POLITECNICO MILANO 1863

Mitigating Environmental and Operational Variability in Bridge SHM via Regime Partitioning and Adaptive Mahalanobis Novelty Scoring

School of Civil, Environmental and Land Management Engineering

Polo Territoriale di Lecco

Civil Engineering For Risk Mitigation

Author: **SeyedSoheil Kashfi**

Student ID: 976203
Advisor: Prof. Alberto Corigliano
Co-advisor: Prof. Alireza Entezami

Academic Year: 2024-2025

Abstract

Monitoring bridges is key to ensuring the structural safety and planning effective maintenance; however, environmental variations and operational (EOV) can alter measured structural responses in ways that resemble real damage, thus increasing false alarms or masking deterioration authentic. In real-world applications, the limited availability of damage data further limits the use of conventional supervised approaches, motivating the development of robust unsupervised strategies for reliable valuation structural conditions in the presence of variable operating conditions.

This thesis proposes an unsupervised hybrid framework that extends the Tuned Spectral Clustering (TSC) and integrates a damage index called Conditioning-guided Adaptive Mahalanobis Squared Distance (CAMSD). The method initially uses TSC to subdivide reference measurements into more homogeneous clusters, then performs the detection of anomalies through CAMSD with a statistical threshold calibrated on undamaged data. The approach is validated on the KW51 dataset (undamaged pre-intervention segment, with real EOV) and on the Z24 dataset under progressive damage and is compared with the classic Mahalanobis Squared Distance (MSD) approach, with variants based on robust and standard-baseline covariance. The results of the various analyses show a improved mitigation of EOV effects, a reduced false alarm rate and a more reliable damage detection, supporting the proposed approach as a solution practical for monitoring bridges.

Keywords: Structural monitoring, unsupervised learning, spectral clustering, EOV mitigation, KW51 bridge, Z24 bridge, anomaly detection.

Abstract in lingua Italiana

Il monitoraggio dei ponti è fondamentale per garantire la sicurezza strutturale e pianificare interventi di manutenzione efficaci; tuttavia, la variabilità ambientale e operativa (Environmental and Operational Variability, EOV) può alterare le risposte strutturali misurate in modo da imitare il danneggiamento reale, incrementando i falsi allarmi oppure mascherando un deterioramento effettivo. Nelle applicazioni in campo reale, la limitata disponibilità di dati di danno riduce ulteriormente l'efficacia degli approcci supervisionati convenzionali, motivando lo sviluppo di strategie robuste non supervisionate per una valutazione affidabile delle condizioni strutturali in presenza di condizioni operative variabili.

La presente tesi propone un framework ibrido non supervisionato che estende il Tuned Spectral Clustering (TSC) e integra un indice di danno denominato Conditioning-guided Adaptive Mahalanobis Squared Distance (CAMSD). In una prima fase, il TSC suddivide le misure di riferimento in cluster più omogenei; successivamente, il rilevamento delle anomalie è eseguito mediante CAMSD, adottando una soglia statistica calibrata su dati indisturbati. L'approccio è validato sul dataset KW51 (segmento pre-intervento non danneggiato, con EOV reali) e sul dataset Z24 in presenza di danneggiamento progressivo, ed è confrontato con il metodo classico basato sulla distanza quadratica di Mahalanobis (MSD), includendo varianti con stima della covarianza standard e robusta. I risultati mostrano una migliore mitigazione degli effetti EOV, una riduzione del tasso di falsi allarmi e una rilevazione del danno più affidabile, supportando la proposta come soluzione pratica per il monitoraggio dei ponti.

Parole chiave: Monitoraggio strutturale; apprendimento non supervisionato; clustering spettrale; mitigazione della variabilità ambientale e operativa (EOV); ponte KW51; ponte Z24; rilevamento di anomalie.

Contents

List of Figures

List of Tables

Acronyms

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
CAMSD	Conditioning-guided Adaptive Mahalanobis Squared Distance
DBC	Density-Based Clustering
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
EEMD	Ensemble Empirical Mode Decomposition
ICEEMDAN	Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
EM	Expectation–Maximization
EMD	Empirical Mode Decomposition
EOV	Environmental and Operational Variability
FFT	Fast Fourier Transform
FNR	False Negative Rate
FPR	False Positive Rate
GMM	Gaussian Mixture Model
GPR	Gaussian Process Regression
IMF	Intrinsic Mode Function
kNN	k-Nearest Neighbours
KW51	KW51 bridge dataset (Belgium)
ML	Machine Learning
MSD	Mahalanobis Squared Distance
OMA	Operational Modal Analysis
PSD	Power Spectral Density
RBF	Radial Basis Function
SC	Spectral Clustering
SHM	Structural Health Monitoring
SPR	Statistical Pattern Recognition
SVM	Support Vector Machine
TSC	Tuned Spectral Clustering
VMD	Variational Mode Decomposition
Z24	Z24 bridge dataset (Switzerland)

List of Symbols

t	Sample index (time/sample number)
n	Number of samples (observations)
p	Feature dimension (number of variables)
v	Number of modal-frequency features (typically $v = 4$)
$\mathbf{x}_t$	Feature vector at sample t
$\tilde{\mathbf{x}}_t$	Standardized feature vector at sample t
$\boldsymbol{\mu}_{\mathrm{tr}}$	Training (baseline) mean vector used for standardization
$\boldsymbol{\sigma}_{\mathrm{tr}}$	Training (baseline) standard deviation vector used for standardization
K	Number of clusters / regimes (or mixture components)
$\mathcal{C}_k$	The k-th cluster (regime)
$\boldsymbol{\mu}_k$	Mean vector / centroid of cluster (or local neighborhood)
$\boldsymbol{\Sigma}_k$	Covariance matrix of cluster (or local neighborhood)
π_k	Mixture weight of the k-th Gaussian component
$p(\mathbf{x})$	Probability density of feature vector $\mathbf{x}$
w_{ij}	Similarity between samples i and j (RBF/graph weight)
$\mathbf{W}$	Similarity (weight) matrix
σ	RBF kernel scale (similarity bandwidth)
$\mathbf{D}$	Degree matrix of the similarity graph
d_i	Degree of node i (row-sum of $\mathbf{W}$)
$\mathbf{L}_{\mathrm{sym}}$	Symmetric normalized graph Laplacian
$\mathbf{L}_{\mathrm{rw}}$	Random-walk normalized graph Laplacian
$\mathbf{U}$	Matrix of selected eigenvectors (spectral embedding basis)
$\mathbf{V}$	Spectral embedding matrix (rows clustered to obtain regimes)
ε	DBSCAN neighborhood radius
minPts	Minimum number of neighbors to define a DBSCAN core point
$\mathcal{N}_\varepsilon(\mathbf{x}_i)$	ε-neighborhood of point $\mathbf{x}_i$
$\mathbf{z}$	Query (monitoring) feature vector
$\mathcal{N}_k(\mathbf{z})$	Local neighborhood of size k around $\mathbf{z}$
k	Neighborhood size used in CAMSD (typically varied in $[k_{\min}, k_{\max}]$)

