOH, 2010; MALHI; YAN; GAO, 2011; SHAO et al., ; REYNDERS; WURSTEN; ROECK, 2014; SANTOS et al., 2015; FIGUEIREDO; CROSS, 2013). However, the required input parameters, as well as the usual constraints related to data distribution, make these approaches hard to employ in real-world monitoring campaigns. In some cases, despite of their model complexity and high computational cost, the training procedures do not guarantees a proper modeling of normal conditions, resulting in poor damage detection performance.

1.2 Related work

Traditionally, in most civil applications, the damage detection process is carried out using physics-based methods (rooted in the structure dyncamics) and parametric approaches. However, in high complex structures, those methods may be impractical due to the level of expertise and time required to their development (LAORY; TRINH; SMITH, 2011; WORDEN et al., 2015). On the other hand, non-parametric approaches rooted in the machine learning field become an alternative, as they are very useful to find hidden patterns in the data and computationally efficient (CATBAS; GOKCE; GUL, 2012; HU et al., 2016). Herein, machine learning-based approaches addressing damage assessment are discussed. In particular, the progressive development of approaches that address PCA-based models is described. Moreover, the most relevant cluster-based methods and their adaptation to damage detection in SHM are also introduced. From the current state-of-the-art the main objectives are subsequently defined.

1.2.1 Traditional approaches for damage detection

Principal component analysis is a common method to perform data normalization and feature classification without directly measure the sources of variability. Yan et al. (YAN et al., 2005a) present a PCA-based approach to model linear environmental and operational influences using only undamaged feature vectors. The number of principal components of the vibration features is implicitly assumed to correspond to the number of independent factors related to the normal variations. A further extension of the proposed method is presented in (YAN et al., 2005b). In this case, a local extension of PCA is used to learn nonlinear relationships by applying a local piecewise PCA in a few regions of the feature space. Although both approaches demonstrate adequate damage detection performance, the use of PCA imposes serious limitations in the employment of the proposed approaches, such as: only linear transformations can be performed through the orthogonal components; as larger the variance of the component, the greater its importance (in some cases this assumption is untrue); and scale variant (SHLENS, 2002).

To overcome the PCA limitations and detect damage in structures under changing environmental and operational conditions, an output-only vibration-based damage detection approach was proposed by Deraemaeker et al. (DERAEMAEKER et al., 2008). Two types of feature extraction based on automated stochastic subspace identification and Fourier transform are used as damage sensitive-features, as well as the environmental effects and the damage detection are carried out by factor analysis (FA) and a statistical process control, respectively. The results demonstrate that when FA is applied to deal with normal variations both type of features provide reliable damage classification results. However, this approach has been tested only using a numerical model of bridge, which does not ensure its performance in real monitoring scenarios. Furthermore, the FA is also able to learn linear influences as linear PCA.

Auto-associative neural network (AANN) is a nonlinear version of PCA intended to perform feature extraction, dimensionality reduction, and damage detection of multivariate data. As demonstrated by Krammer (KRAMER, 1991), the AANN is capable to perform, intrinsically, a nonlinear PCA (NLPCA), as it characterizes the underlying dependency of the identified features on the unobserved operational and environmental factors. Worden (WORDEN, 1997) developed a novelty detection technique by applying AANN to learn the normal condition of structures. Once the model is trained, the residual error tends to increase when damaged cases are presented. A later study (FIGUEIREDO et al., 2011) applied the AANN to model nonlinearities in a laboratory structure under simulated effects of variability. In this study, the AANN was not able to modeling, properly, the normal condition, which was verified in terms of Type I/II errors (false-positive and false-negative indications of damage). Zhou et al. (ZHOU; NI; KO, 2011) proposed a new damage index to avoid the occurrence of errors. However, the results strongly depends

on the type of features extracted.

The aforementioned drawbacks lead the research efforts to kernel-based machine learning algorithms, which have been widely used in structure monitoring. The ones based on support vector machines (SVM) have demonstrated high reliability and sensitivity to damage. A supervised SVM method to detect damage in structures with a limited number of sensors was proposed in (MITA; HAGIWARA, 2003). Khoa et al. (KHOA et al., 2014) proposed an unsupervised adaptation to dimensionality reduction and damage detection in bridges. Santos et al. (SANTOS et al., 2016b) carried out a comparison study on kernel-based methods. The results demonstrated that those SVM-based approaches has been outperformed by kernel PCA (KPCA), in terms of removing environmental and operational effects and classification performance.

The KPCA is an alternative approach to perform NLPCA. The kernel-trick allows to mapping the feature vectors to high dimensional spaces, which provides nonlinear strengths to linear PCA. Cheng et al. (CHENG et al., 2015) applied KPCA to detect damage on concrete dams subjected to normal variations. Similarly, novelty detection methods were proposed in (OH; SOHN; BAE, 2009; YUQING et al., 2015) by applying KPCA as a data normalization procedure. In these approaches, the problems related to the choice of suitable damage index and estimation of some parameters are addressed. However, the issues related to the choice of an optimal kernel bandwidth and the number of retained components were not fully addressed. Reynders et al. (REYNDERS; WURSTEN; ROECK, 2014) developed an alternative approach to detect damage and eliminate the environmental and operational influences in terms of retained components, and presents a complete scheme to assess the previous issues. However, this approach is not able to completely remove the normal effects, as it deals with a fraction of the environmental and operational effects.

1.2.2 Cluster-based approaches for damage detection

Over the years, the approaches based on the MSD has been widely used in real-world monitoring campaigns due to its ability to identify outliers (WORDEN; MANSON, 2007; NGUYEN; CHAN; THAMBIRATNAM, 2014; ZHOU et al., 2015). The MSD-based approach assumes that the normal condition can be modeled by a unique cluster from a multivariate Gaussian distribution. In this context, an abnormal condition is considered as a statistical deviation from the normal pattern learned during training phase as a mean vector and a covariance matrix, allowing to infer whether the data were generated by a source not related to the normal condition. However, as noticed in (FIGUEIREDO; CROSS, 2013), when nonlinearities are present in the observations, the MSD fails in modeling the normal condition of a structure because it assumes the baseline data as a unique multivariate Gaussian distribution.

A two-step damage detection strategy based on GMMs has been developed in (FIGUEIREDO; CROSS, 2013; FIGUEIREDO et al., 2014b; SANTOS et al., 2016a) and applied to long-term monitoring of bridges. In the first step, the GMM-based approach models the main clusters that correspond to the normal and stable set of undamaged conditions, even when normal variations affect the structural response. To learn the parameters of GMMs, the classical maximum likelihood (ML) estimation based on the expectation-maximization (EM) algorithm is adopted in (FIGUEIREDO; CROSS, 2013). This approach applies an expectation step and a maximization step until the log-likelihood converges to a local optimum. Thus, the convergence to the global optimum is not guaranteed.

To overcome the limitations imposed by EM, in (FIGUEIREDO et al., 2014b) the parameter estimation is carried out using a Bayesian approach based on a Markov-chain Monte Carlo method. In (SANTOS et al., 2016a), a genetic-based approach is employed to drive the EM searching towards the global optimum. In these approaches, as the parameters have been learned, a second step is performed to detect damage on the basis of a MSD outlier formation considering the chosen main groups of clusters. The problems concerning these algorithms is related to the number of required parameters to be tuned, as well as their parametric behavior.

Silva et al. (SILVA et al., 2008) proposed a fuzzy clustering approach to detect damage in an unsupervised manner. The principal component analysis and auto-regressive moving average methods are used to data reduction and feature extraction purposes. The normal condition is modeled by two different fuzzy clustering algorithms, the fuzzy c-means clustering and the Gustafson-Kessel (GK) algorithms. The results demonstrated that the GK algorithm outperforms the alternative approach and reveals a better generalization performance. However, the damage severity is not properly assessed and both approaches output a significant number of false-negative indications of damage.

1.3 Justification

The current state-of-the-art methods are limited to a restrict range of applications, in such manner that linear methods can only be applied to structures under linear variations, as well as the nonlinear ones to structures under nonlinear effects. Some reasons are related to the assumptions of the model made during the training phase that are not fulfilled by the collected data (e.g., assumptions of data normality and dimensionality) due to changes in structural response caused by normal operational and environmental effects. Furthermore, the amount of parameters that require tuning are usually high, and depend on the type of structure and effects altering the data amplitude. Thus, the available approaches do not provide general purpose models, requiring a high level of expertise

and knowledge of the structure's dynamic to decide which method is the most suitable one to each kind of application.

1.4 Motivation

Currently, one of the biggest challenges for transition of SHM technology from research to practice refers to the separation of changes in the feature amplitude caused by damage from those caused by changing operational and environmental conditions. Actually, there are two main approaches to separate those changes. The first implies the direct measure of the sources of variability (e.g., live loads, temperature, wind speed, and/or moisture levels), as well as the structural time-response at different locations of the structure. This input-output approach learns the structural conditions by establishing a direct relation between the normal variations and the actual structural condition. However, such parameterized modeling is hard and complex to deploy in real situations due to the complexity to discover and capture all sources of variability, which still not completely understood (REYNDERS; WURSTEN; ROECK, 2014). The second approach, and the one used in this dissertation, attempts to establish the existence of damage for cases when measurements of the operational and environmental factors that influence the structure's dynamic response are not available or can not be obtained. Thus, the main purpose is to eschew the measure of operational and environmental variations and physics-based models such as finite element analysis. The present work is motivated by the need of robust unsupervised methods to identify damages in structures subjected to linear/nonlinear variations using only the time-response data from undamaged condition.

1.5 Objectives

The main objectives of this work is to review, develop, and apply several machine learning algorithms to data normalization purposes in the context of the SPR paradigm, capable to detect damage on structures under unmeasured operational and environmental factors. Furthermore, the focus of this work is on the implementation of algorithms that analyze and learn the distributions of the extracted damage-sensitive features from the raw data, in an effort to determine the structural health condition. To achieve these goals, the particular objectives are listed below:

1. Overcome the limitations imposed by traditional methods based on PCA and cluster analysis through application and development of novel machine learning algorithms for statistical modeling that address the same principle as the traditional ones (i.e., novel methods to improve the current PCA- and cluster-based approaches).

2. Enhance the literature of SHM concerned to damage assessment and identification by comparing the proposed methods with traditional ones.

3. Apply the proposed methods on data sets from real-world structures subjected to rigorous linear/nonlinear effects, as a means of testing their performance and establish comparisons.

Note that even though these procedures might be applied to infrastructure of arbitrary complexity (such as mechanical, aeronautical and naval structures), in this case the procedures are specially posed in the context of bridge applications.

1.6 Original contributions

Many research works have been proposed to improve the damage detection and identification. However, some issues still not completely addressed and treated. In that regard, the main original contributions of this work are the following:

1. Fulfil the gap in literature for robust methods able to remove linear/nonlinear variations for damage detection purposes by proposing two non-parametric algorithms based on deep neural networks and agglomerative clustering, highlighting the fact that the cluster-based method was the one totally designed by the author.

2. The deep neural network can be faced as the first application of deep stacked autoencoders (DSA) as a data normalization procedure in the context of SHM, which it is intended to overcome the limitations related to the traditional PCA-based approaches.

3. The novel agglomerative concentric hypersphere (ACH) algorithm aims to model the normal conditions of a structure by clustering similar observations related to the same structural state at a given period of time. This straightforward method does not require any input parameter, except the training data matrix, as well as none assumptions related to data distribution are made.

4. Two deterministic initialization procedures rooted on eigenvectors/eigenvalues decomposition and an uniform data sampling are presented. Furthermore, a random initialization is also introduced. These mechanisms are intended to support the initial guess for the clustering procedure.

5. Compare the proposed approaches with state-of-the-art ones on the basis of Type I/Type II error (false-positive and false-negative indications of damage, respectively) trade-off in two real-world standard data sets from Z-24 Bridge, Switzerland, and Tamar Bridge, United Kingdom.

The results indicate that the proposed approaches has overcome the traditional ones in terms of damage classification and robustness to deal with nonlinear effects caused by normal variations. Furthermore, the proposed cluster-based algorithm demonstrates to be capable to provide physical interpretations about structural conditions, allowing a better understanding of operational and environmental sources of variability.

1.7 Organization of dissertation

The remainder of this dissertation is organized as follows. Chapter 2 summarizes the SPR paradigm for SHM and presents the main PCA- and cluster-based algorithms for data normalization, as well as their corresponding damage detection strategies. Chapter 3 derives the improved proposed approaches for increasing the performance of traditional methods. Chapter 4 is devoted to describe the data sets used as damage-sensitive features from the Z-24 Bridge and Tamar Bridge, as well as to present the experimental results and carry out discussions. Finally, Chapter 5 summarizes the main conclusions and contributions of this dissertation. Additionally, future research topics and published works are carried out.

2 Statistical pattern recognition for structural health monitoring

The SHM research community agree that all approaches to SHM can be seen in the context of a pattern recognition problem, which aims to provide not only reasonable answers for all possible inputs, but also infer explanation and formalization of the relationships deriving that answers. Thus, the SPR paradigm for the development of SHM applications is usually described as a four phase process, as illustrated in Figure 1 (FIGUEIREDO, 2010; FARRAR; WORDEN, 2013). These phases are briefly described below, giving more attention to the fourth phase, which is the main focus of this work.

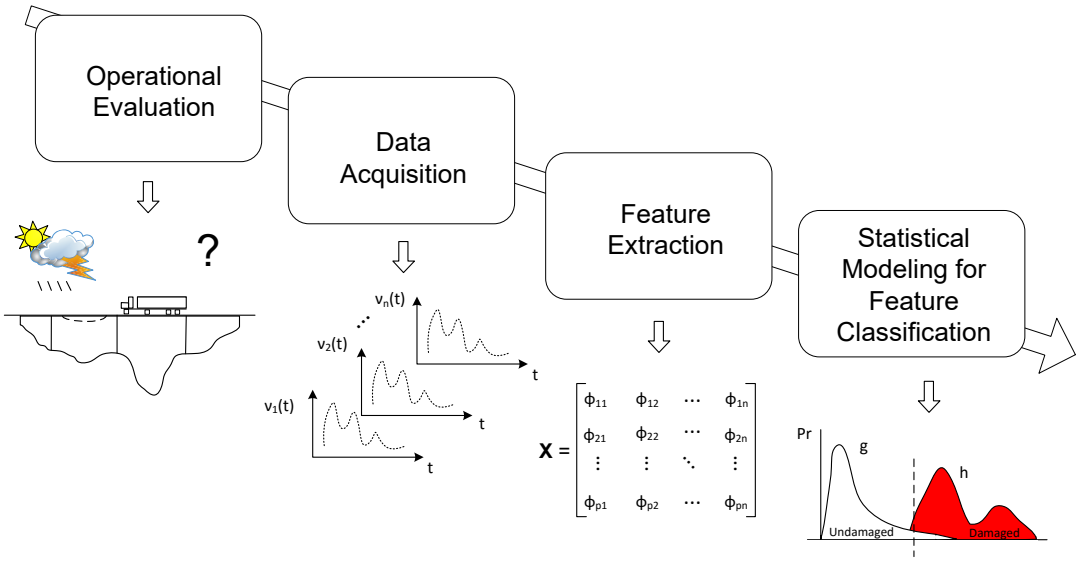

Figure 1 – Flowchart of SPR paradigm for SHM (FIGUEIREDO, 2010).

2.1 Operational evaluation

The main goal of operational evaluation phase it provide answers to four questions regarding the implementation of a monitoring system:

1. What is the life-safety and/or economic justification for performing the structural health monitoring?
2. How is damage defined for the system being investigated and, for multiple damage possibilities, which cases are of the most concern?

3. What are the conditions, both operational and environmental, under which the system to be monitored functions?

4. What are the limitations on acquiring data in the operational environment?

In that regard, operational evaluation seeks to set the boundaries and limitations on the kind of monitored parameters, aiming to determine how the monitoring will be accomplished. It enables to reduce time and cost efforts during posterior phases, allowing to determine the appropriate features to be extracted from the system being monitored and attempts to exploit unique features of the damage that is to be detected. Otherwise, later phases would be carried out without any reliability in the monitoring system designed.

2.2 Data acquisition

The data acquisition portion of the SHM process involves selecting the excitation methods, the sensor types, number and locations, as well as the data acquisition/storage/transmittal hardware. This portion of the process will be application-specific. Economic factors play the major role during acquisition of the hardware to be used for the SHM system. The sensor sensitivity to low level excitation, the data interrogation procedures, as well as the interval at which data should collected are other issues that must be addressed. For example, in applications where life-safety is a critical effort, such as earthquake monitoring, it may be prudent to collect data immediately before and at periodic intervals after a large event. On the other hand, if identify slightly changes in stiffness and geometric properties is the main concern, then it may be necessary to collect data almost continuously at relatively short time intervals once some critical crack has been identified. The kind of strategy is highly dependent of the questions addressed during operational evaluation. All these contents can affect more or less directly the readings collected, regarding the presence and location of damage.

2.3 Damage-sensitive feature extraction

The part of the SHM process that has demanded high research efforts is the identification of data features that allows one to distinguish between undamaged and damaged states of the monitored structure (DOEBLING et al., 1996; SOHN et al., 2004). A damage-sensitive feature is some quantity extracted from the structural response data that is correlated with the presence of damage in a structure (e.g., modal parameters, maximum displacements, regression model parameters and residual errors), indication the presence (or not) of damage in a structure. An adequate damage-sensitive feature will vary consistently in accordance with the level of damage in the structure. However, the

type of feature differs from the kind of structure and the objective of monitoring. Fundamentally, the feature extraction process is based on fitting some model, either physics- or data-based, to the measured response data. The parameters of these models, or the predictive errors associated with them, become the damage-sensitive features. Generally, a degree of signal processing is required in order to extract effective features.

2.4 Statistical modeling for feature classification

Undoubtedly, the portion of the SHM process with least attention in the technical literature is concerned to the development of statistical models for discrimination between features from the undamaged and damaged structures. Statistical modeling for feature classification is concerned with the implementation of algorithms that analyze the distributions of the extracted features in an effort to determine the structural condition at a given period of time. The functional relationship between the selected features and the damage state of the structure is often difficult to define based on physics-based engineering analysis procedures. Therefore, the statistical models are derived using machine learning techniques. These algorithms usually fall into three categories: (i) group classification, (ii) regression analysis, and (iii) outlier or novelty detection. The appropriate algorithm to use depends on the ability to perform supervised or unsupervised learning. In the context of SHM applications, supervised learning is referred to the case where examples of data from damaged and undamaged conditions are available; group classification and regression analysis are often used for this purpose. On the other hand, unsupervised learning arises when only data from the undamaged structure are available for training, where outlier or novelty detection methods are the primary class of algorithms used in this situation. However, for high capital expenditure infrastructures, such as civil ones, the unsupervised learning are often required because only data from the undamaged condition are available.