$d^2(\mathbf{z})$	CAMSD squared distance / anomaly score of $\mathbf{z}$
δ	Novelty index (anomaly score) used for alarming (as plotted in results)
$\mathbf{\Sigma}_{k,\lambda}$	Regularized covariance matrix for neighborhood size k
λ	Covariance regularization parameter (ridge term)
$\kappa(\mathbf{\Sigma})$	Condition number of covariance matrix (numerical conditioning indicator)
α	Target false positive rate (FPR) used for threshold calibration
τ	Decision threshold on novelty/anomaly scores
$Q_{1-\alpha}(\cdot)$	Empirical $(1-\alpha)$-quantile operator
$\mathrm{cov}(X, Y)$	Covariance between variables X and Y
$\lVert\cdot\rVert_2$	Euclidean (ℓ_2) norm
$\mathbf{I}$	Identity matrix
$\mathbb{R}^p$	p-dimensional real vector space
$(\cdot)^\top$	Transpose operator

1 | Structural Health Monitoring

1.1. Introduction

SHM integrates sensing technologies with data analytics to evaluate the condition of a structure and to identify damage as early as possible [8, 26]. In real applications, however, structures are continuously subjected to changing environmental and operational conditions, which can significantly affect their measured dynamic response even in the absence of damage. For instance, temperature variations may lead to expansion or contraction and changes in boundary conditions, which can shift vibration based features; similarly, variations in traffic loading can alter response amplitudes and induce different strain or vibration patterns compared to low traffic periods.

This variability creates major difficulties for reliable damage identification. Environmental and Operational influences may mask damage signatures or, conversely, produce apparent changes that resemble damage, leading to false alarms and missed detections if not treated properly [7].

Common environmental factors that influence structural response include:

- Temperature: thermal expansion/contraction and temperature dependent stiffness or boundary effects that often follow daily or seasonal cycles.
- Wind: time varying loads that can modify excitation levels and increase measured vibrations.
- Humidity: moisture-related effects that may influence apparent mass or stiffness or introduce variability in measured features.
- Solar radiation: thermal gradients that can modify measured modal properties and response features.

Operational variability is also important because load and usage patterns change over time:

- Traffic: varying vehicle flows modify dynamic loading, with congestion often producing repeated or higher stress/response levels.

- Special events: construction, retrofitting, or maintenance activities can temporarily alter structural response and measurement conditions.

These factors often act simultaneously and can generate feature variations comparable to those caused by early-stage damage.

From a practical standpoint, a robust SHM workflow therefore aims (i) to mitigate or model environmental and operational effects and (ii) to establish a reliable baseline model for the undamaged (or reference) condition. New observations can then be compared to this baseline to identify anomalies that may indicate damage [7, 43]. When these steps are implemented carefully, SHM supports continuous and near real time insight into structural condition and can improve maintenance planning and decision making [10]. From a risk mitigation perspective, the SHM objective is not only to detect anomalies, but to do so with a controlled false alarm rate (to avoid unnecessary interventions) while minimizing missed detections (to avoid unsafe continued operation). This thesis addresses the above challenge by adopting an unsupervised monitoring strategy aimed at (i) separating dominant EOV driven regimes in the feature space and (ii) performing novelty detection within an appropriate baseline representation, so that deviations attributable to structural deterioration can be flagged with controlled false alarm behavior.

1.2. Core Principles of SHM

Bridges are critical infrastructure for transportation networks and economic connectivity, therefore making early detection of abnormal structural behavior essential for safety, serviceability, and reduction of unplanned disruptions. A core concept in SHM is the tight integration of (i) sensing hardware, (ii) data acquisition, and communication systems, and (iii) analytical methods that transform measurements into conditions indicators. In practice, SHM implements an inference chain from measurement $\rightarrow$ feature $\rightarrow$ statistical model $\rightarrow$ decision, where sensing hardware and data acquisition provide the measurements, and analytical methods convert them into condition indicators that support operational decisions (alarms, inspection triggers, maintenance actions) [10, 26].

Modern SHM increasingly leverages machine learning because monitoring campaigns produce large datasets, and relationships between response features and operating conditions are often nonlinear and uncertain [10, 32]. In this setting, machine learning can support automated data screening, feature extraction, normalization, clustering of operating regimes, and novelty detection [43, 45]. Many SHM systems can be framed under an

SPR paradigm, typically structured into operational evaluation, data acquisition, feature extraction, and statistical model development.

1.3. Different Levels of SHM

SHM objectives are commonly described in levels of increasing complexity and information requirements [8]. At the most basic level, monitoring data are used to detect whether the structure has deviated from a reference undamaged condition, typically using global response features such as modal properties. Beyond detection, more advanced objectives include localizing where damage is likely to exist and quantifying its severity [8].

In many practical frameworks, damage diagnosis is organized as:

- Level 1 – Early Damage detection: Determines whether damage is present in the structure binary yes/no decision. Methods typically rely on global features such as natural frequencies or mode shapes that can be extracted from sparse sensor arrays, making Level 1 the most practical starting point for operational monitoring.

- Level 2 – Damage localization: Identifies the spatial location of detected damage (e.g., which span, pier, or structural element). This requires either denser instrumentation or more sensitive local features, and often relies on comparing measured mode shapes against reference patterns or using pattern recognition to associate feature changes with specific regions

- Level 3 – Damage quantification: Estimates the severity or extent of damage (e.g., crack depth, stiffness reduction percentage, residual load capacity). Quantification typically requires physics based models (e.g., finite element updating) or calibrated regression models that map feature changes to damage metrics, making it the most demanding and model-dependent level.