In SHM, for general purposes, the training matrix $\mathbf{X} \in \mathbb{R}^{n \times m}$ is composed of n observations under operational and environmental variability when the structure is undamaged, where m is the number of features per observation obtained during the feature extraction phase. The test matrix $\mathbf{Z} \in \mathbb{R}^{v \times m}$ is defined as a set of t observations collected during the undamaged/damaged conditions of the structure. Note that an observation represents a feature vector encoding the structural condition at a given time, and a data cluster represents a set of feature vectors corresponding to a global normal and stable state condition of the structural system. In follow, some of the main machine learning algorithms for statistical modeling are briefly described.

2.4.1 Linear principal component analysis

An output-only technique that has been applied for eliminating environmental influences on features is linear principal component analysis. This multivariate statistical procedure aims to estimate a linear static relationship between the extracted features and the unknown normal influences by reducing the dimensionality of the original input data through a linear projection onto a lower dimensional space. This linear map allows one to remove the normal variations in terms of retained components. In the SHM field, PCA has been used for several purposes (feature selection, feature cleansing and visualization). Herein, PCA is used as a data normalization method.

Assuming the training data matrix $\mathbf{X}$ decomposed in the form of (JOLLIFFE, 2002)

$$\mathbf{X} = \mathbf{T}\mathbf{U}^T = \sum_{i=1}^{m} \mathbf{t}_i \mathbf{u}_i^T, \tag{2.1}$$

where $\mathbf{T}$ is called the scores matrix and $\mathbf{U}$ is a set of m orthogonal vectors, $\mathbf{u}_i$, also called the loadings matrix. The orthogonal vectors can be obtained by decomposing the covariance matrix of $\mathbf{X}$ in the form of $\Sigma = \mathbf{U}\Lambda\mathbf{U}^T$, where Λ is a diagonal matrix containing the ranked eigenvalues λ_i, and $\mathbf{U}$ is the matrix containing the corresponding eigenvectors. The eigenvectors associated with the higher eigenvalues are the principal components of the data matrix and they correspond to the dimensions that have the largest variability in the data. Basically, this method permits one to perform an orthogonal transformation by retaining only the principal components d ($\leq m$), also know as the number of factors. Precisely, choosing only the first d eigenvectors, the final matrix can be rewritten without significant loss of information in the form of

$$\mathbf{X} = \mathbf{T}_d \mathbf{U}_d^T + \mathbf{E} = \sum_{i=1}^{d} \mathbf{t}_i \mathbf{u}_i^T + \mathbf{E}, \tag{2.2}$$

where $\mathbf{E}$ is the residual matrix resulting by the d factors. The coefficients of the linear transformation are such that if the feature transformation is applied to the data set and then reversed, there will be a negligible difference between the original and reconstructed data.

In the context of data normalization, the PCA algorithm can be summarized as follows: the loadings matrix is obtained from $\mathbf{X}$, the test matrix $\mathbf{Z}$ is mapped onto the feature space $\mathbb{R}^d$ and reversed back to the original space $\mathbb{R}^m$, the residual matrix $\mathbf{E}$ is computed as the difference between the original and the reconstructed test matrix

$$\mathbf{E} = \mathbf{Z} - (\mathbf{Z}\mathbf{U}_d)\,\mathbf{U}_d, \tag{2.3}$$

and finally in order to establish a quantitative measure of damage, the structural health condition is discriminated by generating a damage indicator (DI) and classifying through a threshold. For the l feature vector ($l = 1, 2, ..., v$), a DI is adopted in the form of the squared root of the sum-of-square errors (Euclidean norm):

$$\mathrm{DI}(l) = \parallel e_l \parallel . \tag{2.4}$$

If l feature vector is related to the undamaged condition, $z_l \approx \hat{z}_l$ and $DI \approx 0$. On the other hand, if the feature vector comes from the damaged condition, the residual errors increase, and the DI deviates from zero, thereby indicating an abnormal condition in the structure.

Therefore, the classification is performed using a linear threshold for a certain level of significance. In this work, the threshold is defined for 95% of confidence on the DIs taking into account only the baseline data used in the training process. Thus, if the approach has learned the baseline condition, then it is statistically guaranteed approximately 5% of misclassifications in the DIs derived from undamaged observations not used for training.

2.4.2 Auto-associative neural network

The AANN is trained to characterize the underlying dependency of the extracted features on the unobserved operational and environmental factors by treating that unobserved dependency as hidden intrinsic variables in the network architecture. The AANN architecture consists of three hidden layers: the mapping layer, the bottleneck layer, and de-mapping layer.

While PCA is restricted to mapping only linear correlations among variables, AANN can reveal the nonlinear correlations present in the data. If nonlinear correlations exist among variables in the original data, the AANN can reproduce the original data with greater accuracy and/or with fewer factors than PCA. This NLPCA can be achieved by training a feed-forward neural network to perform the identity mapping, where the network outputs are simply the reproduction of network inputs. Thus, this architecture is a special case of autoencoder (Figure 2).

In the context of data normalization for SHM, the AANN (SOHN; WORDEN; FARRAR, 2002) is first trained to learn the correlations between features from the training matrix $\mathbf{X}$. Then, the network should be able to quantify the unmeasured sources of variability that influence the structural response. This variability is represented at the bottleneck output, where the number of nodes (or factors) should correspond to the number of unobserved independent factors (e.g., wind speed, temperature, loading) that influence

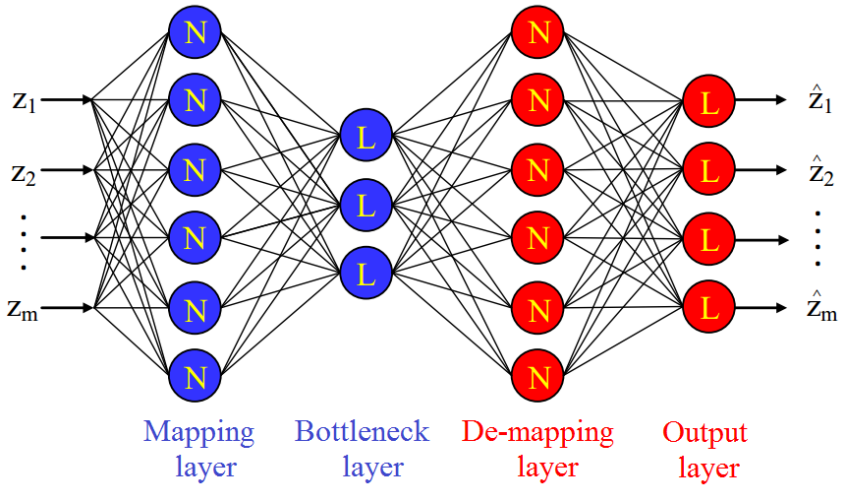

Figure 2 – Schematic representation of an AANN (SOHN; WORDEN; FARRAR, 2002).

the structural response. Second, for the test matrix $\mathbf{Z}$, the residual matrix $\mathbf{E}$ is given by

$$\mathbf{E} = \mathbf{Z} - \hat{\mathbf{Z}}, \tag{2.5}$$

where $\hat{\mathbf{Z}}$ corresponds to the estimated feature vectors that are the output of the network. The DIs are calculated using the Equation (2.4) on the residuals from $\mathbf{E}$. The threshold is defined in accordance to described in Section 2.4.2.

2.4.3 Kernel principal component analysis

An improved approach to perform NLPCA is based on kernel PCA. Proposed by Reynders (REYNDERS; WURSTEN; ROECK, 2014) for eliminating environmental and operational influences from damage-sensitive features, this technique has demonstrated ability to learn more slight nonlinear dependencies than the AANN.

The main intuition behind this method is related to the realization of the linear PCA in some high dimensional feature space $\mathcal{F}$. Since $\mathcal{F}$ is nonlinear in relation to input space, the contour lines of constant projections onto the principal eigenvector become nonlinear in input space. Crucial to KPCA is the fact that there is no need to carry out the map into $\mathcal{F}$. All necessary computations are carried out by the use of a kernel function Q in the input space. Figure 3 compares the space projection on the linear PCA and KPCA.

The radial bases function (RBF) (KEERTHI; LIN, 2003) is the most common kernel function used to nonlinear mapping. This kernel maps examples into a higher dimensional space so it can handle the case when the relation between class labels and

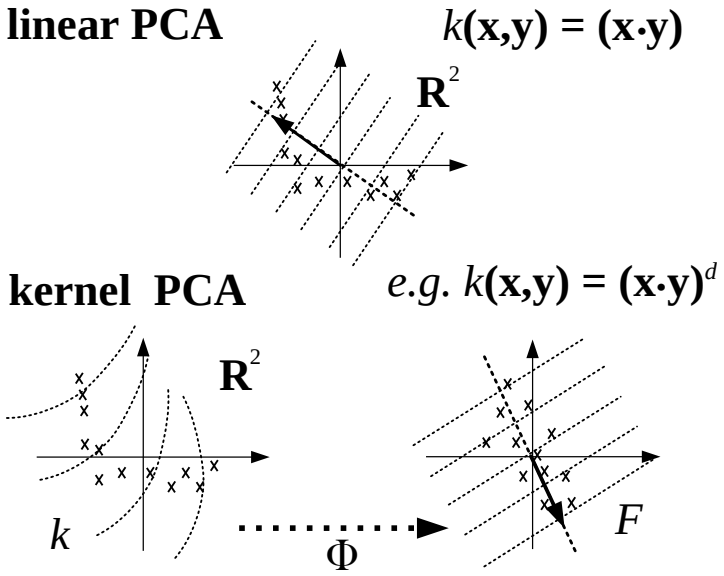

Figure 3 – The basic idea behind KPCA (SCHOLKOPF; SMOLA; MULLER, 1998).

attributes is nonlinear. The RBF kernel can be expressed by (CHANG; LIN, 2011)

$$Q(\mathbf{x}_i, \mathbf{x}_y) = \exp\left(-\gamma \parallel \mathbf{x}_i - \mathbf{x}_j \parallel^2\right), \gamma > 0, \tag{2.6}$$

where γ is a kernel parameter that controls the bandwidth of the inner product matrix Q. An optimal value of γ can be estimated by requiring that the corresponding inner product matrix is maximally informative as measured by Shannon's information entropy. The detailed steps to estimate the optimal value of γ can be found in (REYNDERS; WURSTEN; ROECK, 2014).

In the context of SHM, the input training matrix $\mathbf{X}$ is mapped by ϕ: $\mathbf{X} \rightarrow \mathcal{F}$ to a high dimensional feature space $\mathcal{F}$. The linear PCA is applied on the mapped data $\mathcal{T}_\Phi = \{\phi(\mathbf{x}_1), ..., \phi(\mathbf{x}_n)\}$. The computation of the principal components and the projection on these components can be expressed in terms of dot products, thus the RBF kernel function can be employed. The KPCA trains the kernel data projection

$$\mathbf{U} = A^T Q(\mathbf{x}) + b, \tag{2.7}$$

where A is the projection matrix, b is the bias vector and $Q(\mathbf{x})$ is the kernel function centered in the training vectors. The kernel mean squared reconstruction error, which must be minimized, is defined such that

$$\varepsilon_{KMS}(A, b) = \frac{1}{n} \sum_{i=1}^{n} \parallel \phi(\mathbf{x}_i) - \tilde{\phi}(\mathbf{x}_i) \parallel^2, \tag{2.8}$$

where the reconstructed vector $\tilde{\phi}(\mathbf{x})$ is given as a linear combination of the mapped data $\mathcal{T}_\Phi$

$$\tilde{\phi}(\mathbf{x}) = \sum_{i=1}^{n} \beta_i \phi(\mathbf{x}_i), \;\; \beta = A(\mathbf{u}_i - b). \tag{2.9}$$

In contrast to the linear PCA, the explicit projection from the feature space $\mathcal{F}$ to the input space usually does not exist (SCHOLKOPF; SMOLA, 2001).

For the test matrix $\mathbf{Z}$, the residual matrix $\mathbf{E}$ is given by

$$\mathbf{E} = \phi(\mathbf{Z}) - \tilde{\phi}(\mathbf{Z}), \tag{2.10}$$

where $\phi(\mathbf{Z})$ corresponds to the high dimensional feature vectors and $\tilde{\phi}(\mathbf{Z})$ their corresponding reconstruction after linear PCA. The DIs are calculated using the Equation (2.4) on the residuals from $\mathbf{E}$. The threshold is defined in accordance to described in Section 2.4.2.

2.4.4 Mahalanobis squared-distance

Another well-known method for performing data normalization without any information regarding the environmental and operational influences is based on the Mahalanobis squared-distance (MSD) (WORDEN; MANSON; ALLMAN, 2003; FIGUEIREDO, 2010). The Mahalanobis distance differs from the Euclidean distance because it takes into account the correlation between the variables and it does not depend on the scale of the features. However, this model assumes that data follows a unique multivariate Gaussian distribution, i.e., the data can be modeled by only one Gaussian cluster.

Considering the training data matrix $\mathbf{X}$, a mean feature vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ are estimated. In the context of data normalization, the mean vector and covariance matrix should encode all normal variations represented by the baseline data. Thus, for a test data $\mathbf{z}_l$, the MSD is used as a standard outlier analysis procedure, providing DIs by (WORDEN, 1997)

$$\mathrm{DI}(l) = (\mathbf{z}_l - \boldsymbol{\mu})\, \boldsymbol{\Sigma}^{-1} (\mathbf{z}_l - \boldsymbol{\mu})^T . \tag{2.11}$$

The main assumption is that if a test feature vector is obtained from data collected on the damaged system that includes operational and environmental variability, similar to those from training data, this vector will be further from the mean feature vectors corresponding to the undamaged condition as quantified by the MSD. The classification and threshold definition are realized in accordance to described in Section 2.4.2.

2.4.5 Gaussian mixture models

To overcome the limitations imposed by MSD, Gaussian mixture models are usually employed. This algorithm carries out a model-based clustering, using multivariate finite mixture models, that aims to capture the main clusters (or components) of features that correspond to the normal and stable state conditions of a structure under operational and environmental conditions. Compared to MSD, the GMM assumes that the data can be modeled by a set of multivariate Gaussian distributions, allowing to learn nonlinear relationships (Figure 4).

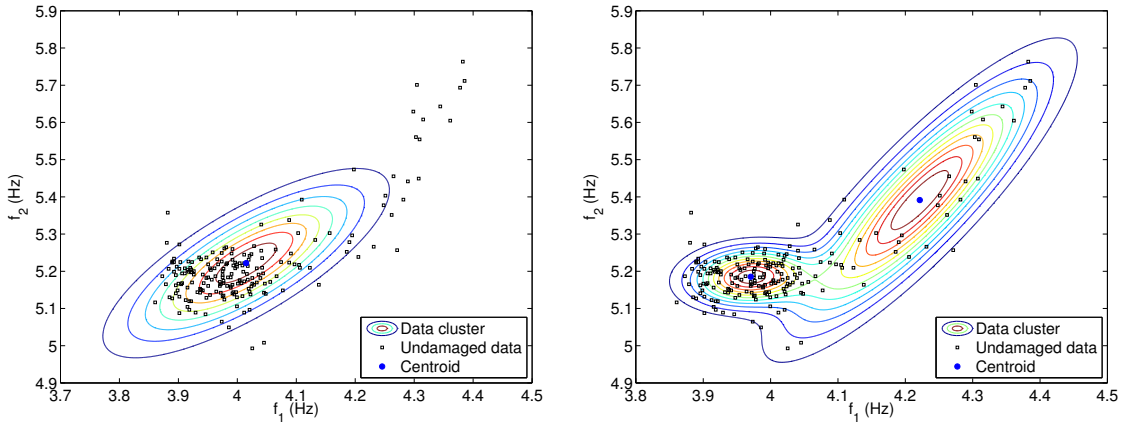

Figure 4 – Comparison between MSD (left) and GMM (right) models (FIGUEIREDO et al., 2014a).

Therefore, a finite mixture model, $p(\mathbf{x}|\boldsymbol{\Theta})$, is the weighted sum of $K > 1$ components $p(\mathbf{x}|\boldsymbol{\theta}_k)$ in $\mathbb{R}^m$,(MCLACHLAN; PEEL, 2000)

$$p(\mathbf{x}|\boldsymbol{\Theta}) = \sum_{k=1}^{K} \alpha_k p(\mathbf{x}|\boldsymbol{\theta}_k), \tag{2.12}$$

where α_k corresponds to the weight of each component. These weights are positive $\alpha_k > 0$ with $\sum_{k=1}^{K} \alpha_k = 1$. For a GMM, each component $p(\mathbf{x}|\boldsymbol{\theta}_k)$ is represented as a Gaussian distribution,

$$p(\mathbf{x}|\boldsymbol{\theta}_k) = \frac{\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right\}}{(2\pi)^{m/2}\sqrt{\det(\boldsymbol{\Sigma}_k)}}, \tag{2.13}$$

being each component denoted by the parameters, $\boldsymbol{\theta}_k = \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$, composed of the mean vector, $\boldsymbol{\mu}_k$ and the covariance matrix, $\boldsymbol{\Sigma}_k$. Thus, a GMM is completely specified by the set of parameters $\boldsymbol{\Theta} = \{\alpha_1, \alpha_2, \ldots, \alpha_K, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_K\}$.

The EM algorithm is the most widespread local search method used to estimate the parameters of the GMMs (DEMPSTER; LAIRD; RUBIN, 1977; MCLACHLAN; PEEL, 2000). This method consists of an expectation step and a maximization step which are

alternately applied until the log-likelihood (LogL), $\log p\left(\mathbf{X}|\boldsymbol{\Theta}\right) = \log \prod_{i=1}^{n} p\left(\mathbf{x}_i|\boldsymbol{\Theta}\right)$, converges to a local optimum (DEMPSTER; LAIRD; RUBIN, 1977). The performance of the EM algorithm depends directly on the choice of the initial parameters $\boldsymbol{\Theta}$, which may implies many replications of this method during an execution (FIGUEIREDO; JAIN, 2002). To select the best GMM by means of goodness-of-fit and parsimony, the Bayesian information criterion (BIC) is used and minimized,(BOX; JENKINS; REINSEL, 2008)

$$\mathrm{BIC} = -2\log p\left(\mathbf{X}|\boldsymbol{\Theta}\right) + \left\{Km\left[\left(\frac{m+1}{2}\right)+1\right] + K - 1\right\}\log\left(n\right). \tag{2.14}$$

Similar to Akaike information criterion (AIC), BIC uses the optimal LogL function value and penalizes for more complex models, i.e., models with additional parameters. The penalty term of BIC is a function of the training data size, and so it is often more severe than AIC.