- Level 4 – Prognosis (prediction): At Level 4 the goal shifts from diagnosis to forecasting future deterioration and remaining service life by linking the current inferred state to deterioration models and expected operational demands.

As the level increases, the task generally becomes more demanding, requiring more informative measurements, better sensor placement, and stronger modeling assumptions. For example, global vibration features may support detection, but localization and quantification can be challenging in large or complex structures, especially when instrumentation is sparse and environmental effects are strong.

In addition to diagnosis, prognosis aims to forecast future deterioration and remaining service life by linking the current state of the structure to expected loading and environmental conditions over time. Such predictions typically rely on long term monitoring records and on deterioration models that connect measured indicators to future performance.

This thesis focuses on Level 1 (damage detection) because operational bridge monitoring campaigns typically provide limited labeled data and sparse sensor coverage, making reliable binary detection a necessary precondition before attempting more advanced localization or quantification under field conditions.

1.4. Influence of EOV

Damage detection in operational structures is strongly complicated by EOV. Many vibration based indicators change not only due to damage but also due to temperature, wind, humidity, traffic intensity, and boundary-condition effects. This issue is particularly important when methods are validated primarily in numerical simulations or controlled experiments, because real structures often exhibit coupled and time-varying EOV effects that are difficult to replicate. As a result, EOV can generate response changes that are comparable in magnitude to those caused by damage, increasing the likelihood of False Positive Rate (FPR) when EOV is not addressed.

A key response to this challenge is data normalization, i.e., transforming features so that variations driven by EOV are reduced relative to variations driven by structural change; this is widely recognized as essential because EOV can mask subtle damage signatures in vibration based monitoring.

EOV factors often occur simultaneously, which further complicates interpretation. Temperature effects are among the most prominent in bridges; long term monitoring campaigns on benchmark structures such as KW51 and Z24 bridges show clear correlations between modal properties and temperature evolution. This sensitivity can mislead damage detection if a change in modal features is interpreted as damage when it is actually driven by temperature gradients or other environmental conditions [7]. When environmental measurements are incomplete or unavailable, output-only strategies (e.g., clustering operating regimes in the feature space) are often adopted as a practical surrogate for explicit EOV modeling.

1.5. SPR

Statistical Pattern Recognition (SPR) enables automated condition assessment by transforming measured features into statistical models and decision rules for detecting deviations from an undamaged baseline. Within an SHM framework, SPR is commonly implemented through a sequence of steps that extends the canonical four-part paradigm (operational evaluation, data acquisition, feature extraction, and statistical model development) [10]. Within an SHM framework, SPR is commonly implemented through a sequence of steps, including:

1. Operational assessment and definition of monitoring objectives,
2. Data acquisition and quality control,
3. Feature extraction and dimensionality reduction,
4. Feature classification, clustering, and novelty detection.

This pipeline supports a consistent interpretation of monitoring data and helps translate raw sensor measurements into actionable indicators of structural performance.

1.5.1. Operational Assessment

Operational assessment clarifies why SHM is needed and how the system will be used in practice. It includes defining damage scenarios, performance requirements, limitations of deployment, and the intended decision-making process (e.g., alarms, inspections, maintenance planning) [26]. Economic considerations are also important because the monitoring strategy should provide measurable benefits compared to conventional inspection and maintenance approaches.

1.5.2. Data Acquisition and Sensing

Reliable measurement of a structure's dynamic response is a fundamental requirement in SHM [26]. The sensing and data acquisition stage includes selecting suitable sensors, defining their locations, and choosing appropriate acquisition hardware, while also accounting for practical constraints such as installation effort and cost [19]. In SHM, the objective is to collect measurements that are sufficiently sensitive to structural changes and can support consistent tracking as deterioration evolves.

For bridge applications, multiple sensors are typically distributed across key parts of the structure (e.g., deck, piers, and abutments) to capture vibrations at different positions and improve the representativeness of the measured response [3]. Because bridges operate under varying conditions, the sensing system should provide stable measurements under traffic loading and environmental actions, so that changes in the recorded signals can be interpreted with confidence. Furthermore, data cleansing is an essential pre-processing step, especially when machine learning–based methods are used for SHM. It focuses on detecting and correcting issues such as outliers, inconsistencies, and missing or corrupted samples, which may arise from sensor faults, environmental disturbances, or transmission problems [10]. Without adequate cleansing, these errors can bias subsequent feature extraction and modeling, reducing the reliability of structural condition assessment.

1.5.3. Feature Extraction

Feature extraction compresses large volumes of raw measurements into a smaller set of indicators that are expected to be sensitive to structural change [8]. In vibration based SHM, features may include modal properties, spectral characteristics, and time-frequency descriptors, as well as statistical summaries derived from response time series [20]. [20]. Because EOV also influences many features, a key goal is to select or transform features so that damage related variability is not overwhelmed by environmental/operational effects. Moreover, data normalization aims to reduce non-damage related variability, improving the separability between healthy and potentially damaged conditions. When relevant environmental measurements are available, regression style normalization or filtering approaches can be applied; when they are incomplete, output only approaches based on clustering and unsupervised normalization can be used [45, 46]. In this thesis, the output only clustering paradigm is adopted because operational bridge monitoring campaigns like KW51, Z24, typically provide eigenfrequency time series without concurrent environmental measurements across all monitoring periods. The proposed TSC-CAMSD framework, therefore partitions baseline data into homogeneous operating regimes via clustering, then performs novelty detection within each regime to mitigate EOV induced false alarms without requiring explicit temperature or load data. In both cases, the objective is to obtain a representation where deviations from the baseline are more likely attributable to damage [7]. In the case studies considered in this thesis, the primary features are vibration based modal quantities (e.g., modal frequencies) extracted from long term monitoring data and evaluated under significant EOV influence.

1.5.4. Feature Classification and Novelty Detection

Once features are extracted and ideally normalized, SHM decisions are made by classification, clustering, or novelty detection [10]. In many civil applications, labeled damage data are scarce, so unsupervised methods are often used to represent the baseline condition and to flag outliers as potential anomalies. Distance based indicators, including Mahalanobis type measures, provide a practical way to quantify deviation from baseline behavior under environmental effects [43, 45]. In this context, clustering groups data into regimes with similar behavior, while novelty detection assigns a score or decision indicating how unlikely a new observation is under the baseline model.

1.6. Machine Learning for Enhanced SHM

Machine learning techniques can substantially improve the efficiency and accuracy of SHM by automating the interpretation of large volumes of vibration data and highlighting patterns or anomalies that may indicate structural deterioration [10]. Compared with approaches that depend heavily on manual inspection of results, these algorithms can be updated as new measurements become available, supporting early warning and long term monitoring [22]. Their ability to process high rate data streams enables timely condition assessment and can support maintenance decision making that reduces the probability of undetected deterioration progressing to unsafe states. In addition, machine learning is well-suited for capturing complex and nonlinear dependencies in monitoring data that may be difficult to represent with conventional models [32]. Finally, these data-driven tools can be integrated into broader decision-support workflows, increasing their practical value beyond standalone damage detection.

1.7. Model-driven and Data-driven Approaches in SHM

SHM methodologies are commonly grouped into model-driven and data driven approaches, depending on whether the decision-making is primarily anchored in physics-based modeling or learned directly from monitoring data. In bridge applications, this distinction is not merely conceptual: it strongly affects the type of uncertainty that dominates the damage-detection task and the practicality of the resulting pipeline in field conditions.

Model-driven SHM aims to represent the structure through a physics-based model (typically a finite-element formulation) and to infer damage by comparing predicted and measured responses. In principle, this class of methods is appealing because it supports direct mechanical interpretation and may enable damage localization and quantification when the model is accurate and the excitation is well characterized. However, for full-scale civil infrastructure, the modeling assumptions are often challenged by uncertain boundary conditions, evolving operational constraints, imperfect knowledge of material properties, and additional effects such as soil–structure interaction. These sources of uncertainty become even more critical when EOV alters the dynamic characteristics of the structure, meaning that discrepancies between model and data may reflect EOV or modeling error rather than actual damage. In practice, many operational SHM solutions adopt hybrid strategies that combine physics-based modeling with data-driven learning. For example, simplified mechanical models may be used to constrain plausible behavior, while machine learning captures residual patterns, compensates for unmodeled EOV effects, or supports regime partitioning in high-dimensional feature spaces.

Data-driven SHM, in contrast, relies on extracting information directly from measured data and learning statistical regularities of the undamaged condition. This perspective is particularly well suited to vibration based bridge monitoring, where features such as modal parameters are known to vary with environmental and operational conditions, and where labeled damaged-state data are rarely available in practice. For these reasons, many modern SHM solutions adopt an SPR workflow: a baseline dataset is collected under normal operation, features are extracted and standardized, and an unsupervised model is trained to detect deviations from the learned baseline behavior. In this thesis, emphasis is placed on unsupervised, non-parametric strategies because they match realistic monitoring constraints and avoid relying on damage labels or strict distributional assumptions.

Within this data-driven perspective, the approach developed in this work combines two complementary components. First, a clustering stage partitions baseline data into operating regimes to mitigate EOV by ensuring that new observations are compared against an appropriate local reference. Second, anomaly detection is carried out using a conditioning-guided adaptive Mahalanobis-type distance (CAMSD), which produces a novelty index that can be thresholded for early warning while maintaining numerical stability through covariance-conditioning checks and regularization. Because high-fidelity modeling of full-scale bridges under evolving boundary conditions and uncertain operational loads is difficult in practice, this thesis emphasizes data-driven, unsupervised strategies designed to remain effective under pronounced EOV.

1.8. Thesis Structure

This thesis is organized as follows. Chapter 1 introduces bridge SHM and the challenge of environmental and operational variability (EOV). Chapter 2 reviews unsupervised learning and clustering methods used for regime identification in SHM. Chapter 3 summarizes the sensing context and vibration-feature extraction background relevant to the adopted datasets. Chapter 4 presents the KW51 and Z24 case studies and the feature sets used.

Chapter 5 details the proposed TSC–CAMSD framework, including regime partitioning and novelty scoring with threshold calibration. Chapter 6 reports experimental results, comparisons, and performance metrics on both datasets. Chapter 7 concludes the work, discusses limitations, and outlines future developments.

1.9. Objectives and contributions

The objective of this thesis is to develop and evaluate an unsupervised SHM workflow for bridge monitoring data that is robust to EOV and suitable for early-warning damage detection. Specifically, the thesis:(i) frames the monitoring problem within a regime-aware baseline modeling perspective; (ii) evaluates unsupervised techniques to partition dominant operating regimes in feature space; and (iii) applies novelty detection to quantify deviations from baseline behavior with controlled false-alarm performance. The proposed workflow is validated on two benchmark bridge datasets (KW51 and Z24), emphasizing interpretability and deployment-oriented decision thresholds.

2 | Machine Learning Methodologies

2.1. Utilizing Machine Learning for SHM

Machine learning algorithms can process and interpret large streams of sensor measurements in real time, which supports continuous monitoring and earlier detection of potential damage [13]. When integrated into SHM, these methods can go beyond simple anomaly flagging by providing indications about the probable damage type, its severity, and the potential implications for structural performance [8]. In many applications, models are calibrated using historical records that describe normal behavior and, when available, damaged conditions, so that new measurements can be evaluated against learned patterns [22].

By learning characteristic features associated with different damage mechanisms, these algorithms can deliver consistent predictions when new data are provided. Compared with traditional approaches, machine learning is particularly useful when monitoring data is high-dimensional and influenced by changing EOV conditions [10]. As SHM systems increasingly generate large volumes of vibration data, automated data-driven models become attractive due to their computational efficiency and their ability to scale to long term monitoring [12].