For the damage detection process, and for each observation $\mathbf{z}_l$, one needs to estimate K DIs. Basically, for each component k discovered during training phase

$$\mathrm{DI}_k(l) = \left(\mathbf{z}_l - \boldsymbol{\mu}_k\right)\boldsymbol{\Sigma}_k^{-1}\left(\mathbf{z}_l - \boldsymbol{\mu}_k\right)^T, \tag{2.15}$$

where $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ represent the parameters from all the observations of the k-th component, when the structure is undamaged even though under varying operational and environmental conditions. Note that, each component is related to a specific source of variability (e.g., traffic loading, wind speed, temperature and boundary conditions), which allows to provide physical meanings for each component. Finally, for each observation, the DI is given by the smallest DI estimated on each component

$$\mathrm{DI}(l) = \min\left(\mathrm{DI}_1, \mathrm{DI}_2, \cdots, \mathrm{DI}_K\right). \tag{2.16}$$

Thus, it is expected that if the algorithm has learned the baseline condition, i.e., the identified components suitably represent the undamaged and normal condition under all possible operational and environmental conditions, for a defined threshold, the approach should outputs less than 5% of false-positive alarms for undamaged data from test matrix. Additionally, the threshold is defined as described in Section 2.4.2.

3 Proposed machine learning algorithms for data normalization

In SHM, the purpose of machine learning algorithms is to enhance the damage detection in structures subjected to the presence of varying operational and environmental conditions under which the system response is measured. The previous chapter was dedicated to the description of the SPR paradigm and the main statistical methods for data normalization reported in literature. Herein, two novel machine learning algorithms for statistical modeling are proposed. These approaches intend to improve the performance of traditional methods based on PCA and cluster analysis.

3.1 Deep learning algorithms

Deep learning methods aim at learning feature hierarchies of features from higher levels of the hierarchy formed by the composition of lower level features. They include learning methods for a wide range of deep architectures using graphical models with many levels of hidden variables (HINTON; SALAKHUTDINOV, 2006; BENGIO, 2009; ERHAN et al., 2010). However, the main trend of this field is based on neural networks with massive amount of hidden layers (BENGIO et al., 2007; RANZATO et al., 2007; VINCENT et al., 2008; WESTON et al., 2012).

The big challenge for development of this field concerns to the training algorithms used to build deep architectures. In virtually all instances of deep learning, the objective function is a highly non-convex function of the parameters, with the potential for many distinct local minima in the model parameter space. The main problem is that not all of these minima provide equivalent results, due to the standard training algorithms (based on random initialization) guide the searching towards areas with poor generalization performance (BENGIO et al., 2007).

The breakthrough of the area was achieved by (HINTON; SALAKHUTDINOV, 2006; RANZATO et al., 2007; BENGIO et al., 2007) with development of a two-phase training scheme: greedy layer-wise unsupervised pre-training followed by supervised fine-tuning. First, each layer is pre-trained with an unsupervised algorithm that learns a nonlinear transformation of its input (the output of the previous layer) and captures the more relevant variations. Second, for the final training phase, the deep architecture is fine-tuned with respect to a supervised training criterion using gradient-based optimization.

This novel training scheme has dramatically increased the performance of such methods, driving intense research efforts to the field (ERHAN et al., 2010), mainly to the unsupervised branch. In this context, methods based on stacked autoencoders has demanded the main attention on the last years.

3.1.1 Autoencoders

An autoencoder is a neural network that is trained to attempt to copy its input to the output. Internally, it has a hidden layer $\mathbf{h}$ that describes a code used to represent the input. The network may be viewed as consisting of two parts: an encoder function $\mathbf{h} = f(\mathbf{x})$ and a decoder that produces a reconstruction $\hat{\mathbf{x}} = g(\mathbf{h})$. This architecture is presented in Figure 5. If an autoencoder succeeds in simply learning to set $g(f(\mathbf{x})) = \mathbf{x}$ everywhere, then it is not especially useful. Instead, autoencoders are designed to be unable to learn perfectly the input patterns. This is made by adding constraints to perform only approximate copies of the input patterns into the output. Thus, the model is forced to prioritize the most important characteristics of the data, allowing to learn useful properties (GOODFELLOW; BENGIO; COURVILLE, 2016).

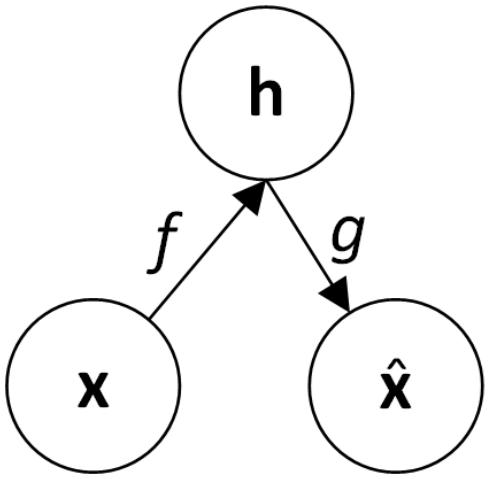

Figure 5 – General scheme of a simple autoencoder (GOODFELLOW; BENGIO; COURVILLE, 2016).

Traditionally, autoencoders were used for dimensionality reduction or feature learning. Recently, theoretical connections between autoencoders and latent variable models have brought autoencoders to the forefront of generative modeling. A common manner to obtain useful representations is to constrain $\mathbf{h}$ to have smaller dimension than $\mathbf{x}$. The case where $\mathbf{h}$ have a greater dimension than $\mathbf{x}$ has not demonstrated to result in something more useful than learn a simple identity function. An autoencoder whose code dimension is less than the input dimension is called undercomplete. This representation forces the autoencoder to capture the most salient features of the training data.

The learning process is described simply as minimizing a loss function

$$L\left(\mathbf{x}, g\left(f\left(\mathbf{x}\right)\right)\right) = \frac{1}{\lambda n} \sum_{\forall \mathbf{x} \in \mathbf{X}} \left\| g\left(f\left(\mathbf{x}\right)\right) - \mathbf{x} \right\|^2, \tag{3.1}$$

where L is a loss function penalizing $g\left(f\left(\mathbf{x}\right)\right)$ for being dissimilar from $\mathbf{x}$, such as the mean squared error, and λ is a hyperparameter which controls the asymptote slope. In this case, an undercomplete autoencoder learns to span the principal subspace of the training data, i.e., the same subspace as PCA. Autoencoders with nonlinear encoder functions f and nonlinear decoder functions g can thus learn a more powerful nonlinear generalization than KPCA, and even better than AANN.

3.1.2 Stacked autoencoders

Multilayer autoencoders are feed-forward neural networks with an odd number of hidden layers (DEMERS; COTTRELL, 1992; HINTON; SALAKHUTDINOV, 2006) and shared weights between the encoder and decoder layers (asymmetric network structures may be employed as well). The middle hidden layer has a number of nodes equal to the number of factors to be retained, d, as in PCA. The network is trained to minimize the mean squared error between the input and the output of the network (ideally, the input and the output are equal).

The goal of training is to perform, in the middle hidden layer (i.e., the bottleneck layer), a lower dimensional representation of the input data, in such manner that it preserves as much data structure as possible. The lower dimensional representation can be obtained by extracting the node values in the middle hidden layer for a given point used as input. In order to allow the autoencoder to learn a nonlinear mapping between the high-dimensional and low-dimensional data representation, sigmoid activation functions are generally used (except in the middle layer, where a linear activation function is usually employed).

The term *stacked autoencoders* relates to the concept of stacking simple modules of functions or classifiers, as proposed and explored in (WOLPERT, 1992; BREIMAN, 1996; BENGIO et al., 2007), to compose a strong model. These simple models are independently trained to learn specific tasks, and then they are "stacked" on top of each other in order to learn more complex representations of the input patterns. Experimentally, stacked autoencoders yield much better dimensionality reduction than corresponding shallow or linear autoencoders (HINTON; SALAKHUTDINOV, 2006).

Typically, these networks are represented by many layers of nonlinear mapping, characterizing a deep architecture with a high number of connections. Therefore, the usual training based on backpropagation algorithms converge slowly and are likely to get stuck in local minima. Thus, the previously presented two-phase training scheme is employed to

overcome these problems. An example of the two-phase training is shown schematically in Figure 6.

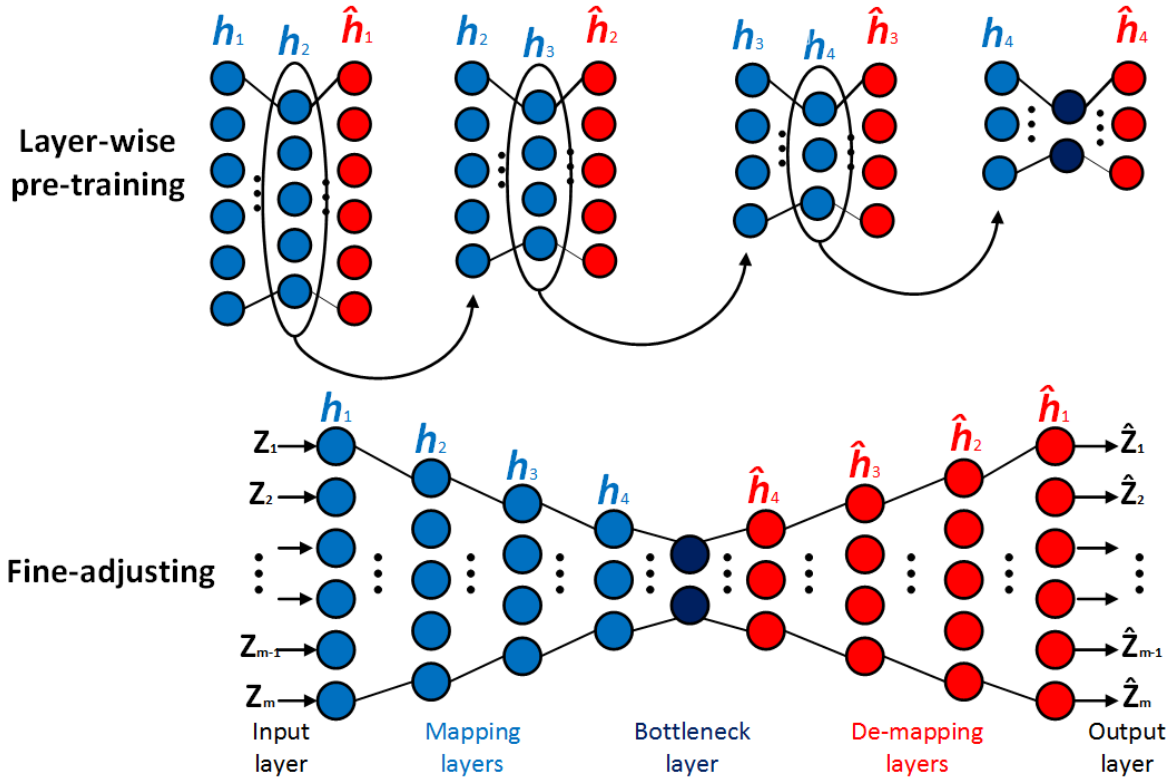

Figure 6 – Unsupervised layer-wise pre-training and fine adjustment of a nine-layer deep architecture.

In the context of SHM, the DSA allows to learn more reliable representations of the input vectors due to successively nonlinear transformations applied. The robust training tends to generate a more powerful model capable to generalize the normal conditions from training to the test data, performing better filtering of environmental and operational variations. Furthermore, the problems related to traditional PCA-based approaches are circumvented by this algorithm (i.e., assumptions of data normality and definition of hyper parameters are not performed). The criterion to automatically define the number of hidden nodes in each layer, as well as more details about the proposed architecture, can be found in (MAATEN; POSTMA; HERIK, 2009).

Herein, for SHM, a nine-layer DSA is trained to represent, in the bottleneck layer, low level features from training matrix $\mathbf{X}$. These new features must characterize the hidden factors that changed the underlying distribution of the structural dynamic response. For the test matrix $\mathbf{Z}$, the residual matrix $\mathbf{E}$ is built as stated in Equation (2.5), and the resulting DIs are calculated by Equation (2.4). The threshold is defined in accordance to defined in Section 2.4.2.

3.1.3 Deep autoencoders and traditional principal component analysis

In Table 1, the approaches based on PCA are summarized by five general properties: (1) ability to derive nonlinear mappings and learning nonlinear relationships in the data, (2) direct mapping between the original dimensional and the low dimensional space, (3) strengths to perform adequate learning of linear and nonlinear dependencies, (4) extraction of features that are strongly correlated to the data in original space, and (5) the computational complexity regarding the main procedures of each technique. All properties are discussed below.

Table 1 – Comparison of the different PCA-based approaches.

Characteristic	DSA	PCA	AANN	KPCA
Nonlinear mapping	✓	✗	✓	✓
Direct mapping	✓	✓	✓	✗
Generalization	✓	✗	✗	✗
Robust feature extraction	✓	✗	✗	✗
Complexity	$\mathcal{O}(limw)$	$\mathcal{O}(d^3)$	$\mathcal{O}(imw)$	$\mathcal{O}(n^3)$

The property 1 is concerned to the learning of nonlinear relationships in the data, which for real monitoring scenarios where structure dynamics is highly influenced by nonlinear variations is a crucial issue. In that regard, as linear PCA performs only linear orthogonal transformations it is not able to learn, properly, nonlinear sources of varying normal conditions, resulting in inaccurate damage detection performance in many applications. On the other hand, the approaches based on DSA, AANN and KPCA can handle this matter by different mechanisms, which ensure different levels of removing normal variations.

From property 2 one can infer that some techniques are not able to specify a direct mapping from the original dimension to the low dimensional space (or vice versa). It can be pointed out as a disadvantage for two main reasons: (1) it is not possible to generalize for held-out or new test data without performing a new mapping, as well as any insights about the amount of normal variability retained from the mapping/demapping operation can not be inferred. From that point of view, the approach based on KPCA is the only one that does not fall into this property, due to the learning and mapping performed in high dimensional space.

Generalization is related to the capacity of a technique to accurately learn patterns and predict outcome values from previously unseen data. Thus, property 3 is regarded to indicate whether technique is able or not to perform, correctly, the assessment of data arising from new measurements. As demonstrated by (SANTOS et al., 2015), the approaches based on PCA, AANN and KPCA have several limitations when evaluating new structural conditions, mainly in the cases of extreme operational and environmental factors. In contrast, deep neural networks have the ability to properly evaluate new data.

That characteristic can be assigned to the robust training phase that allows a high quality feature extraction.

One of the main qualities of deep neural networks is the robust nonlinear mapping/demapping of data, deriving compressed representation of the input variables. In comparison, the traditional approaches to perform PCA do not allow a proper learning of the relationships between the input variables mainly due to the training algorithms and specific procedures that limit the kind of features to be extracted (e.g., linear PCA applies only orthogonal transformations, as well as the backpropagation algorithm and kernel mapping may not drive the parameter tuning towards global optimum). In that regard, property 4 indicates whether techniques provide robust mapping/demapping to compose strong features.

For property 4, Table 1 provides insight into the computational complexity of the computationally most expensive algorithmic components. The computational complexity is a issue of great importance to its practical applicability. If the memory or computational resources needed are too large, their application become infeasible. Thus, the corresponding computational complexity is determined by the properties of the dataset such as the number of samples and their dimensionality, as well as the input parameters of each algorithm, such as the target dimensionality d and the number of iterations i (for iterative techniques). In the case of neural networks, w indicates the size of the model by the number of weights and l the number of layers. The computationally most demanding part of PCA is the eigenvalue and eigenvector decompositions, which is performed using a method in $\mathcal{O}(d^3)$. Due to the kernel projection, KPCA performs an eigenanalysis of an $n \times n$ matrix by solving a semi-definite problem subjected to n constraints, requiring a learning algorithm in $\mathcal{O}(n^3)$. Both DSA and ANN employ backpropagation, which has a computational complexity of $\mathcal{O}(imw)$. However, as DSA has a pre-training phase, its complexity order increases to $\mathcal{O}(limw)$.

From the discussion of the five general properties above, it is possible to derive some considerations: (1) some techniques do not provide a direct mapping between the original space and the mapped one, (2) when the number of factors to be retained is close to the number of samples, $d \approx n$, nonlinear techniques have computational disadvantages compared to linear PCA, and (3) a considerable number of nonlinear techniques suffer from high demands of computational efforts. From these observations, it is clear that nonlinear techniques impose considerable demands on computational resources, as compared to linear PCA. Considering the DSA-based approach, its application must be concerned to the level of improvement in the data normalization and damage detection performance, since the high computational resources required can limit the applicability for real monitoring scenarios.

3.2 Agglomerative clustering

The agglomerative clustering algorithms are part of a hierarchical clustering strategy which treats each object as a single cluster, and iteratively merges (or agglomerates) subsets of disjoint groups, until some stop criterion is reached (e.g., number of clusters equal to one) (KAUFMAN, 1990; MAIMON; ROKACH, 2010). These bottom-up algorithms create suboptimal clustering solutions, which are typically visualized in the form of a dendrogram representing the level of similarity between two adjacent clusters, allowing to rebuild step-by-step the resulting merging process. Any desired number of clusters can be obtained by cutting the dendrogram properly.

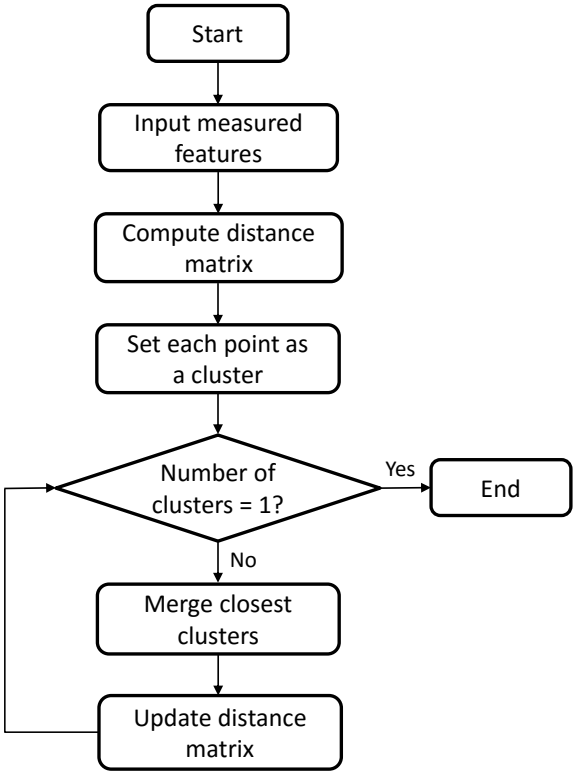

Figure 7 – Flow chart of agglomerative hierarchical clustering.