In SHM, machine learning strategies are commonly grouped into supervised and unsupervised learning. Supervised learning requires data from both undamaged and damaged states, whereas unsupervised learning is typically trained using only data from the healthy condition [42]. A key limitation of supervised learning in civil infrastructure is that realistic damaged data are rarely available, since intentionally inducing damage in full-scale structures is not feasible. Therefore, unsupervised methods are often preferred in practice, because they establish a baseline from the undamaged state and then evaluate new data by comparing them against this reference [9]. In this setting, the workflow generally consists of a training phase to learn the baseline model and a monitoring (validation) phase in which deviations from the baseline are used to detect, localize, or quantify potential damage [11].

Overall, integrating machine learning into SHM improves both the speed and consistency of structural assessment, and it can support damage categorization so that maintenance Actions can be prioritized based on the suspected severity of the identified issue [16]. At the same time, practical challenges remain, particularly the availability of representative training data and the continued need for engineering judgment when interpreting results [36]. Finally, the general objective is to learn an automated computational model from training data and then assess the current structural condition using test data and the learned model, which can be implemented through either supervised or unsupervised learning strategies [22].

2.2. Fundamentals of Machine Learning

In machine learning, uncertainty appears in different ways: for example, deciding what prediction is most plausible given past observations, selecting an appropriate model to explain the available data, or determining which additional measurement would be most useful to reduce ambiguity. Probabilistic machine learning provides a principled framework to represent and propagate uncertainty in predictions by modeling distributions rather than single-point estimates. In SHM, this is valuable because operational decisions (e.g., inspection triggers) require not only an alarm score but also an indication of how uncertain the prediction is under limited or changing data. Importantly, probabilistic models provide predictive uncertainty (e.g., posterior variance), which should not be conflated with guaranteed decision confidence unless calibration is explicitly assessed. [33, 34].

This perspective is closely connected to statistics because both fields model randomness and draw inferences from data; however, probabilistic machine learning places stronger emphasis on learning predictive models that can be updated as new information becomes available. In addition, Probabilistic formulations can also support sequential data acquisition strategies, but this aspect is not pursued in the present work.

Within SHM, several classes of machine learning methods can support damage assessment. Supervised learning uses labeled examples to learn mappings between features and known structural states, while unsupervised learning aims to discover patterns and detect anomalies using unlabeled data, which is often more realistic for civil infrastructure monitoring [9, 42].

2.2.1. Supervised Learning Techniques

All machine learning approaches first require a suitable set of training data. In supervised learning, both the inputs and the corresponding outputs are available during data collection; therefore, the training set is formed by input–output pairs. Each input is associated with a known target value or label, which acts as a supervision signal and guides the learning process. When enough representative pairs are available, supervised methods often provide strong performance. However, assembling labeled datasets commonly depends on manual annotation or expert judgment, which can be costly and time-consuming in practical applications [22]. Supervised learning in SHM is typically applied through classification and regression. Classification assigns observations to discrete categories (e.g., damaged vs. undamaged), whereas regression predicts continuous quantities (e.g., a damage index). Among commonly used supervised algorithms, Support Vector Machines (SVMs) are widely adopted for classification due to their margin-based decision rule and strong generalisation in high-dimensional settings [4]. Tree-based models, including Random Forests, are also used because they are flexible, can capture nonlinear interactions, and often provide useful variable-importance summaries [2].

2.2.2. Unsupervised Learning Techniques

Unsupervised learning addresses problems in which only input data are available during training, meaning that no output labels are provided with the collected measurements. An effective unsupervised algorithm must therefore define a criterion for grouping observations based only on the information contained in the inputs, even though the true labels are unknown [22]. The central difficulty is that "similarity" must be inferred without supervision, which generally makes unsupervised learning more challenging than supervised approaches. At the same time, building an unlabeled training dataset is often more economical in practice, since it avoids the additional effort and cost of manual annotation [22]. For this reason, there remains a strong need for reliable strategies that can learn useful structure from unlabeled data. In machine learning, training data are used to construct mathematical models for decision-making, and model selection typically involves a trade-off between simplicity and complexity [32]. Simple models (e.g., linear formulations) are usually computationally efficient and can be trained robustly even with limited data, but they may fail to capture complex patterns as the variability of the dataset increases. More flexible models can represent richer relationships, yet they often require more data and greater computational resources to achieve stable performance [32]. In SHM, a key advantage of unsupervised learning is that it can operate without labeled damaged-state data, which makes it a practical and cost-effective option for large-scale monitoring of civil infrastructure [9]. In SHM practice, unsupervised novelty detection typically assumes that the training data represent predominantly undamaged behavior. This assumption must be stated explicitly because contamination of the baseline (i.e., including damaged or anomalous data during training) can inflate thresholds and increase missed detections.

2.2.3. Parametric and Nonparametric Models

Parametric models assume a fixed functional form described by a finite set of parameters, whereas nonparametric models adapt their effective complexity to the amount of available data. This distinction is relevant in SHM because the relationship between EOV and response features may be nonlinear, and flexible models can improve robustness when sufficient baseline data exist. GPR is a nonparametric Bayesian approach that defines a prior over functions through a covariance kernel and yields predictive means and variances. It has been widely used for EOV normalization because it can model nonlinear trends and quantify predictive uncertainty; however, standard implementations scale cubically with the number of training points unless approximations are adopted [40].

2.3. Machine Learning Approaches in SHM

In SHM, both supervised and unsupervised machine learning approaches are used for damage detection. Supervised methods (e.g., SVMs and classifiers based on neural networks) require labeled examples of structural states. In contrast, unsupervised approaches, including clustering, density estimation, and reconstruction-based models such as autoencoders, are useful when damaged-condition labels are limited or unavailable. Therefore, the choice between supervised and unsupervised strategies depends on data availability, the monitoring objective, and the acceptable trade-off between false alarms and missed detections.

In this thesis, the emphasis is placed on unsupervised approaches for damage detection in SHM, motivated by the practical difficulty of obtaining reliable labeled damage data for full-scale civil infrastructure. Before presenting the particular unsupervised technique adopted in this work, a brief overview of widely used unsupervised methods in SHM is provided. This background helps clarify how clustering-based strategies can support damage identification without requiring labeled training data.

2.4. Clustering algorithms

Clustering partitions feature space into groups that can be interpreted as distinct operating regimes, which helps to construct robust baselines and reduce false alarms under EOV. In SHM applications, clustering is typically embedded in a two-phase procedure: during a baseline period, features are collected under varying conditions to form training data; during monitoring, new features are compared to the baseline (directly or via regime assignment) to indicate abnormal behavior [45, 49]. Figure 2.1 provides a broad visual overview of common clustering methods.