The common flow chart of an agglomerative clustering procedure is summarized in Figure 7. Initially, every observation is defined as a centroid. Then, a similarity matrix composed of the distances between each cluster is computed to determine which clusters can be merged. Usually, this agglomerative process is repeated until only one cluster remains. As described in (BERRY; LINOFF, 1997), when the cluster diameter is small, the corresponding data group is defined more precisely as this group is composed by few members strongly correlated. The fundamental assumption is that small clusters are more coherent than large ones (BERRY; LINOFF, 1997; MANNING; RAGHAVAN; SCHüTZE, 2008).

3.2.1 Agglomerative concentric hyperspheres

The ACH algorithm is an agglomerative cluster-based technique, working in the feature space, composed of two main steps: (1) an off-line initialization and (2) bottom-up clustering procedure. Depending on the type of initialization mechanism, the algorithm becomes completely deterministic or random. The initialization has a direct influence on the algorithm performance. Hereafter, all clusters are merged, in an iterative manner, by evaluating the boundary regions that limit each cluster through inflation of a concentric hypersphere. These two steps allow to automatically discover the number of clusters and, therefore, no input parameters are required. Note that, none information regarding data distribution is required. The entire clustering procedure is divided into three main phases:

(i) *Centroid displacement.* For each cluster, its centroid is dislocated to the position with higher observation density, i.e. the mean of its observations.

(ii) *Linear inflation of concentric hyperspheres.* Linear inflation occurs on each centroid by progressively increasing an initial hypersphere radius,

$$R_0 = \log_{10}\left(\|\mathbf{c}_i - \mathbf{x}_{max}\|^2 + 1\right), \tag{3.2}$$

where $\mathbf{c}_i$ is the centroid of the i-th cluster, used as a pivot, and $\mathbf{x}_{max}$ is its farthest observation, such that $\|\mathbf{c}_i - \mathbf{x}_{max}\|^2$ is the radius of the cluster centered in $\mathbf{c}_i$. The radius grows up in the form of an arithmetic progression (AP) with common difference equal to R_0. The creation of new hyperspheres is set by a criterion based on the positive variation of the observation density between two consecutive inflations, defined as the inverse of variance; otherwise the process is stopped.

(iii) *Cluster merging.* If there is more than one centroid inside the inflated hypersphere, all centroids are merged to create an unique representative centroid positioned at the mean of the centroids' position. On the other hand, if only the pivot centroid is within the inflated hypersphere, this centroid is assumed to be on the geometric center of a cluster, thus the merging is not performed.

For completeness, Figure 8 presents an example of the algorithm applied to a three-component scenario with a five-centroid initial solution. First, in Figure 8a, the centroids are moved to the center of their clusters, as indicated in the first phase. In Figures 8b and 8c, two centroids are merged to form one cluster, as they are within the same inflated hypersphere. On the other hand, in Figure 8d only the pivot centroid is located in the center of a cluster, therefore the ACH algorithm does not perform the merge process. In the case where the merging occurs, all centroids analyzed before are evaluated again to

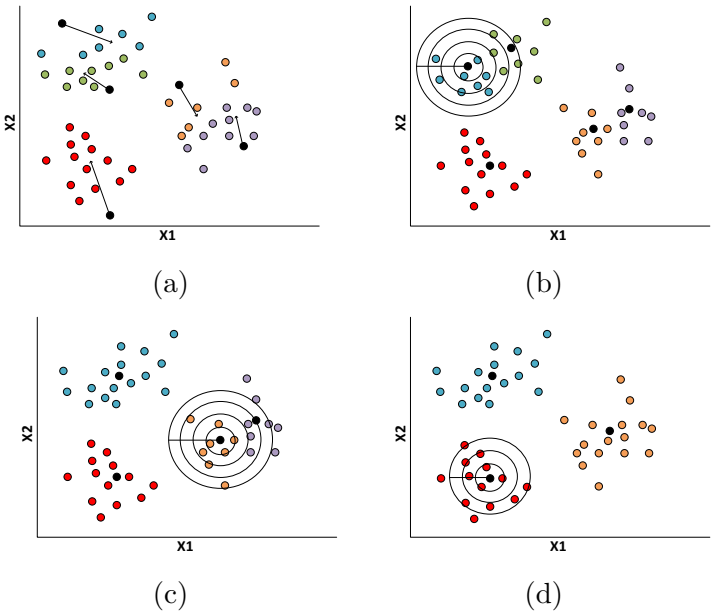

Figure 8 – ACH algorithm using linear inflation running in a three-component scenario.

infer if the new one is not badly positioned in another cluster or closer to a boundary region.

The Algorithm 1 summarizes the proposed method. Initially, it identifies the cluster to which each observation belongs and moves the centroids to the mean of their observations. Then, a hypersphere is built on the pivot centroid and it is inflated until the observation density decreases. Finally, the merges of all centroids within the hypersphere is performed, by replacing these centroids by their mean. The process is repeated until convergence, i.e. there is no centroid merging after the evaluation of all centroids or the final solution is composed by only one centroid.

Note that, the main goal of the clustering step is to maximize the observation density related to each cluster. In other words, to locate the positions with maximum observation concentration, also known as mass center, in such manner that when a hypersphere starts to inflate its radius, it reaches the decision boundaries of the cluster. This process is also described by maximizing the cost function

$$\begin{aligned} \max \quad & \sum_{k=1}^{K} \left(\sum_{\mathbf{x}_i \in \mathbf{c}_k} \frac{\|\mathbf{x}_i - \mathbf{c}_k\|^2}{\mathbf{N}_k} \right)^{-1}, \\ \text{s.t.} \quad & n = \sum_{k=1}^{K} \mathbf{N}_k, \\ & 1 \leq \mathbf{N}_k \leq n, \end{aligned} \tag{3.3}$$

where $\mathbf{c}_k$ is the k-th centroid, $\mathbf{x}_i$ the i-th observation assigned to the k-th cluster and $\mathbf{N}_k$ is the number of observations in the cluster k. The clustering procedure naturally carries out the optimization of the cost function by means of density evaluation, thus its direct

```
calcIndexes(C, X)
while not cover all elements of C do
    move(C, X)
    c_pivo = nextCenter(C)
    radius_init = calcRadius(c_pivo, X)
    radius, density_0, density_1, delta_0, delta_1 = 0
    repeat
        radius = radius + radius_init
        H = calcHypersphere(C, c_pivo, X, radius)
        density_0 = density_1
        density_1 = calcDensity(H)
        delta_0 = delta_1
        delta_1 = |density_0 − density_1|
    until (delta_0 > delta_1);
    reduce(C, H)
    if merging occurred then
        calcIndexes(C, X)
    end if
end while
```

Algorithm 1: Summary of the ACH algorithm.

computation is not necessary. The convergence is guaranteed by gradual decreasing of the observation density as the hypersphere keeps inflating. More theoretical details and complexity analysis are provided in Appendix A and B, respectively.

3.2.1.1 Initialization procedures

Three procedures can be employed to choose the initial centroids, depending on the application. The *random* initialization is performed by choosing $p < n$ distinct observations from the training matrix as initial centroids. This is quite similar to the initialization procedure often used in the K-means algorithm (MACQUEEN, 1967).

To accomplish a deterministic clustering, two non-stochastic initializations are presented as well. The first one performs an eigenvector decomposition to create as many centroids as the number of observations in the training set, through a *divisive* procedure, quite similar to the one described in (HAMERLY; ELKAN, 2003). Primarily, the mean value of all data points is divided in other two points generated by

$$y_{new} = y \pm \mathbf{u}_i \sqrt{2\frac{\mathbf{t}_{i,i}}{\pi}}, \tag{3.4}$$

where $\mathbf{u}_i$ and $\mathbf{t}_{i,i}$ are the most significant eigenvector and eigenvalue, respectively, and y the point being divided. Each new point is divided in other two points in opposite directions placed around dense regions of the feature space. The process is repeated until the number of points is equal to p; then they are used as initial centroids. At the end of this

divisive procedure, each point has moved towards the region of higher concentration of observations, benefiting posterior clustering approaches. The second non-stochastic initialization divides the training matrix *uniformly* and chooses equally spaced observations as the initial centroids. The gap between each chosen observation is a factor of the number of training observations, usually equal to $\lceil n/p \rceil$. The selected observations are used as initial centroids. The parameter p, in all cases, can be equal to $\lceil n/2 \rceil$.

The damage detection process is carried out for each test observation $\mathbf{z}_l$ by estimating K DIs, as well as described in Section 2.4.5. However, in contrast to the GMM-based approach, for each component k discovered during training phase, the DI is calculated using euclidean distance. Thus, for each observation, the smallest distance estimated on each cluster learned is used as the actual DI.

4 Experimental results and analysis

In this section, the performance of the proposed approaches is evaluated with state-of-the-art ones based on PCA and cluster analysis. The methods are compared on the basis of Type I and Type II errors and their capabilities to filter linear/nonlinear changes, when dealing with operational and environmental effects, on standard data sets from Z-24 and Tamar Bridges.

Initially, the test bed structures are described briefly. The set of parameters chosen for each approach, as well as the main results, are presented by addressing each specific scenario. In every cases, the analysis are carried out by taking into account the results of PCA-based approaches separately from the cluster-based ones. Finally, for overall analysis purpose, all methods are compared from their performances on both scenarios.

Additionally, to determine which initialization procedure is more suitable to be employed with the ACH algorithm, a comparative study using the Z-24 Bridge data set is also carried out. The referred data set was chosen due to the extreme nonlinear effects, providing a reliable baseline to derive conclusions.

4.1 Test bed structures and data sets

In this work, the applicability and comparison between the proposed and state-of-the-art approaches are evaluated using the damage-sensitive features extracted from the Z-24 and Tamar Bridges. In the case of Z-24 Bridge, the standard data sets are unique in the sense that they combine one-year monitoring of the healthy condition, operational and environmental variability and realistic damage scenarios. In a different manner, a monitoring system was carried out on the Tamar Bridge during almost two-years, generating only data sets related to undamaged scenarios. In follow, these test structures and their data sets are highlighted.

4.1.1 Z-24 Bridge data sets

The Z-24 Bridge was a post-tensioned concrete box girder bridge composed of a main span of 30 m and two side-spans of 14m, as shown in Figure 9. The bridge, before complete demolition, was extensively instrumented and tested with the purpose of providing a feasibility tool for vibration-based SHM in civil engineering (PEETERS J. MAECK, 2001). A long-term monitoring test was carried out, from 11 November 1997

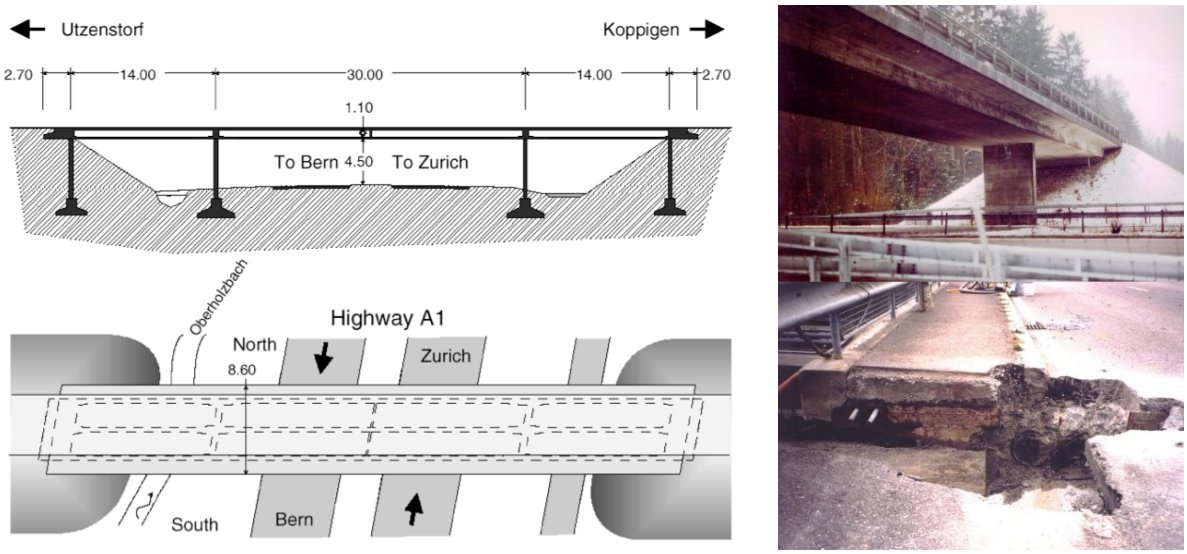

Figure 9 – Z-24 Bridge scheme (left) and picture (top right), as well as a damage scenario introduced by anchor head failure (bottom right).

until 10 September 1998, to quantify the operational and environmental variability present on the bridge and to detect damage artificially introduced, in a controlled manner, in the last month of operation. Every hour, during 11 minutes, eight accelerometers captured the mechanical vibrations of the bridge as well as an array of sensors measured environmental parameters, such as temperature at several locations.

Progressive damage tests were performed in one-month time period (from 4 August to 10 September 1998) before the demolition of the bridge to prove that realistic damage has a measurable influence on the bridge dynamics,(PEETERS J. MAECK, 2001) as summarized in Table 2. Note that the continuous monitoring system was still running during the progressive damage tests, which permits one to validate the SHM system to detect cumulative damage on long-term monitoring.

In this case, the natural frequencies of the Z-24 Bridge are used as damage-sensitive features. They were estimated using a reference-based stochastic subspace identification method on time series from the accelerometers (PEETERS; ROECK, 1999). The first four natural frequencies estimated hourly from 11 November 1997 to 10 September 1998, with a total of 3932 observations, are highlighted in Figure 10. The first 3470 observations correspond to the damage-sensitive feature vectors extracted within the undamaged structural condition under operational and environmental influences. The last 462 observations correspond to the damage progressive testing period, which is highlighted, especially in the second frequency, by a clear drop in the magnitude of the frequency.

Note that the damage scenarios are carried out in a sequential manner, which cause a cumulative degradation of the bridge. Therefore, in this work is assumed that the bridge operates within its undamaged condition (baseline or normal condition), under operational and environmental variability, from 11 November 1997 to 4 August 1998 (1–

Table 2 – Structural damage scenarios introduced progressively (details in (FIGUEIREDO et al., 2014a)).

Date	Description
04-08-1998	Reference measurement I (before any damage scenario)
09-08-1998	After installation of the settlement system
10-08-1998	Pier settlement = 2 cm
12-08-1998	Pier settlement = 4 cm
17-08-1998	Pier settlement = 5 cm
18-08-1998	Pier settlement = 9.5 cm
19-08-1998	Foundation tilt
20-08-1998	Reference measurement II (after removal of the settlement system)
25-08-1998	Spalling of concrete (12 m^2)
26-08-1998	Spalling of concrete (24 m^2)
27-08-1998	Landslide at abutment
31-08-1998	Concrete hinge failure
02-09-1998	Anchor head failure I
03-09-1998	Anchor head failure II
07-09-1998	Tendon rupture I
08-09-1998	Tendon rupture II
09-09-1998	Tendon rupture III

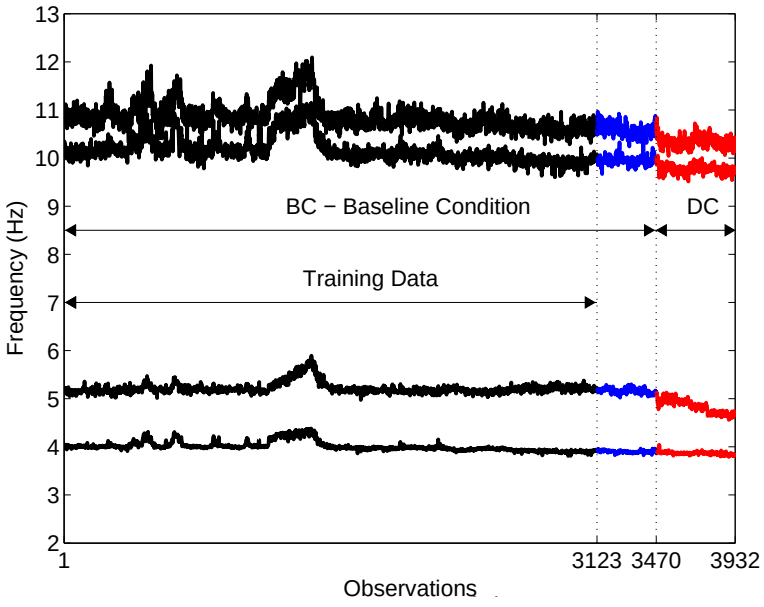

Figure 10 – First four natural frequencies of Z-24 Bridge. The observations in the interval 1-3470 are the baseline/undamaged condition (BC) and observations 3471-3932 are related to damaged condition (DC) (FIGUEIREDO et al., 2014b).

3470 observations). On the other hand, the bridge is considered damaged from 5 August to 10 September 1998 (3471–3932 observations). For the baseline condition period, the observed jumps in the natural frequencies are associated to the asphalt layer, in cold periods, which contributes, significantly, to the stiffness of the bridge. In this work, the

existence of a bilinear behaviour in the natural frequencies for below and above freezing temperature is nonlinearity (PEETERS; ROECK, 2001).

For generalization purposes, the feature vectors were split into the training and test matrices. As shown in Figure 10, the training matrix, $\mathbf{X}^{3123\times4}$, is composed of 90% of the feature vectors from the undamaged condition. The remaining 10% of the feature vectors are used during the test phase to make sure that the DIs do not fire off before the damage starts. The test matrix, $\mathbf{Z}^{3932\times4}$, is composed of all the data sets, even the ones used during the training phase.

4.1.2 Tamar Bridge data sets

The Tamar Bridge (depicted in Figure 11) is situated in the south-west of the United Kingdom and connects Saltash (Cornwall) with the city of Plymouth (Devon). This bridge is a major road across the River Tamar and plays a significant role in the local economy. Initially, in 1961, the bridge had a main span of 335 m and side spans of 114 m. If the anchorage and approach are included, the overall length of the structure is 643 m. The bridge stands on two concrete towers with a height of 73 m and a deck suspended at mid-height (CROSS et al., 2013).

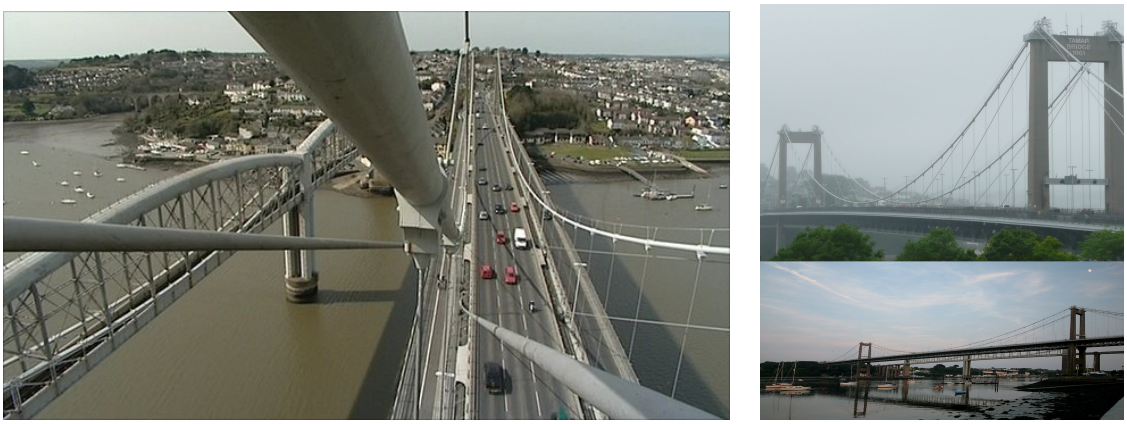

Figure 11 – The Tamar Suspension Bridge viewed from cantilever (left) and River Tamar margin (right).

Since 1961 the bridge structure was a steel truss supported vertically by a pair of suspension cables. To meet a European Union Directive, where bridges should be capable of carrying lorries of up to 40 tonnes, the bridge underwent a strengthening and widening upgrade scheme, which was completed in 2001 (KOO et al., 2013). The upgrade consisted of adding cantilevered lanes either side of the truss to provide a total of four lanes for traffic and a footpath for pedestrians. The heavy composite deck was replaced by an orthotropic steel deck and eight pairs of stay cables connected to the towers were added to support the increased weight of the deck.

To track the effects of the upgrade, various sensor systems were installed to extract monitoring data such as tensions on stays, accelerations, wind velocity, temperature, deflection and tilt. Eight accelerometers were implemented in orthogonal pairs for four stay cables and three sensors measured deck accelerations. The time series were stored with a sampling frequency of 64 Hz at 10 minutes intervals (CROSS et al., 2013). The data collected in the period from 1 July 2007 to 24 February 2009 (602 observations) were then passed directly to a computer-based system and via a reference-based stochastic subspace identification technique (PEETERS, 1999), the natural frequencies were estimated. The first five natural frequencies obtained during the feature extraction phase are illustrated in Figure 12.

Herein, there is no damaged observations known in advance (KOO et al., 2013), and so it is assumed that all observations are extracted from the undamaged condition. Therefore, only Type I errors can be identified. From a total amount of 602 observations, the first 363 ones are used for statistical modeling in the training process (corresponding to one-year monitoring from 1 July 2007 to 30 June 2008) and the entire data sets are used in the test process, yielding a training matrix $\mathbf{X}^{363\times5}$ (1–363 observations) and a test matrix $\mathbf{Z}^{602\times5}$ (1–602 observations).

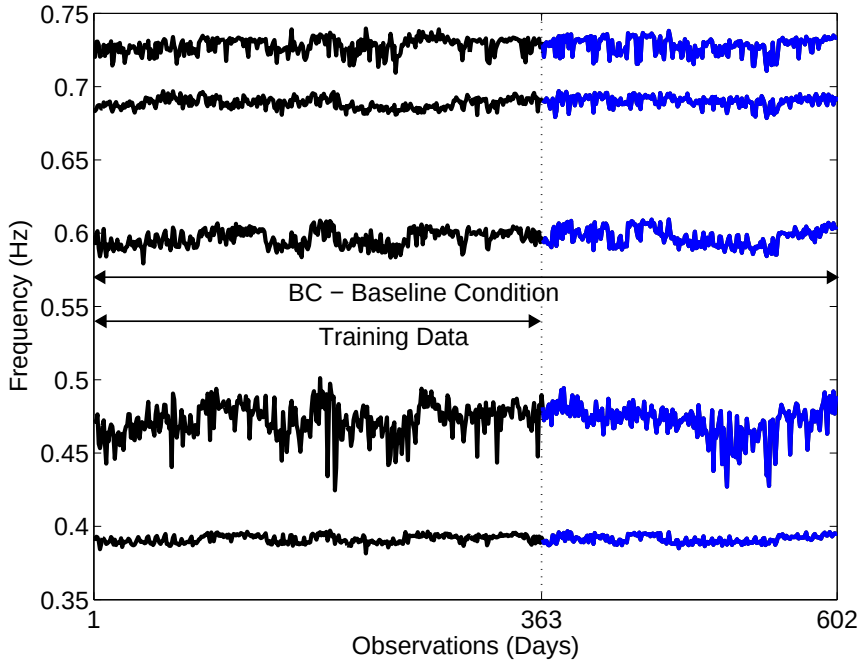

Figure 12 – First five natural frequencies obtained in the Tamar Bridge. The observations in the interval 1–363 are used in the statistical modeling while observation 364-602 are used only in the test phase (FIGUEIREDO et al., 2012).

4.2 Parameter tuning

Although the ACH- and MSD-based approaches do not require any input parameters (except the training matrix), the remain ones work through some predefined parameters. One of the challenges for training deep networks is the large number of design choices, such as connectivity, architecture and optimization method.

The choice of optimization also comes with a number of hyperparameters such as learning rate and momentum. However, some automatic approaches have been developed to addresses many of these issues. Herein, for the DSA-based approach, an undercomplete nine-layer architecture using complete layer-wise connections was adopted, as well as the number of units in each layer was automatically defined on basis of the input data dimension. Therefore, for the Z-24 Bridge, the number of units per layer are 10, 7, 3 and 2, as well as for the Tamar Bridge they are 11, 7, 3 and 2. Note that as the DSA is composed of mapping and demapping layers the number of units in the first three mapping layers are repeated in the last three demapping layers. Additionally, the stop criteria were the number of epochs and gradient convergence, defined as 1000 and 10^{-7}, respectively. All constraints regarding the adopted architecture and parameter setting can be verified in (MAATEN; POSTMA; HERIK, 2009).

For the AANN-based approach, a Levenberg–Marquardt back-propagation algorithm was used to train the network. Several trainings with different initial conditions were performed through approach described in (KRAMER, 1991) to increase the probability that the global minimum was achieved. Therefore, for the Z-24 Bridge, the network has 6 units in each mapping and demapping layer and 4 units in the bottleneck layer. In counterpart, for the Tamar Bridge, the network has 9 units in the mapping and demapping layers, and 8 units in the bottleneck layer. The nodes in the bottleneck layer represent the underlying unobserved variables driving the changes in the features such as the ambient temperature (FIGUEIREDO; CROSS, 2013).

For linear PCA, the amount of variability retained is 95%, which can account approximately all normal variations. In the case of KPCA, the optimal kernel bandwidth can be computed by the maximization of the information entropy (REYNDERS; WURSTEN; ROECK, 2014). Thus, for the Z-24 Bridge, the optimal kernel parameter estimated was 0.384, as well as for the Tamar Bridge the value was 0.0142. In both scenarios, as performed by (SANTOS et al., 2016c), 99% of variance was retained, which can account for approximately all data variations in kernel space.

Finally, for the GMM-based approach, as recommended in (FIGUEIREDO; CROSS, 2013), to alleviate the drawbacks of the EM algorithm, during each execution of the EM, 10 repetitions are performed, each one running until 1000 iterations be reached or the LogL converges. Also, the number of clusters explored via BIC was in the range $K \in 2, \cdots, 15$.

4.3 Damage detection with hourly data set from the Z-24 Bridge

4.3.1 PCA-based approaches

For evaluation of the DSA-based approach with PCA-based ones, the number of Type I and Type II errors for the test matrix are presented in Table 3. Considering a level of significance of 5%, the DSA algorithm outputted the less amount of false alarms (even more than the KPCA-based approach). Specifically, the DSA was able to perform the best performance in terms of normalization of the normal variations with a number of Type I errors less than 5% of the entire undamaged data. It is important to note that when data normalization is well performed the corresponding number of misclassifications on data derived from normal conditions must be 5%, at most, due to the threshold definition. If an algorithm misclassify more than 5% of undamaged data, then it demonstrates problems to learn the normal variations.

Table 3 – Number and percentage of Type I/II errors for PCA-based approaches using the hourly data set from the Z-24 Bridge.

Approach	Type I	Type II	Total
DSA	165 (4.76%)	4 (0.87%)	169 (4.30%)
PCA	173 (4.99%)	372 (80.52%)	545 (13.86%)
AANN	196 (5.65%)	16 (3.46%)	212 (5.39%)
KPCA	180 (5.19%)	4 (0.87%)	184 (4.68%)

The linear PCA provides the worst result in terms of error trade-off establishment, impacting the total amount of errors (more than 13% of misclassifications). Although it demonstrates a reasonable data normalization performance, its unacceptable sensitivity to damage occurrence limits its application on real-world monitoring scenarios, mainly when life-safety issues are a critical effort. On the other hand, the AANN demonstrates problems for modeling the normal condition, resulting in more than 5% of Type I errors and a poor damage classification performance, reaching more than 3% of Type II errors and a total amount up to 5%.

In general terms, the DSA-based approach attains the best results when compared to alternative ones, as can be noted by the improved performance achieved for minimizing the Type I errors, maintaining a reasonable number of Type II errors. Comparing the DSA-based approach to the ones based on AANN and KPCA, it performs a better modeling of normal condition due to the employed architecture and training algorithm. The DSA maps the input data to a greater dimensional space and performs a gradual dimensionality reduction, allowing to better model salient relationships. For SHM, this kind of strength derive a better fitting of normal variations, resulting in robust damage assessment. The two-phase training scheme drives the parameter tuning towards the global

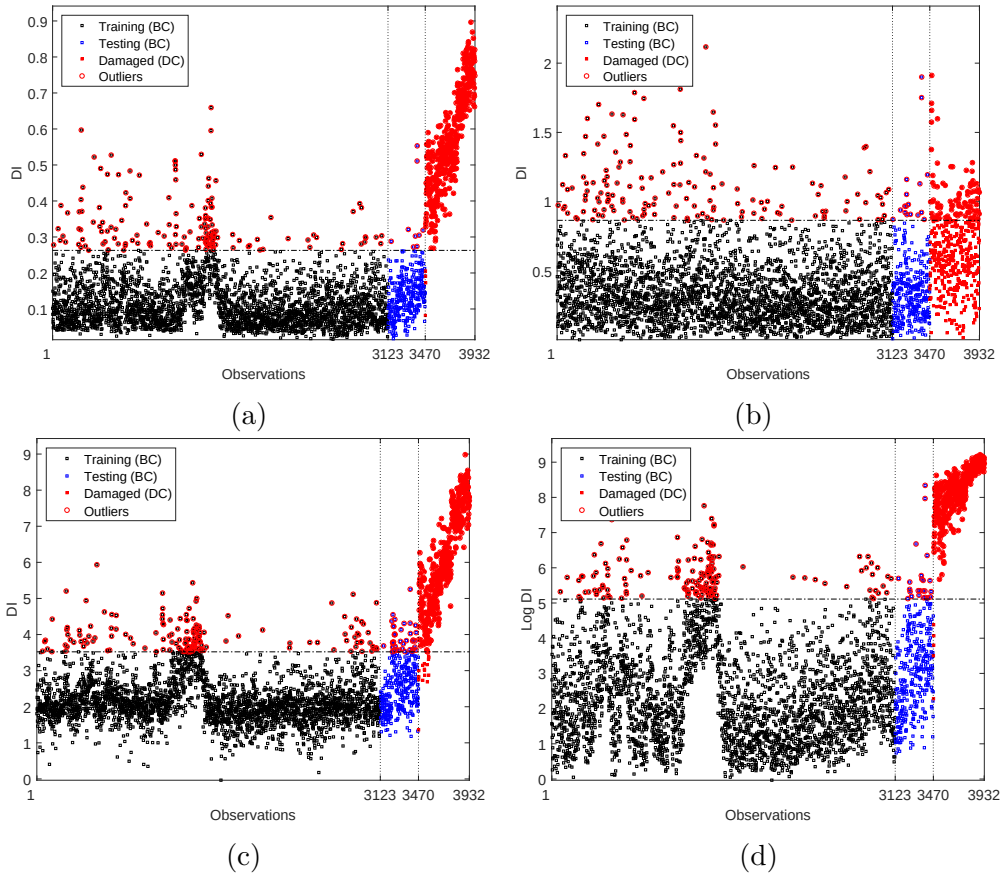

Figure 13 – Damage indicators along with a threshold defined over the training data for Z-24 Bridge: (a) DSA-, (b) PCA-, (c) AANN- and (d) KPCA-based approaches.

optimum, providing a robust model capable to perform with high reliability classification on data not used during training phase.

To verify the performance of DSA-based approach to remove operational and environmental effects, the DIs, derived from the entire test data, are shown in Figure 13. In this case, the only approaches capable to establish a reliable monotonic relationship between the DIs and the gradual increasing in the level of damage were the DSA- and AANN-based approaches. On the other hand, the KPCA fails to establish this relationship, as well as linear PCA. The main reason for their poor normalization performance is related to the freezing effects arising from temperature variations, which cause problems to learn, properly, the normal effects. Thus, such techniques are not recommended to deployment in real applications.

4.3.2 Comparative study of the initialization procedures

To carry out a comparative study, the ACH was independently executed for the three initialization procedures based on random, uniform and divisive strategies, as described in Subsection 3.2.1.1. Table 4 summarizes the Type I and Type II errors for all ACH initializations. The random initialization becomes the algorithm more sensitive to

detect abnormal conditions as expressed by the low number of Type II errors (4); however, it is penalized with a high number of Type I errors (220), demonstrating a loss of generalization capability. An alternative behavior is reached when deterministic initialization procedures are applied. Basically, the uniform initialization demonstrates a high degree of generalization and robustness to fit the normal condition at the cost of losing sensitivity to detect anomalies, as given by a high number of Type II errors (19). On the other hand, the divisive initialization establishes a trade-off between generalization and sensitivity, reaching a low number of Type II errors (6) and maintaining an acceptable number of Type I errors (188), which indicates effectiveness to model the normal condition and to overcome the nonlinear effects. Furthermore, one can figure out that for levels of significance around of 5% both random and divisive initializations are indicated, mainly when the minimization of Type II errors is a critical effort.

Table 4 – Number of clusters and percentage of Type I and Type II errors for each ACH initialization procedure using the hourly data set from the Z-24 Bridge.

Initialization	K	Type I	Type II	Total
Random	5	220 (6.34%)	4 (0.86%)	224 (5.69%)
Uniform	6	159 (4.58%)	19 (4.11%)	178 (4.52%)
Divisive	3	188 (5.41%)	6 (1.29%)	194 (4.93%)

The DIs for the entire test data, along with a threshold defined over the training data, is evidenced in Figure 14, regarding each initialization procedure. If data normalization is well performed, it is statistically guaranteed less than 5% of misclassifications in the DIs derived from undamaged observations not used for training. Therefore, excepting when ACH is initialized with the uniform procedure, it outputs a monotonic relationship in the amplitude of DIs related to the level of degradation accumulated on the bridge along the time. In Figure 14b the freezing effects are highlighted by more evident peaks in the DIs related to the data used in the training phase, indicating that the uniform initialization does not allow an appropriate filtering of nonlinear effects. Note that, a nonlinear effect is not necessarily related to a damaged condition; it can arise from a normal variation of physical parameters of the structure not taken into account during training phase. On the other hand, when damage is presented in the form of an orthogonal component that diverges from the normal condition under common operational and environmental factors, it is detected as a non-observed effect, and thus an anomaly condition.

In relation to the number of data clusters, the ACH was able to find six, five, and three clusters when coupled with uniform, random and divisive initializations, respectively. Furthermore, the random and divisive initializations demonstrate to be more appropriate due to their potential to benefit the clustering step, providing a proper learning of the normal condition, even when operational and environmental variability is present. However,

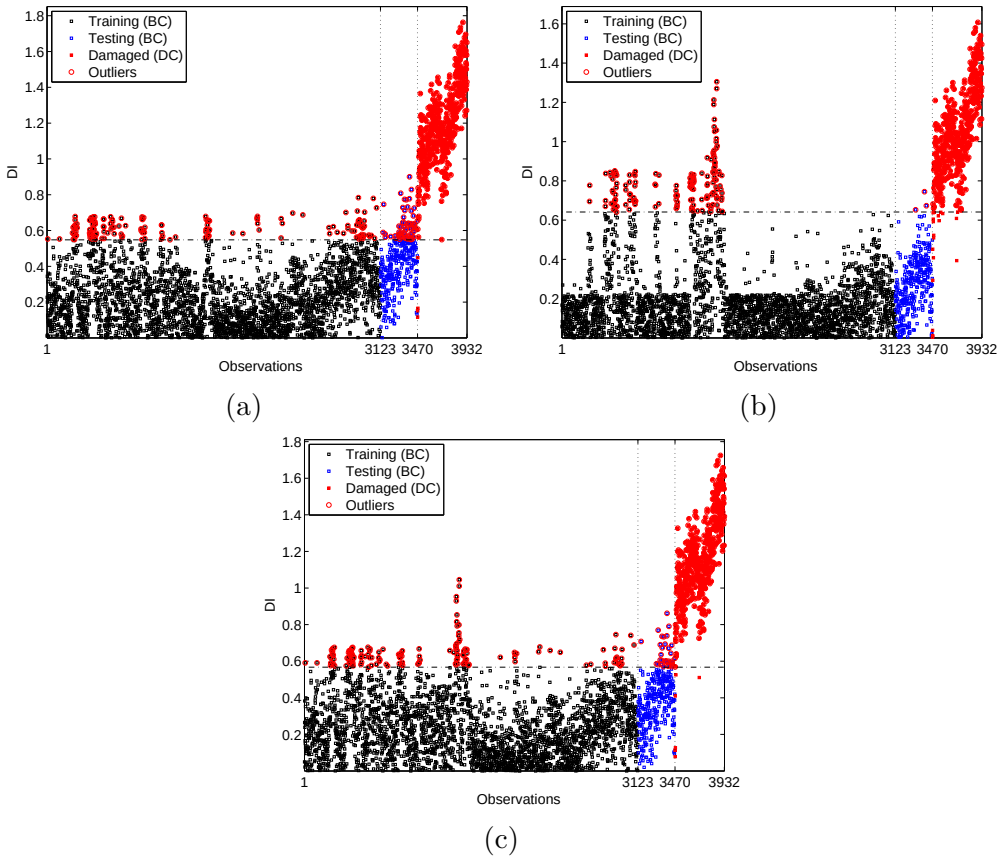

Figure 14 – ACH damage indicators for different initialization procedures along with a threshold defined over the training data: (a) random-, (b) uniform-, and (c) divisive-based initialization procedures.

considering the best model as the one that establishes a trade-off between minimization of the number of errors using less clusters as possible, the divisive initialization is the most suitable model, as it accomplishes reliable results using a small number of clusters, being more indicated when one wants to reach a balance of sensitivity and specificity rates.