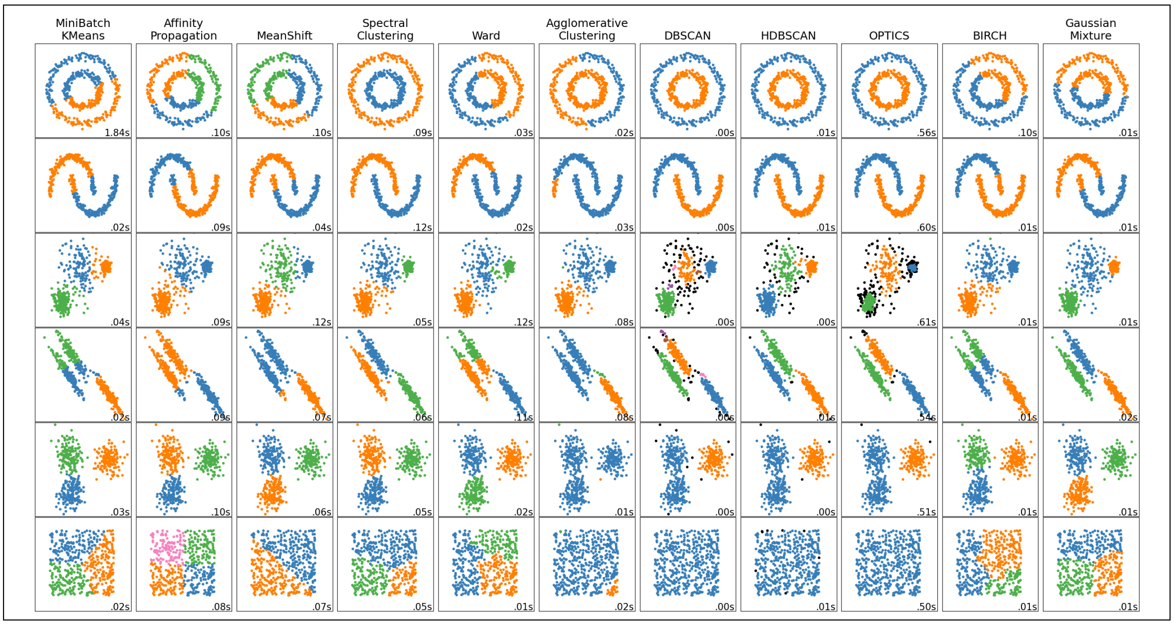


Figure 2.1: Overview of common clustering algorithm families, adapted from [48].

2.4.1. K-means Clustering

K-means is one of the most widely used partitional clustering algorithms due to its simplicity and computational efficiency. Given feature vectors $\{\mathbf{x}_i\}_{i=1}^n$ with $\mathbf{x}_i \in \mathbb{R}^d$, the method partitions the observations into K clusters by minimising the within-cluster sum of squares [27]:

$$\min_{\{\mathcal{C}_k\}_{k=1}^K} \sum_{k=1}^{K} \sum_{\mathbf{x}_i \in \mathcal{C}_k} \|\mathbf{x}_i - \boldsymbol{\mu}_k\|_2^2 , \tag{2.1}$$

where $\boldsymbol{\mu}_k$ is the centroid of cluster $\mathcal{C}_k$. In practice, the standard algorithm alternates between assignment to the nearest centroid and centroid updates until convergence [25].

In SHM, K-means is often used as a baseline tool to group data into operating regimes (e.g., seasonal or diurnal patterns) when the feature space is approximately separable under Euclidean distance [27].

A practical requirement of K-means is specifying the number of clusters K. In operational SHM datasets, K is commonly selected using exploratory criteria such as the elbow rule on the within-cluster sum of squares, or by comparing solutions over a range of candidate values. Since the objective is non-convex, K-means may converge to a local minimum; therefore, multiple random initialisations (restarts) are often used, and the solution with the lowest objective value is retained [21].

Because K-means is distance-based, feature scaling is essential: variables should be standardised (e.g., z-score or robust scaling) so that clustering is not dominated by the largest-variance feature. Despite its popularity, K-means has limitations that are relevant to bridge monitoring. First, the algorithm implicitly favors spherical clusters and may perform poorly when EOV produces elongated or curved manifolds (e.g., frequency–temperature dependencies). Second, the centroid estimate is sensitive to outliers, meaning that sensor faults or transient events can distort the cluster centres and degrade the baseline regimes. For these reasons, K-means is often treated as a baseline method in SHM rather than a final solution when strong EOV is present.

2.4.2. Hierarchical Clustering

Hierarchical clustering constructs a nested sequence of partitions represented by a dendrogram. This is advantageous in SHM because it allows the analyst to explore regime structure at multiple resolutions (e.g., daily variations nested within seasonal trends). Unlike k-means, hierarchical clustering does not require specifying the number of clusters K in advance; instead, K is selected by cutting the dendrogram at a chosen dissimilarity level [41, 49].

In agglomerative (bottom-up) hierarchical clustering, each observation starts as its own cluster, and the algorithm repeatedly merges the two closest clusters according to a linkage rule. Let $d(\mathbf{x}, \mathbf{y})$ be the point-to-point dissimilarity (often Euclidean distance after feature scaling). For two clusters A and B, common linkage criteria are:

$$D_{\text{single}}(A,B) = \min\{d(\mathbf{a},\mathbf{b}) : \mathbf{a} \in A,\ \mathbf{b} \in B\}, \tag{2.2}$$

$$D_{\text{complete}}(A,B) = \max\{d(\mathbf{a},\mathbf{b}) : \mathbf{a} \in A,\ \mathbf{b} \in B\}, \tag{2.3}$$

$$D_{\text{average}}(A,B) = \frac{1}{|A|\,|B|} \sum_{\mathbf{a}\in A} \sum_{\mathbf{b}\in B} d(\mathbf{a},\mathbf{b}), \tag{2.4}$$

$$\Delta_{\text{Ward}}(A,B) = \frac{|A|\,|B|}{|A|+|B|}\, \|\boldsymbol{\mu}_A - \boldsymbol{\mu}_B\|_2^2\,, \tag{2.5}$$

where $\boldsymbol{\mu}_A$ and $\boldsymbol{\mu}_B$ are the cluster centroids. Ward's method (minimum-variance linkage) merges the pair that yields the smallest increase in the total within-cluster sum of squares [35, 55].