4.3.3 Cluster-based approaches

For comparison purposes with well known cluster-based algorithms from the literature, the novel ACH coupled with divisive initialization is chosen to accomplish a

Table 5 – Number of clusters and percentage of Type I/II errors for cluster-based approaches using the hourly data set from the Z-24 Bridge.

Approach	K	Type I	Type II	Total
ACH	3	188 (5.41%)	6 (1.29%)	194 (4.93%)
MSD	1	162 (4.66%)	191 (41.34%)	353 (8.97%)
GMM	7	210 (6.05%)	10 (2.16%)	220 (5.59%)

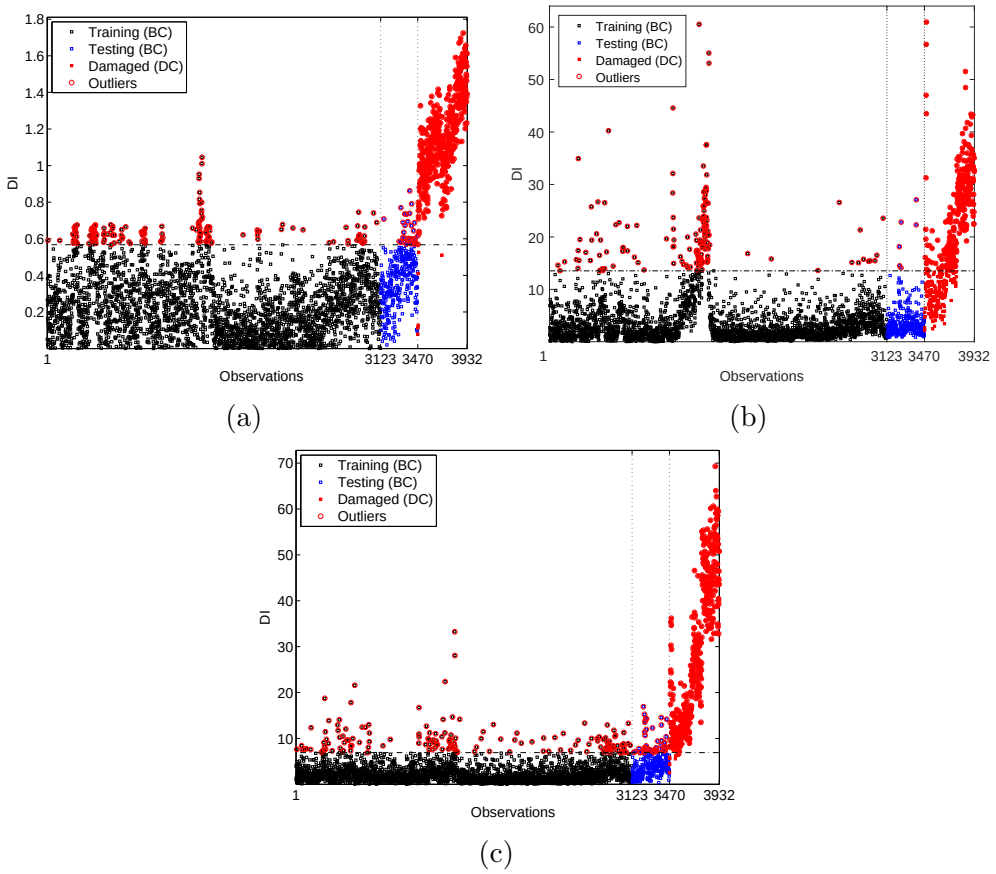

Figure 15 – Damage indicators along with a threshold defined over the training data for Z-24 Bridge: (a) ACH-, (b) MSD-, and (c) GMM-based approaches.

study with MSD- and GMM-based approaches. Therefore, to quantify the classification performance, the Type I and Type II errors for the test matrix are presented in Table 5. Basically, for a level of significance around 5%, the ACH presents the best results in terms of total number of misclassification, less than 5% of the entire test data. In turn, the GMM shows an intermediate performance, attaining 5.59% of misclassified observations. Concerning the ACH and GMM, one can verify that both provide a high sensitivity to damage, although the ACH presents the smaller amount of misclassifications. In terms of generalization, the ACH attains the best results when compared to GMM, as can be inferred by the minimization of Type I errors. Nevertheless, the MSD provides the worst result, misclassifying roughly 9% of the entire test data, demonstrating an inappropriate level of sensitivity to damage, which for high capital expenditure engineering structures is unacceptable due to the drastic consequences it may cause (e.g., undetected failures may cause human losses).

To evaluate the ACH performance to model the normal condition and establish comparisons, the DIs, taking into account the entire test data, are shown in Figure 15. The ACH outputs a monotonic relationship in the amplitude of the DIs related to the damage level accumulation, whereas the GMM fails to stablish this relationship. In the

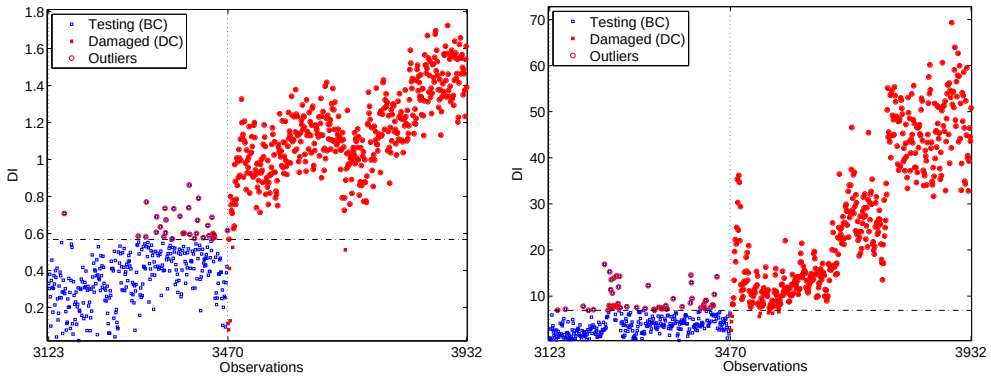

Figure 16 – Emphasis on monotonic relationship between the level of damage and the amplitude of the damage indicators for ACH- (a) and GMM-based (b) approaches.

case of the MSD-based approach, patterns in the DIs caused by the freezing effects can be pointed out, which indicate this approach is not able to attenuate, properly, the effects of environmental variations. Thus, it demonstrates to be not effective to model the normal condition.

Furthermore, for detaching the monotonic relationship derived by ACH, Figure 16 highlights the range of data not used for training purpose. The ACH maintains a constant monotonic relationship with the gradual level of damage and the amplitude of the DIs, even when operational and environmental variability is present; however, the GMM misclassified undamaged observations throughout the track of observations not used in the training phase, indicating a not proper learning of the normal condition. On the other hand, all misclassified undamaged observations accomplished by ACH are grouped in a well known fuzzy region that may exists in the boundary frontiers of the quasi-circular clusters. This is explained by the nature of ACH-based approach. Although the ACH aims to find out radially symmetric clusters, some data groups describe quasi-circular groups of similar observations that present in their decision boundaries sparse regions accomplishing observations deriving to a gradual change of structural state.

In terms of number of clusters, the GMM finds seven clusters ($K = 7$) related to Gaussian components. However, the ACH accomplishes the best results with only three clusters ($K = 3$), indicating the GMM has generalization problems, which can be explained by a tendency of overfitting caused by a high number of clusters. When evaluating cluster-based approaches, a trade-off between good fitting and a low number of clusters is required, as the high number of clusters may lead to an overfitting; conversely, low number of clusters may conduct to an underfitting.

4.4 Damage detection with daily data set from the Tamar Bridge

4.4.1 PCA-based approaches

The data sets from the Z-24 Bridge are unique, as it was known a priori the existence of damage. On the other hand, the data sets from the Tamar Bridge represent the most common situation observed in real-world SHM applications on bridges, as there is no indication of damage in advance.

Following the same procedure carried out in the previous section, the number of Type I errors for the test matrix are presented in Table 6. The Type II errors are not summarized herein as there is no indications about structural damage. The total number of Type I errors is 27 (4.49%), 25 (4.15%), 30 (4.98%) and 73 (12.13%) for the DSA-, PCA-, AANN- and KPCA-based approaches, respectively. Therefore, as the percentage of errors given by the DSA is close to the 5% level of significance assumed in the training process, one concludes that the DSA-based approach offers a model that properly learns the normal variations, as well as linear PCA.

Table 6 – Number and percentage of Type I errors for PCA-based approaches using the daily data set from the Tamar Bridge.

Approach	Type I
DSA	27 (4.49%)
PCA	25 (4.15%)
AANN	30 (4.98%)
KPCA	73 (12.13%)

In contrast to the Z-24 Bridge, the Tamar was not influenced by extreme nonlinear variations, allowing to learn its dynamic behavior using linear models, such as PCA (FIGUEIREDO et al., 2012). Thus, the nonlinear approaches (DSA, AANN and KPCA) tends to overfitting the training data, resulting in poor normalization performance of undamaged data not used for training. However, the DSA was the only method capable to model the normal condition with high degree of reliability, reaching a total amount of Type I errors very close to the ones outputted by linear PCA. It can be pointed out to the robust training algorithm employed in the DSA-based approach, which provides an adequate learning of the normal condition without compromise its generalization capabilities. This characteristic has been noticed as one of the major advantages of two-phase training algorithms for deep neural networks (ERHAN et al., 2010).

For an overall analysis purpose, the DIs for all observations in the test matrix are shown in Figure 17. For the KPCA-based approach, a concentration of outliers in the data not used in the training phase is observed, suggesting an inappropriate modeling of

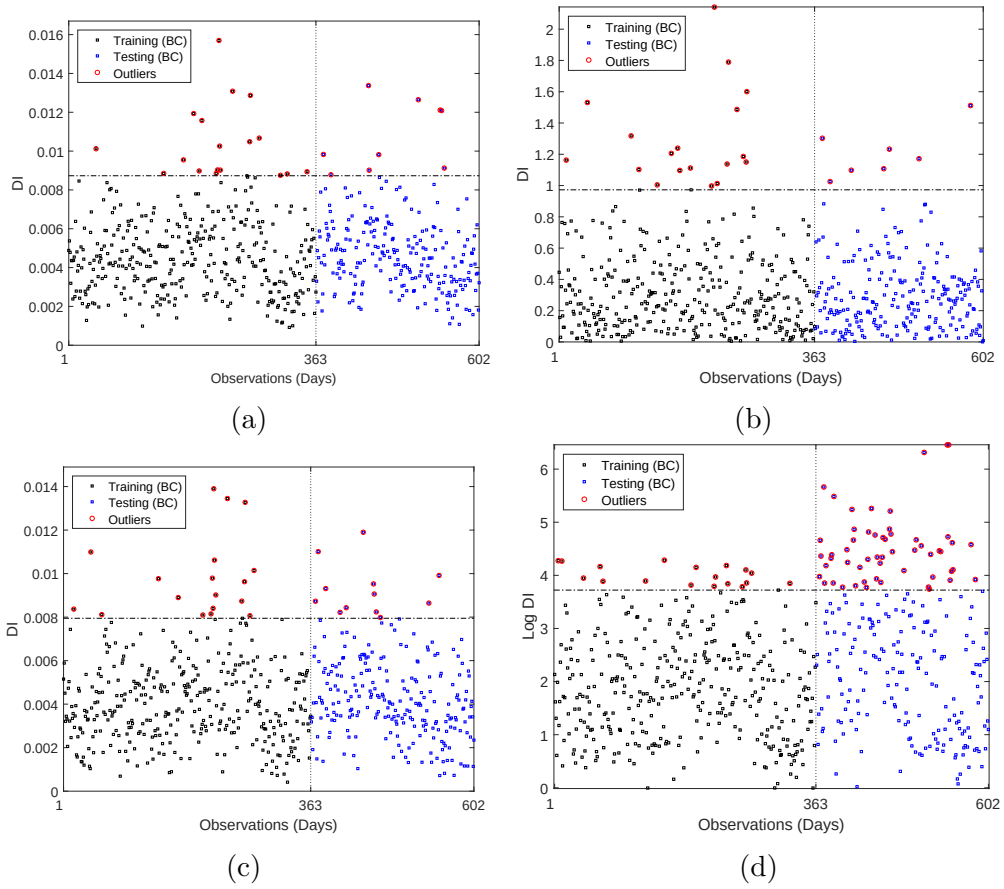

Figure 17 – Damage indicators along with a threshold defined over the training data for Tamar Bridge: (a) DSA-, (b) PCA-, (c) AANN, and (d) KPCA-based approaches.

the normal condition. On the other hand, the DSA-, PCA- and AANN-based approaches seems to output a random pattern among the expected outlier observations, especially among the ones not used in training process, suggesting a properly understanding of the normal condition by the models defined by both algorithms. Note that, in this case, there is no indications about the existence of neither damage nor extreme operational and environmental variability in the data set. Thus, nonlinear patterns are not expected in the corresponding DIs.

Furthermore, the importance of this result is rooted on the fact that this scenario is close to the ones found in real-world monitoring, where there is no indications of damage a priori, which permits one to reduce the number of false alarms and increase the reliability of the SHM system.

4.4.2 Cluster-based approaches

Herein, the same evaluation procedure used previously for Z-24 Bridge is employed. Thus, the ACH-based approach coupled with divisive initialization is compared to the MSD- and GMM-based ones. Therefore, Table 7 summarizes the number of Type I errors for each approach using test matrix. In this case, the MSD has performed the

Table 7 – Number of clusters and percentage of Type I errors for cluster-based approaches using the daily data set from the Tamar Bridge.

Approach	K	Type I
ACH	1	33 (5.48%)
MSD	1	23 (3.82%)
GMM	3	69 (11.46%)

best result in terms of classification performance, reaching 3.82% of Type I errors. However, the ACH outputted a reasonable number of Type I errors (5.48%) by modeling the structural dynamic response as only one data cluster, as well as MSD implicitly assumes. It indicates that the linear behavior of the structure allows to model its normal condition using only linear approaches. As the MSD-based approach learns only linear influences its performance has overcome the alternative methods. Due to the ACH algorithm applies an agglomerative learning procedure it is able to verify this linear behavior by modeling the structural response as only one data cluster. On the other hand, the GMM algorithm, which assumes the data formed by more than one Gaussian component, modeled the normal condition using three clusters in an attempt to learn nonlinear influences. However, as the data is better modeled by only one cluster, the GMM-based approach does not provide an appropriate modeling of normal condition, resulting in more than 11% of misclassifications.

The DIs obtained from the test matrix are highlighted in Figure Figure 18. It shows that for the GMM-based approach, a concentration of outliers in the data not used in the training phase is observed, suggesting an inappropriate modeling of the normal condition. On the other hand, the ACH-based approach seems to output a random pattern among the expected outlier observations, especially among the ones not used in training process, suggesting a properly understanding of the normal condition by the unique cluster defined over the training data. When comparing the ACH- and MSD-based approaches, one can figure out that the MSD has performed a more adequate filtering of the normal variations than ACH. This behavior is expected due to the properties of the data set. As demonstrated by (FIGUEIREDO et al., 2012), the data arising from Tamar Bridge follows, approximately, a unique multivariate normal distribution, benefiting the modeling by MSD. Although the ACH does not perform any normal assumption, its general results demonstrate high reliability, leading a reasonable number of misclassifications by learning a simple model and verifying the same linear relationship as MSD.

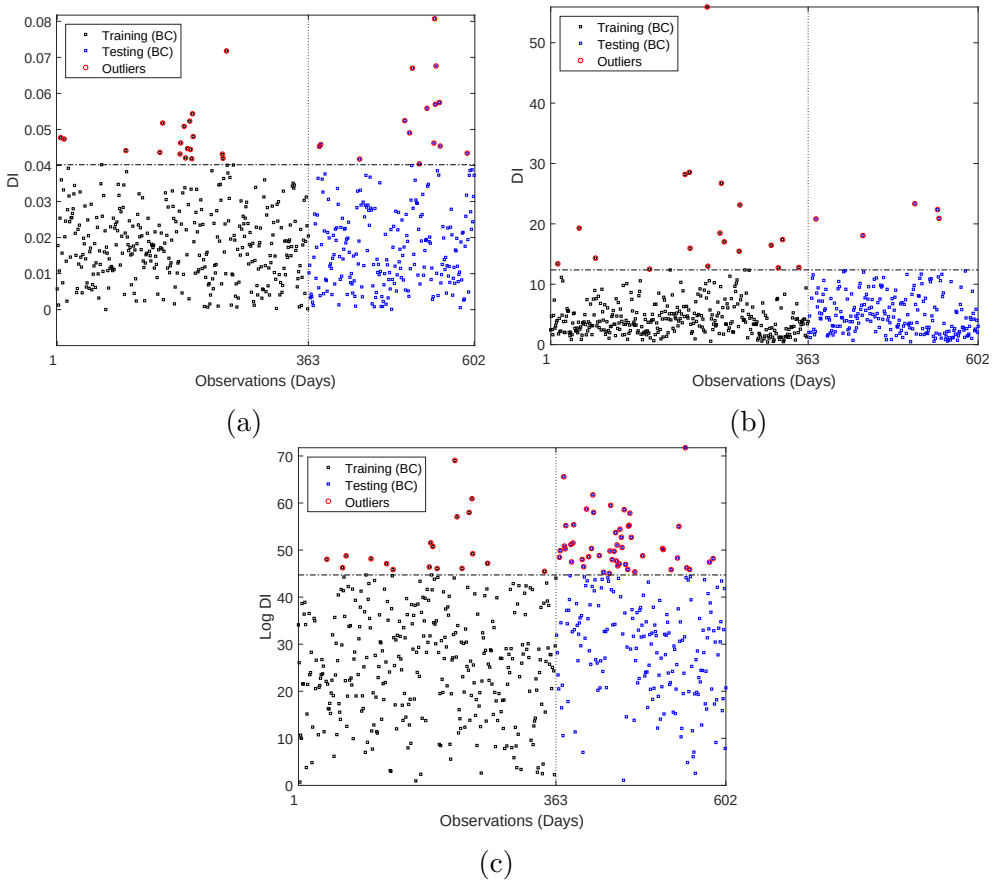

Figure 18 – Damage indicators along with a threshold defined over the training data for Tamar Bridge: (a) ACH-, (b) MSD-, and (c) GMM-based approaches.