The linkage choice is important in SHM because baseline regime geometry is strongly influenced by EOV and measurement noise. Single linkage can capture elongated structures but is sensitive to chaining and noise; complete linkage tends to produce compact clusters but may fragment elongated regimes; average linkage provides an intermediate behavior; Ward linkage often yields compact partitions by construction [41, 49].

A drawback of hierarchical clustering is its computational and memory cost for large datasets (standard implementations scale on the order of $O(n^2)$) and its sensitivity to outliers, which can create misleading early merges and distort the dendrogram [35, 41]. Therefore, hierarchical clustering is often used for exploratory regime analysis or for moderate-sized datasets, sometimes complemented by robust scaling, outlier handling, or robust distance measures.

2.4.3. Density-Based Clustering (DBC)

Density-based clustering (DBC) defines clusters as contiguous regions of high point density separated by regions of low density. This is appealing in SHM because baseline data can form non-spherical sets in feature space, while monitoring data may include isolated points caused by damage, sensor drift, or transient operational events [43, 49]. In contrast to partitional methods that force every point to belong to a cluster, density-based methods naturally label low-density points as noise, which aligns well with novelty-detection goals [41].

A widely used density-based algorithm is DBSCAN (Density-Based Spatial Clustering of Applications with Noise), which identifies core points using a local neighborhood criterion and expands clusters through density-reachability [18]. Let $\{\mathbf{x}_i\}_{i=1}^{n}$ be feature vectors and define the ε-neighborhood of point $\mathbf{x}_i$ as

$$\mathcal{N}_\varepsilon(\mathbf{x}_i) = \left\{\mathbf{x}_j : \|\mathbf{x}_j - \mathbf{x}_i\|_2 \leq \varepsilon\right\}. \tag{2.6}$$

A point $\mathbf{x}_i$ is a core point if

$$|\mathcal{N}_\varepsilon(\mathbf{x}_i)| \geq \text{minPts}, \tag{2.7}$$

where minPts is the minimum number of neighbours required to form a dense region. Points that are not core points but lie within the ε-neighborhood of a core point are border points, while remaining points are labeled as noise (outliers). DBSCAN forms clusters as maximal sets of density-connected points, starting from a core point and iteratively adding density-reachable neighbours [18].

In SHM, DBSCAN is useful when the number of regimes is unknown and when the dataset contains clear noise/outlier regions [49]. However, DBSCAN can be sensitive to parameter selection (especially ε) and may struggle when cluster densities vary substantially across operating conditions because ε acts as a global density threshold [30]. In practice, careful feature scaling is required, and when strongly variable densities are expected, hierarchical density-based variants such as HDBSCAN can be considered as a more robust alternative [31].

2.4.4. Gaussian Mixture Models (GMMs)

GMMs represent the distribution of features as a weighted combination of Gaussian components [32]. This probabilistic representation is useful in SHM because baseline data may be multimodal due to different operating regimes (e.g., temperature ranges or traffic patterns), and a single Gaussian model may not describe the healthy state adequately [32, 49]. In addition, GMMs provide soft memberships (posterior probabilities of belonging to each component), which can reduce hard boundary effects when operating conditions transition smoothly [32].

Formally, a GMM models the feature density as

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k), \qquad \sum_{k=1}^{K} \pi_k = 1,\ \pi_k \geq 0, \tag{2.8}$$

where π_k are the mixture weights, and $(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ are the mean vector and covariance matrix of the k-th Gaussian component.

In practice, GMM parameters are typically estimated via the expectation–maximisation

(EM) algorithm [6]. The number of mixture components can be selected using likelihood-based information criteria such as AIC and BIC [1, 47], or by validation on baseline consistency [32]. Within SHM pipelines, GMMs can be used either as a clustering tool for regime identification or as a probabilistic reference model, where observations with low likelihood (or low posterior under the baseline components) are flagged as novel [5, 13].

A known limitation is sensitivity to initialisation and susceptibility to outliers, since extreme points can inflate covariance estimates and distort the fitted components [32]. For field monitoring, this motivates data cleansing, robust scaling, covariance regularisation, or coupling GMM-based regime modeling with outlier detection to protect baseline modeling [43, 49]. Overall, GMMs provide a strong probabilistic baseline model, but they may struggle when the healthy-state data lie on strongly nonlinear manifolds [49].

2.4.5. Spectral Clustering (SC)

Spectral clustering (SC) is designed for datasets where the natural cluster structure is nonlinear and cannot be captured well by centroid-based methods. Instead of clustering directly in the original feature space, SC constructs a similarity graph between observations and uses an eigenvector-based embedding of a graph Laplacian to reveal separable structure in a lower-dimensional representation [52, 54]. This is particularly relevant in SHM, where EOV effects can generate curved or intertwined trajectories in feature space (e.g., modal frequency trends with temperature) [38, 49].

Given feature vectors $\{\mathbf{x}_i\}_{i=1}^n$, a common choice is to build a weighted similarity matrix $W \in \mathbb{R}^{n\times n}$ using a Gaussian kernel:

$$W_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{2\sigma^2}\right), \qquad W_{ii} = 0, \tag{2.9}$$

where σ controls the similarity scale [37, 52]. Let D be the degree matrix with $D_{ii} = \sum_j W_{ij}$. Two widely used normalised Laplacians are the symmetric Laplacian

$$L_{\text{sym}} = I - D^{-1/2} W D^{-1/2}, \tag{2.10}$$

and the random-walk Laplacian

$$L_{\text{rw}} = I - D^{-1} W, \tag{2.11}$$

both of which are commonly preferred over the unnormalised Laplacian in practice [52, 54].

A standard implementation (Ng–Jordan–Weiss) proceeds as follows [37, 52]:

1. Construct the similarity graph (i.e., compute W) and degree matrix D.
2. Compute the first K eigenvectors of L_{sym} to form $U \in \mathbb{R}^{n\times K}$.
3. Row-normalise U (each row has unit ℓ_2 norm) to obtain an embedding in $\mathbb{R}^K$.
4. Apply k-means to the embedded points to obtain the final partition into K clusters.