4.5 Overall analysis

In general, for the Z-24 Bridge, one can figure out that the proposed approaches have outperformed the corresponding state-of-the-art ones, in terms of minimization of misclassifications and removing of nonlinear effects derived, mainly, from colder periods. In this case, both the DSA- and ACH-based approaches performed the better balance between Type I/II errors, as can be seen in Table 8. The KPCA- and GMM-based approaches tried to establish the same balance, however the limitations imposed by their model properties embarrass the proper learning of normal conditions. In the case of the novel cluster-based approach, it results can be pointed out as the remarkable one, as it reaches better results compared to the traditional ones based on PCA and cluster analysis using a very simple model. Regarding the DSA-based approach, its modeling demonstrated to be the best one in terms of minimization of Type I/II errors, performing even better than ACH-based approach. However, in comparison to the ACH, the DSA model does not provide any intuitions about the sources of variability, in contrast to the novel approach that allows to infer the main global structural state conditions of the structure by clustering similar observations regarded to the main influences at a present time period.

Table 8 – Comparison with the best machine learning approaches.

Data set	Approach	Type I	Type II	Total
Z-24 Bridge	DSA	165 (4.76%)	4 (0.87%)	169 (3.30%)
	ACH	188 (5.41%)	6 (1.29%)	194 (4.93%)
	KPCA	180 (5.19%)	4 (0.87%)	184 (4.68%)
Tamar Bridge	DSA	27 (4.49%)	———	27 (4.49%)
	ACH	33 (5.48%)	———	33 (5.48%)
	PCA	25 (4.15%)	———	25 (4.15%)
	MSD	23 (3.82%)	———	23 (3.82%)

One important note is the carried out analysis are based on the fact that for real-world applications a suitable approach should not only minimize the false-negative indications of damage, for safety issues, but also minimize as much as possible the number of false-positive indications of damage, regarding economic factors.

On the other hand, from the Tamar Bridge results, one can infer that the absence of nonlinear variations has embarrassed the modeling performed by nonlinear methods, such as KPCA-, AANN- and GMM-based approaches, which are not robust enough to learn linear variations without overfit the training data. On the other hand, the linear approaches, based on PCA and MSD have learned, adequately, the normal conditions. This behavior is a natural one due to the specific purposes derived for each technique (e.g., nonlinear approaches to model nonlinearities and linear ones to model linear influences). However, the approaches based on DSA and ACH reached a similar performance than the ones based on linear PCA and MSD, allowing to conclude that the proposed techniques provide a robust modeling capable to learn linear and nonlinear influences with similar ability. Thus, both approaches demonstrate to be general purpose methods, which can be employed in scenarios where the structural manager has few informations regarding the sources of variability and the actual condition of structure.

5 Conclusions, future research and published works

5.1 Main conclusions

The correct monitoring and evaluation of structures with an arbitrary level of complexity depends mainly on robust approaches to separate the changes in sensor readings caused by damage from those caused by changing operational and environmental conditions. Herein, in an effort to address the issues related to this challenge, novel machine learning algorithms were proposed for modeling the normal condition of structures by posing the SHM as an SPR paradigm. Therefore, this dissertation is mainly concerned with damage detection by means of output-only approaches. Even though these approaches can be employed for any kind of structures, the methods were especially posed for SHM bridge applications.

Specifically, this work significantly contributes to the SHM field by proposing the first application of deep learning algorithms in the context of data normalization and feature classification, as well as it proposes a novel and original agglomerative procedure for clustering the normal condition by inflation of concentric hyperspheres. This novel agglomerative concentric hypersphere (ACH) algorithm evaluates the spacial geometry and sample density of each cluster, and can be though as an important advance to cluster-based methods, as it does not require any input parameters (except training matrix) and excludes the need for measures related to the sources of variability. Furthermore, it does not assume any particular distribution to the data, allowing its application in real-world monitoring scenarios without prior knowledge from the type of structural response. On the other hand, the applied deep stacked autoencoder (DSA) arises as an improved algorithm to perform NLPCA, circumventing the drawbacks related to the traditional PCA-based approaches (e.g., data normality assumptions, definition of hyperparameters and issues related to the model of normal condition in structures with different behaviors when subjected to linear/nonlinear variations). Both algorithms have overcome some of the best approaches available in literature up to this moment.

The proposed approaches were compared with traditional ones based on PCA (linear PCA, AANN and KPCA) and cluster analysis (MSD and GMM), through their application on two conceptually different but real-world data sets, from the Z-24 and Tamar Bridges, located in Switzerland and United Kingdom, respectively. The structures

were subjected to strong known environmental and operational influences, which cause structural changes due mainly to nonlinear effects of freezing and boundary conditions like thermal expansions and contractions.

In terms of overall analysis, as verified on the test bed structures, the proposed approaches demonstrate to be: (i) as robust as their respective traditional ones to detect the existence of damage; and (ii) potentially more effective to model the baseline condition and to remove the effects of the operational and environmental variability, as suggested by the minimization of misclassifications on the data from both structures. In a global perspective is concluded that DSA has the best classification performance in terms of minimization of Type I/II errors, indicating to be the most appropriated method when the main goal is attenuate misclassifications. Additionally, the ACH algorithm also demonstrated high reliability to model the normal condition and minimize misclassifications.

Even though that the DSA provides the most robust model, the ACH has the advantage to provide a model capable to carry out physical interpretations related to the sources of variability that alter the structural responses (i.e., each cluster can be related to a different source of variability changing the structural properties at a given period). In this context, the DSA provides a black-box model, which does not provide physical meanings, and not even contribute to the increasing of knowledge related to the nature and behavior of the structure. At this point, we can pose the ACH algorithm in the context of the well-known theorem that *there is no free lunch*, in which machine learning algorithms are classified in one of two classes: (1) specialized methods for some category of problems and (2) methods that maintain a reasonable performance in the solution of most part of problems. Thus, the ACH fits the category in which results are often acceptable, independently of the structure complexity, with an addendum of providing a model that allows physical interpretation. The chosen method for deployment in real monitoring scenarios depends highly on the monitoring goals (e.g., minimization of misclassifications is the only concern or model interpretation is also a major issue).

5.2 Future research topics

In the future it is intended to further explore the capabilities of deep learning algorithms for feature extraction, as a manner to diminish the level of expertise required to extract damage-sensitive features, which currently varies from the kind of structure and nature of damage to be detected. Also, the deep neural networks can be used in the context of time-series analysis for prediction of unusual behaviors that may be related to damage occurrence in a structural system.

For the ACH-based approach, novel distance metrics and density functions can be

used to improve the estimation of data clusters, as well as increase the damage detection performance. Other real-world or simulated data sets with different type of normal effects and damage can be tested using the proposed approaches, as a manner to verify their general performance to overcome most kind of operational and environmental influences.

5.3 Published works

The main original works that support this dissertation are addressed in follow. The articles are organized in accordance to the period of publication and type of media (journal or conference).

Original works published in international journals:

1. MOISÉS SILVA; ADAM SANTOS; REGINALDO SANTOS; ELOI FIGUEIREDO; CLAUDOMIRO SALES; JOÃO C. W. A. COSTA. "Agglomerative concentric hypersphere clustering applied to structural damage detection", Mechanical Systems and Signal Processing, 2017.

2. ADAM SANTOS; REGINALDO SANTOS; MOISÉS SILVA; ELOI FIGUEIREDO; CLAUDOMIRO SALES; JOÃO C. W. A. COSTA. "A Global Expectation-Maximization Approach Based on Memetic Algorithm for Vibration-Based Structural Damage Detection", IEEE Transactions on Instrumentation and Measurement, 2017.

3. ADAM SANTOS; MOISÉS SILVA; REGINALDO SANTOS; ELOI FIGUEIREDO; CLAUDOMIRO SALES; JOÃO C. W. A. COSTA. "A global expectation-maximization based on memetic swarm optimization for structural damage detection", Structural Health Monitoring, 2016.

4. ADAM SANTOS; ELOI FIGUEIREDO; MOISÉS SILVA; REGINALDO SANTOS; CLAUDOMIRO SALES; JOÃO C. W. A. COSTA. "Genetic-based EM algorithm to improve the robustness of Gaussian mixture models for damage detection in bridges", Structural Control & Health Monitoring, 2016.

5. MOISÉS SILVA; ADAM SANTOS; ELOI FIGUEIREDO; REGINALDO SANTOS; CLAUDOMIRO SALES; JOÃO C. W. A. COSTA. "A novel unsupervised approach based on a genetic algorithm for structural damage detection in bridges", Engineering Applications of Artificial Intelligence, 2016.

Original works published in international conferences:

1. MOISÉS SILVA; ADAM SANTOS; REGINALDO SANTOS; ELOI FIGUEIREDO; CLAUDOMIRO SALES; JOÃO C. W. A COSTA. "A structural damage detection

technique based on agglomerative clustering applied to the Z-24 Bridge", 8th European Workshop on Structural Health Monitoring (EWSHM), 2016, Bilbao.

2. ADAM SANTOS; MOISÉS SILVA; REGINALDO SANTOS; ELOI FIGUEIREDO; CLAUDOMIRO SALES; JOÃO C. W. A. COSTA. "Output-only structural health monitoring based on mean shift clustering for vibration-based damage detection", 8th European Workshop on Structural Health Monitoring (EWSHM), 2016, Bilbao.

Bibliography

BENGIO, Y. Learning deep architectures for ai. *Found. Trends Mach. Learn.*, Now Publishers Inc., Hanover, MA, USA, v. 2, n. 1, p. 1–127, Jan 2009. ISSN 1935-8237. Cited on page 20.

BENGIO, Y. et al. Greedy layer-wise training of deep networks. In: SCHöLKOPF, P. B.; PLATT, J. C.; HOFFMAN, T. (Ed.). *Advances in Neural Information Processing Systems 19*. [S.l.]: MIT Press, 2007. p. 153–160. Cited 2 times on pages 20 and 22.

BERRY, M. J.; LINOFF, G. *Data Mining Techniques: For Marketing, Sales, and Customer Support*. New York, NY, USA: John Wiley & Sons, Inc., 1997. ISBN 0471179809. Cited on page 26.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. *Time Series Analysis: Forecasting and Control*. 4th. ed. Hoboken NJ, United States: John Wiley & Sons, Inc., 2008. Cited on page 19.

BREIMAN, L. Stacked regressions. *Machine Learning*, v. 24, n. 1, p. 49–64, 1996. ISSN 1573-0565. Cited on page 22.

CATBAS, F. N.; GOKCE, H. B.; GUL, M. Nonparametric analysis of structural health monitoring data for identification and localization of changes: Concept, lab, and real-life studies. *Structural Health Monitoring*, v. 11, n. 5, p. 613–626, July 2012. Cited 3 times on pages 1, 2, and 3.

CHANG, C.-C.; LIN, C.-J. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, v. 2, n. 3, p. 1–27, 2011. Cited on page 16.

CHENG, L. et al. The health monitoring method of concrete dams based on ambient vibration testing and kernel principle analysis. *Journal of Shock and Vibration*, v. 2015, 2015. ISSN 1875-9203. Cited on page 5.

CROSS, E. et al. Long-term monitoring and data analysis of the tamar bridge. *Mechanical Systems and Signal Processing*, v. 35, n. 1–2, p. 16 – 34, 2013. ISSN 0888-3270. Cited 2 times on pages 34 and 35.

D'ANGELO, G.; RAMPONE, S. Feature extraction and soft computing methods for aerospace structure defect classification. *Measurement*, v. 85, p. 192–209, 2016. Cited on page 2.

DEMERS, D.; COTTRELL, G. W. Non-linear dimensionality reduction. In: *Advances in Neural Information Processing Systems 5*. [S.l.: s.n.], 1992. p. 580–587. Cited on page 22.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 39, n. 1, p. 1–38, 1977. Cited 2 times on pages 18 and 19.

DERAEMAEKER, A. et al. Vibration-based structural health monitoring using output-only measurements under changing environment. *Mechanical Systems and Signal Processing*, v. 22, n. 1, p. 34 – 56, 2008. ISSN 0888-3270. Cited on page 4.

DOEBLING, S. et al. *Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: A literature review*. [S.l.], 1996. Cited on page 11.

ERHAN, D. et al. Why does unsupervised pre-training help deep learning? *Machine Learning Research*, v. 11, p. 625–660, Mar 2010. ISSN 1532-4435. Cited 3 times on pages 20, 21, and 43.

FARRAR, C.; WORDEN, K. *Structural Health Monitoring: A Machine Learning Perspective*. Hoboken NJ, United States: John Wiley & Sons, Inc., 2013. Cited 2 times on pages 1 and 10.

FARRAR, C. R.; DOEBLING, S. W.; NIX, D. A. Vibration-based structural damage identification. *Philosophical Transactions of the Royal Society: Mathematical, Physical & Engineering Sciences*, v. 359, n. 1778, p. 131–149, 2001. Cited on page 2.

FARRAR, C. R.; SOHN, H.; WORDEN, K. *Data Normalization: A Key For Structural Health*. [S.l.], 2001. 12 p. Cited on page 2.

FIGUEIREDO, E.; CROSS, E. Linear approaches to modeling nonlinearities in long-term monitoring of bridges. *Journal of Civil Structural Health Monitoring*, v. 3, n. 3, p. 187–194, 2013. Cited 4 times on pages 3, 5, 6, and 36.

FIGUEIREDO, E.; MOLDOVAN, I.; MARQUES, M. B. *Condition Assessment of Bridges: Past, Present, and Future - A Complementary Approach*. Portugal: Universidade Católica Editora, 2013. Cited on page 1.

FIGUEIREDO, E. et al. Machine learning algorithms for damage detection under operational and environmental variability. *Structural Health Monitoring*, v. 10, n. 6, p. 559–572, 2011. Cited 2 times on pages 2 and 4.

FIGUEIREDO, E. et al. Applicability of a markov-chain monte carlo method for damage detection on data from the z-24 and tamar suspension bridges. *Proceedings of the 6th European Workshop on Structural Health Monitoring*, p. 747–754, 2012. Cited 4 times on pages IX, 35, 43, and 45.

FIGUEIREDO, E. et al. A Bayesian approach based on a Markov-chain Monte Carlo method for damage detection under unknown sources of variability. *Engineering Structures*, v. 80, n. 0, p. 1–10, 2014. Cited 4 times on pages IX, XI, 18, and 33.

FIGUEIREDO, E. et al. A Bayesian approach based on a Markov-chain Monte Carlo method for damage detection under unknown sources of variability. *Engineering Structures*, v. 80, n. 0, p. 1–10, 2014. Cited 3 times on pages IX, 6, and 33.

FIGUEIREDO, E. J. F. *Damage Identification in Civil Engineering Infrastructure under Operational and Environmental Conditions*. PhD thesis (Doctor of Philosophy in Civil Engineering) — Faculdade de Engenharia, Universidade do Porto, Porto, Portugal, 2010. Cited 3 times on pages IX, 10, and 17.

FIGUEIREDO, M. A. T.; JAIN, A. K. Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 24, n. 3, p. 381–396, 2002. Cited on page 19.

GHOSH, S. K. D. S. Comparative analysis of k-means and fuzzy c-means algorithms. *International Journal of Advanced Computer Science and Applications*, v. 4, n. 4, 2013. Cited on page 63.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. Cited 2 times on pages IX and 21.

HAMERLY, G.; ELKAN, C. Learning the k in k-means. In: *In Neural Information Processing Systems*. [S.l.]: MIT Press, 2003. p. 2003. Cited on page 29.

HINTON, G. E.; SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. *Science*, v. 313, n. 5786, p. 504–507, Jul 2006. Cited 2 times on pages 20 and 22.

HOSSEINABADI, H. Z. et al. Wavelet network approach for structural damage identification using guided ultrasonic waves. *IEEE Transactions on Instrumentation and Measurement*, v. 63, n. 7, p. 1680–1692, Jul 2014. Cited on page 2.

HSU, T.-Y.; LOH, C.-H. Damage detection accommodating nonlinear environmental effects by nonlinear principal component analysis. *Structural Control and Health Monitoring*, v. 17, n. 3, p. 338–354, 2010. Cited on page 3.

HU, W.-H. et al. Comparison of different statistical approaches for removing environmental/operational effects for massive data continuously collected from footbridges. *Structural Control and Health Monitoring*, 2016. ISSN 1545-2263. Cited 2 times on pages 1 and 3.

JOLLIFFE, I. *Principal Component Analysis*. [S.l.]: 2nd Edition. Springer, 2002. Cited on page 13.

KAUFMAN, P. J. R. L. *Finding groups in data: an introduction to cluster analysis*. 9th. ed. [S.l.]: Wiley-Interscience, 1990. ISBN 9780471878766,0471878766. Cited on page 26.

KEERTHI, S. S.; LIN, C.-J. Asymptotic Behaviors of Support Vector Machines with Gaussian Kernel. *Neural Computation*, v. 15, n. 7, p. 1667–1689, 2003. Cited on page 15.

KHOA, N. L. et al. Robust dimensionality reduction and damage detection approaches in structural health monitoring. *Structural Health Monitoring*, 2014. Cited on page 5.

KOO, K. Y. et al. Structural health monitoring of the Tamar suspension bridge. *Structural Control and Health Monitoring*, v. 20, n. 4, p. 609–625, 2013. Cited 2 times on pages 34 and 35.

KRAMER, M. A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, v. 37, n. 2, p. 233–243, 1991. Cited 2 times on pages 4 and 36.

KULLAA, J. Eliminating Environmental or Operational Influences in Structural Health Monitoring using the Missing Data Analysis. *Journal of Intelligent Material Systems and Structures*, v. 20, n. 11, p. 1381–1390, Jul 2009. Cited on page 2.

KULLAA, J. Distinguishing between sensor fault, structural damage, and environmental or operational effects in structural health monitoring. *Mechanical Systems and Signal Processing*, v. 25, n. 8, p. 2976 – 2989, 2011. ISSN 0888-3270. Cited on page 2.

LAORY, I.; TRINH, T. N.; SMITH, I. F. Evaluating two model-free data interpretation methods for measurements that are influenced by temperature. *Advanced Engineering Informatics*, v. 25, n. 3, p. 495 – 506, 2011. ISSN 1474-0346. Special Section: Engineering informatics in port operations and logistics. Cited on page 3.

MAATEN, L. V. D.; POSTMA, E.; HERIK, J. Van den. *Dimensionality reduction: a comparative review*. Tilburg, Netherlands, 2009. Cited 2 times on pages 23 and 36.

MACQUEEN, J. B. Some methods for classification and analysis of multivariate observations. In: CAM, L. M. L.; NEYMAN, J. (Ed.). *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability*. [S.l.]: University of California Press, 1967. v. 1, p. 281–297. Cited on page 29.

MAIMON, O.; ROKACH, L. *Data Mining and Knowledge Discovery Handbook*. 2. ed. [S.l.]: Springer US, 2010. (Texts and Monographs in Physics). Cited on page 26.

MALHI, A.; GAO, R. X. PCA-based feature selection scheme for machine defect classification. *IEEE Transactions on Instrumentation and Measurement*, v. 53, n. 6, p. 1517–1525, Dec 2004. Cited on page 3.

MALHI, A.; YAN, R.; GAO, R. X. Prognosis of Defect Propagation Based on Recurrent Neural Networks. *IEEE Transactions on Instrumentation and Measurement*, v. 60, n. 3, p. 703–711, March 2011. Cited on page 3.

MANNING, C. D.; RAGHAVAN, P.; SCHüTZE, H. *Introduction to Information Retrieval*. New York, NY, USA: Cambridge University Press, 2008. ISBN 0521865719, 9780521865715. Cited on page 26.

MCLACHLAN, G. J.; PEEL, D. *Finite Mixture Models*. United States: John Wiley & Sons, Inc., Wiley Series in Probability and Statistics, 2000. Cited on page 18.

MITA, A.; HAGIWARA, H. Quantitative damage diagnosis of shear structures using support vector machine. *KSCE Journal of Civil Engineering*, v. 7, n. 6, p. 683–689, 2003. ISSN 1976-3808. Cited on page 5.

NGUYEN, T.; CHAN, T. H.; THAMBIRATNAM, D. P. Controlled Monte Carlo data generation for statistical damage identification employing Mahalanobis squared distance. *Structural Health Monitoring*, v. 13, n. 4, p. 461–472, 2014. Cited on page 5.

OH, C. K.; SOHN, H.; BAE, I.-H. Statistical novelty detection within the yeongjong suspension bridge under environmental and operational variations. *Smart Materials and Structures*, v. 18, n. 12, p. 125022, 2009. Cited on page 5.

PEETERS, B. *System Identification and Damage Detection in Civil Engineering*. PhD thesis (Doctorate) — Katholieke Universiteit Leuven, 1999. Cited on page 35.

PEETERS, B.; ROECK, G. D. Reference-based stochastic subspace identification for output-only modal analysis. *Mechanical Systems and Signal Processing*, v. 13, n. 6, p. 855 – 878, 1999. ISSN 0888-3270. Cited on page 32.

PEETERS, B.; ROECK, G. D. One-year monitoring of the Z24-Bridge: environmental effects versus damage events. *Earthquake Engineering & Structural Dynamics*, v. 30, n. 2, p. 149–171, Feb 2001. Cited 2 times on pages 2 and 34.

PEETERS J. MAECK, G. d. R. B. Vibration-based damage detection in civil engineering: excitation sources and temperature effects. *Smart Materials and Structures*, v. 10, p. 518–27, 2001. Cited 2 times on pages 31 and 32.

PURARJOMANDLANGRUDI, A.; GHAPANCHI, A. H.; ESMALIFALAK, M. A data mining approach for fault diagnosis: An application of anomaly detection algorithm. *Measurement*, v. 55, p. 343–352, 2014. Cited on page 3.

RANZATO, M. aurelio et al. Efficient learning of sparse representations with an energy-based model. In: SCHöLKOPF, P. B.; PLATT, J. C.; HOFFMAN, T. (Ed.). *Advances in Neural Information Processing Systems 19*. [S.l.]: MIT Press, 2007. p. 1137–1144. Cited on page 20.

REYNDERS, E.; WURSTEN, G.; ROECK, G. D. Output-only structural health monitoring in changing environmental conditions by means of nonlinear system identification. *Structural Health Monitoring*, v. 13, n. 1, p. 82–93, Jan 2014. Cited 6 times on pages 3, 5, 7, 15, 16, and 36.

SANTOS, A. et al. Genetic-based em algorithm to improve the robustness of gaussian mixture models for damage detection in bridges. *Structural Control and Health Monitoring*, p. 00–00, 2016. ISSN 1545-2263. Cited on page 6.

SANTOS, A. et al. Machine learning algorithms for damage detection: Kernel-based approaches. *Journal of Sound and Vibration*, v. 363, p. 584–599, 2016. Cited on page 5.

SANTOS, A. et al. A global expectation-maximization based on memetic swarm optimization for structural damage detection. *Structural Health Monitoring*, v. 15, n. 5, p. 610–625, 2016. Cited on page 36.

SANTOS, A. D. F. et al. Applicability of linear and nonlinear principal component analysis for damage detection. In: *Proc. IEEE I2MTC*. Pisa, Italy: [s.n.], 2015. p. 869–875. Cited 2 times on pages 3 and 24.

SCHOLKOPF, B.; SMOLA, A.; MULLER, K.-R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, v. 10, n. 5, p. 1299–1319, 1998. Cited 2 times on pages IX and 16.

SCHOLKOPF, B.; SMOLA, A. J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge MA, United States: The MIT Press, 2001. Cited on page 17.

SHAO, R. et al. The fault feature extraction and classification of gear using principal component analysis and kernel principal component analysis based on the wavelet packet transform. Cited on page 3.

SHLENS, J. *A tutorial on Principal Component Analysis*. [S.l.], 2002. Cited on page 4.

SILVA, S. da et al. Structural damage detection by fuzzy clustering. *Mechanical Systems and Signal Processing*, v. 22, n. 7, p. 1636 – 1649, 2008. ISSN 0888-3270. Cited on page 6.

SOHN, H. Effects of environmental and operational variability on structural health monitoring. *Philosophical Transactions of the Royal Society: Mathematical, Physical & Engineering Sciences*, v. 365, n. 1851, p. 539–560, 2007. Cited on page 2.

SOHN, H. et al. A review of structural health monitoring literature: 1996–2001. *Los Alamos National Laboratory, Los Alamos, NM*, 2004. Cited on page 11.

SOHN, H.; WORDEN, K.; FARRAR, C. R. Statistical Damage Classification Under Changing Environmental and Operational Conditions. *Journal of Intelligent Material Systems and Structures*, v. 13, n. 9, p. 561–574, 2002. Cited 4 times on pages IX, 2, 14, and 15.

VAN, M.; KANG, H.-J. Wavelet Kernel Local Fisher Discriminant Analysis With Particle Swarm Optimization Algorithm for Bearing Defect Classification. *IEEE Transactions on Instrumentation and Measurement*, v. 64, n. 12, p. 3588–3600, Dec 2015. Cited on page 2.

VINCENT, P. et al. Extracting and composing robust features with denoising autoencoders. In: *Proceedings of the 25th International Conference on Machine Learning*. New York, NY, USA: ACM, 2008. p. 1096–1103. Cited on page 20.

WENZEL, H. *Health Monitoring of Bridges*. United States: John Wiley & Sons, Inc., 2009. Cited on page 1.

WESTON, J. et al. *Deep Learning via Semi-supervised Embedding*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. 639–655 p. Cited on page 20.

WOLPERT, D. H. Stacked generalization. *Neural Networks*, v. 5, n. 2, p. 241–259, Feb 1992. ISSN 0893-6080. Cited on page 22.

WORDEN, K. Structural fault detection using a novelty measure. *Journal of Sound and Vibration*, v. 201, n. 1, p. 85 – 101, 1997. ISSN 0022-460X. Cited 2 times on pages 4 and 17.

WORDEN, K. et al. Structural health monitoring: from structures to systems-of-systems. *IFAC-PapersOnLine*, v. 48, n. 21, p. 1 – 17, 2015. ISSN 2405-8963. Cited 2 times on pages 1 and 3.

WORDEN, K. et al. The fundamental axioms of structural health monitoring. *Philosophical Transactions of the Royal Society: Mathematical, Physical & Engineering Sciences*, v. 463, n. 2082, p. 1639–1664, 2007. Cited on page 1.

WORDEN, K.; MANSON, G. The application of machine learning to structural health monitoring. *Philosophical Transactions of the Royal Society: Mathematical, Physical & Engineering Sciences*, v. 365, n. 1851, p. 515–537, 2007. Cited 2 times on pages 2 and 5.

WORDEN, K.; MANSON, G.; ALLMAN, D. Experimental validation of a structural health monitoring methodology: Part i. novelty detection on a laboratory structure. *Journal of Sound and Vibration*, v. 259, n. 2, p. 323 – 343, 2003. ISSN 0022-460X. Cited on page 17.

WORDEN, K.; MANSON, G.; FIELLER, N. R. J. Damage detection using outlier analysis. *Journal of Sound and Vibration*, v. 229, n. 3, p. 647–667, 2000. Cited on page 3.

XIANG, J.; ZHONG, Y.; GAO, H. Rolling element bearing fault detection using PPCA and spectral kurtosis. *Measurement*, v. 75, p. 180–191, 2015. Cited on page 3.

YAN, A.-M. et al. Structural damage diagnosis under varying environmental conditions - part i: A linear analysis. *Mechanical Systems and Signal Processing*, v. 19, n. 4, p. 847 – 864, 2005. ISSN 0888-3270. Cited on page 4.

YAN, A.-M. et al. Structural damage diagnosis under varying environmental conditions - part ii: local pca for non-linear cases. *Mechanical Systems and Signal Processing*, v. 19, n. 4, p. 865 – 880, 2005. ISSN 0888-3270. Cited on page 4.

YAN, A.-M. et al. Structural damage diagnosis under varying environmental conditions-Part I: A linear analysis. *Mechanical Systems and Signal Processing*, v. 19, n. 4, p. 847–864, Jul 2005. Cited on page 3.

YUQING, Z. et al. Nc machine tools fault diagnosis based on kernel pca and -nearest neighbor using vibration signals. *Journal of Shock and Vibration*, v. 2015, 2015. ISSN 1875-9203. Cited on page 5.

ZHOU, H.; NI, Y.; KO, J. Structural damage alarming using auto-associative neural network technique: Exploration of environment-tolerant capacity and setup of alarming threshold. *Mechanical Systems and Signal Processing*, v. 25, n. 5, p. 1508 – 1526, 2011. ISSN 0888-3270. Cited on page 4.

ZHOU, H. et al. Modeling of wind and temperature effects on modal frequencies and analysis of relative strength of effect. *Wind and Structures*, v. 11, n. 1, p. 35–50, Jan 2008. Cited on page 2.

ZHOU, Y.-L. et al. Damage detection in structures using a transmissibility-based mahalanobis distance. *Structural Control and Health Monitoring*, v. 22, n. 10, p. 1209–1222, 2015. ISSN 1545-2263. Cited on page 5.

A Theoretical properties

To provide a reliable proof of convergence for the ACH algorithm, it is necessary to provide some intuition about data distribution in the expendable case and the corresponding behavior of ACH. Since $\mathbf{X} \in \mathbb{R}^{n \times m}$ is composed of the training examples and $\mathbf{C}$ is composed of K sets of values disposed in the feature space, it becomes possible to infer two propositions inherent to the model with greater computational cost. The first proposition concerns to the maximum number of necessary iterations to the most external loop (lines 2 to 19) before convergence.

Proposition 1. *Assuming that* $\mathbf{C}$ *is composed of non empty clusters* $\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_K \in \mathbb{R}^m$*, admitting the same operations such as* $\mathbf{c}_i \subseteq \mathbf{C}$ *and* $\mathbf{X}$ *has only one real data cluster to be defined, then among the* $K!$ *possible permutations of centroids there is at least one which makes necessarily* $\frac{K^2+K-2}{2}$ *iterations before the algorithm converges.*

Proof. Since $\mathbf{C}$ admits anyone of the $K!$ combinations of its elements, there is unless one to keep the centroids distributed on the feature space in a such manner that the algorithm needs $K + (K-1) + (K-2) + \cdots + 2 = \frac{K^2+K-2}{2}$ loops to determine only one cluster describing the actual data shape. This occurs due to the algorithm merges only two components per iteration, in the worst case, forcing the algorithm to check all the components previously verified. □

The second proposition derives from the first and establishes a limit of iterations to the most internal loop (lines 7 to 14 in Algorithm 1), defining the number of hyperspheres in a same component.

Proposition 2. *Being the increment value of the hypersphere radius defined by Equation 3.2 (Section 3.2.1) and* $\mathbf{c}_i$ *is close to the geometric center of the component, then the maximum number of hyperspheres* H_y *before the algorithm converges is given by*

$$H_y \leq \left\lceil \frac{\max(\|\mathbf{c}_i - \mathbf{x}\|)}{R_0} \right\rceil, \ \forall \mathbf{x} \in \mathbf{X}. \tag{A.1}$$

Proof. When a centroid is positioned on the center of a real component (or in its neighbourhood), the hypersphere radius increases as an AP with a common difference equal to R_0. Thus, one can naturally conclude that the hypersphere radius is never greater than the component radius. When the hypersphere reaches the border of the component, more

sparser are the observations, which reduces the sample density compared to the last ACH iteration, leading to the convergence of the algorithm.

□

B Complexity proof

Based on previous propositions, this section provides the asymptotic complexity proof. However, before to start the analysis, some required cost informations are introduced. To estimate a superior limit it is necessary associate a maximum cost value to the execution of each instruction line. For each simple line (e.g., arithmetic operations and logic comparisons) it is assumed a constant value equal to one. On the other hand, to the lines with function calls, the cost is calculated based on some analytical considerations.

Initially, one should analyze the lines with constant cost. The line 1 classifies each observation as belonging to a cluster, which gives a cost equal to $K \times n$. In similar manner, to dislocate K centroids it is imperative to evaluate n observations in a m-dimensional space. Thus, the line 3 assumes a cost equal to the product between the feature space dimension m and the number of training data points n added to the number of centroids K. The line 4 is responsible for selecting the current pivot centroid. If none merges occurred in the last iteration, the next centroid in the set C is selected, otherwise the first centroid is chosen. To this line a constant cost is also assumed.

To compute a component radius, in the worst case, it is necessary evaluate $n - 1$ observations. In this case, the line 5 has a complexity of $n - 1$. The line 9 indicates which points are inside the hypersphere, being necessary analyze all the n points in the training matrix $\mathbf{X}$, deriving a complexity equal to n. In a similar manner, to compute the sample density of a hypersphere, the line 11 needs a maximum of $n \times m$ iterations before convergence.

The function in the line 15 analyzes all the centroids in each iteration to define which ones can be merged. This process results in a complexity equal to $\mid \mathbf{C} \mid$. In the line 17, a new function call to *calcIndexes* is made. As the number of centroids may be reduced over the iterations, this line cost depends on the cardinality of the set $\mathbf{C}$. However, asymptotically one can apply the same cost assumed to the line 1.

To understand the maximum complexity estimated in the line 9, the Proposition 2 discussed previously is required. It is assumed that, when there is only one cluster defined by the data and $K > 1$, the maximum number of built hyperspheres depends of the component radius. Therefore, the number of iterations in the line 9, in the worst case, is equal to $\left\lceil \frac{\max(\|\mathbf{c}_i - \mathbf{x}\|)}{R_0} \right\rceil$. In a solution with K centroids it is possible to infer successive merges performed two by two until one centroid remains. In this case, after each merge, all the components are revalidated. Thereby, the complexity is equivalent to an AP with

common difference and initial term equal to two and one, respectively.

Adding the cost of all terms and multiplying those inside loops by the maximum number of iterations in Proposition 1 derives

$$F(K,n,m) = \left(\frac{K^2+K-2}{2}\right)\left((n-1) + 7 + H_y(mn+n+4) + \frac{(n+1)(K^2+K-2)}{2} \right) + nK + nm + K,$$

ordering and excluding components of less asymptotic order result in

$$\begin{aligned} F(K,n,m) &= \left(\frac{nK^2 - K^2 + 7K^2 + nmK^2H_y + 4K^2H_y + 2nK + 2nm + 2K}{2}\right) \\ &\quad + \left(\frac{nK^4 + K^4 + 2nK^3 + 2K^3 + K^2 + 4}{4}\right) \\ &\quad - \left(\frac{3nK^2 + 4nK + 4K + 4n}{4}\right), \\ &< nmK^2H_y + nK^4 + 2nK^3 + nK^2 + K^4 + 4K^2H_y + 2K^3 \\ &\quad + 6K^2 + 2nm + 2nK + K^2 + 2K + 4 - 3nK^2 - 4nK - 4n - 4K, \\ &< nmK^2H_y + nK^4 + 2nK^3 + nK^2 + K^4 + 4K^2H_y + 2K^3 + 2nm \\ &\quad + 2nK + 6K^2 + K^2 + 2K. \end{aligned}$$

Initially, one may suppose the term nK^4 as the one with highest complexity order. However, the asymptotic curve of the term nmK^2H_y is greater due to $K \ll m$. Substituting H_y and R_0

$$\begin{aligned} F(K,n,m) &= nmK^2H_y, \\ &= nmK^2 \left\lceil \frac{\max(\|\mathbf{c}_i - \mathbf{x}\|)}{R_0} \right\rceil, \\ &= nmK^2 \left\lceil \frac{\max(\|\mathbf{c}_i - \mathbf{x}\|)}{\log_{10}(\|\mathbf{c}_i - \mathbf{x}_{max}\| + 1)} \right\rceil, \\ &\simeq nmK^2 \frac{\max(\|\mathbf{c}_i - \mathbf{x}\|)}{\log_{10}(\|\mathbf{c}_i - \mathbf{x}_{max}\| + 1)}. \end{aligned}$$

In the worst case, due to only one distribution fits the entire data (i.e., $K = 1$) it is assumed $D = \max(\|\mathbf{c}_i - \mathbf{x}\|) = \|\mathbf{c}_i - \mathbf{x}_{max}\|$, then

$$
\begin{aligned}
F(K, n, m) &= nmK^2 \frac{D}{\log_{10}(D+1)}, \\
&= nmK^2 \log_{10}(D+1)^{-D}.
\end{aligned}
$$

Finally, one can conclude the algorithm computational complexity as

$$\mathcal{O}\left(nmK^2\right).$$

When $K \approx n$, the asymptotic complexity becomes a third order function. This is a little worse than most of the traditional cluster-based algorithms in literature, such as k-means and fuzzy c-means (GHOSH, 2013). However, its agglomerative characteristic allows to model, at the same time, the data shapes and discover the optimal number of clusters. This can be pointed out as an advance over other clustering approaches that require offline mechanisms to infer the number of clusters.