(Alternative normalised-cut formulations based on L_{rw} are also widely used [50, 52].)

From a practical perspective, SC requires defining the similarity construction (kernel and/or graph type) and selecting its parameters (e.g., σ, and possibly a k-nearest-neighbour or ε-neighborhood graph). If the similarity scale is too small, the graph may fragment into weakly connected components; if it is too large, distinct regimes can merge. Feature scaling is therefore important because Euclidean distances directly determine the graph weights. Computationally, SC can be more expensive than k-means due to the eigen-decomposition step; in practice, this is mitigated by using sparse graphs and iterative eigensolvers, so SC is often applied to moderate-size datasets or implemented with efficient approximations when needed [52].

Given the strong EOV driven nonlinearity observed in long term bridge features, this thesis prioritizes regime-partitioning methods that can capture non-convex structure. For this reason, spectral clustering is adopted as the main regime-identification tool, while other clustering families are treated as baselines for comparison. The resulting regime-aware baseline is then coupled with a novelty-detection stage (Chapter 5) to provide an early-warning indicator with controlled false-alarm behavior.

Method	**Shape flexibility**	**K required**	**Outlier robustness**	**Typical cost**	**SHM suitability**
K-means	Spherical	Yes	Low	Low	Fast baseline regime split; sensitive to scaling and outliers
Hierarchical	Linkage-dependent	No	Prioritizesrate	High	Exploratory regime hierarchy; useful at moderate n
Density-based (DBSCAN family)	Arbitrary	No	Moderate	Moderate	Regimes + noise/outliers; sensitive to density variation
Gaussian mixture models (GMMs)	Ellipsoidal	Yes	Low	Moderate	Probabilistic regimes; soft assignments; outlier-sensitive unless robustified
Spectral clustering	Arbitrary (graph-based)	Yes	Moderate	High	Nonlinear regimes/manifolds; performance depends on graph construction

Table 2.1: Clustering methods for SHM: typical capabilities and practical trade offs.

3 | Data Acquisition for Bridge

3.1. Tools and Sensors for Data Recording

In bridge SHM, the dynamic response is most commonly recorded using piezoelectric accelerometers, while other sensor types (e.g., force-balance accelerometers and strain gauges) are also adopted depending on the target response and installation constraints [19, 26]. In addition, fibre-optic sensing based on Bragg gratings has been investigated as a promising option to expand the number of measurement channels with practical deployment benefits [26, 51].

For most full-scale bridges, excitation is typically provided by ambient sources such as traffic and wind, so dynamic measurements are often collected under operational loading rather than controlled laboratory inputs. When controlled excitation is required, it can be introduced using devices such as shakers or impact-based methods, although this is not always feasible in real bridge conditions Because long term monitoring produces large volumes of data. Practical SHM campaigns also require careful strategies for data storage, organization, and retrieval in order to support reliable trend tracking over time [8, 10].

Finally, it is important to recognize two persistent challenges in vibration based SHM. First, low-frequency global modal properties (often the easiest to identify in field tests) may show limited sensitivity to localized damage. Second, these same features can be strongly affected by environmental and operational variability, which can mask damage or generate false indications unless appropriate normalization and filtering techniques are applied. These issues, together with practical constraints such as accessibility, cost, and optimal sensor placement, must be considered when designing and operating an in-situ monitoring system [3, 26].

Time-series analysis provides a statistical framework for studying sequential data and is commonly used for preliminary data inspection, model identification, parameter estimation, model validation, feature extraction, and forecasting. A practical first step is to visualize the signal in the time domain to understand its general behavior and guide the selection of an appropriate processing strategy.

In vibration based SHM, measured responses are often classified as stationary or non-stationary depending on whether key statistical properties (e.g., mean and variance) remain approximately constant over time. Stationary records are typically associated with relatively stable operating conditions and are useful for establishing baseline behavior, while non-stationary records arise when loads or environmental conditions vary, requiring analysis methods that can handle time-varying characteristics. Recognizing these categories is important in SHM because the choice of processing and modeling strategy depends strongly on whether the response can be treated as stationary [20]. In long term bridge monitoring, non-stationarity is frequently driven by EOV; therefore, regime-aware modeling is a practical strategy to reduce EOV masking effects before novelty detection.

3.2. Frequency-domain and Time-frequency Analysis for Feature Extraction

To interpret acceleration time histories, frequency-domain analysis is often employed by transforming the signal to identify dominant frequency components and track shifts that may reflect changes in structural properties. A widely used approach is the Fast Fourier Transform (FFT), which can be used to compute spectral representations such as the power spectral density (PSD) and to identify dominant frequency components and their associated energy content.

In addition, time-frequency analysis is useful when frequency content varies over time, since it allows the evolution of spectral components to be examined over specific intervals. Tracking the evolution of frequency characteristics and comparing them with a baseline can help highlight deviations that may be consistent with damage mechanisms such as cracking, joint loosening, or corrosion [8, 39]. However, the same frequency features may also vary substantially under EOV, which motivates baseline segmentation and EOV aware modeling before novelty detection.

3.3. Converting Sensor Data into SHM Insights

It is important to emphasize that sensors do not measure damage directly. Instead, they record the structural response to environmental and operational actions, or to controlled inputs when actuators are used within the monitoring setup. Whether these measurements can indicate the existence, location, or progression of damage depends on the selected sensing technology and on the damage mechanism under investigation [8, 26].

To convert raw measurements into meaningful information about structural condition, data interrogation procedures are required, including feature extraction and statistical modeling for classification and decision-making. For an SHM system to be effective in practice, the sensing hardware, acquisition strategy, and processing pipeline should be designed together, so that the collected data are well-suited to the intended interpretation and assessment tasks [10, 19].

4 | Case studies

This chapter applies the proposed unsupervised statistical-learning framework to two representative bridge structures. The first case, the KW51 steel bowstring railway bridge in Belgium, provides a long term operational dataset with real traffic, environmental variability, and a documented retrofit intervention, but without controlled damage scenarios. The second case, the Z24 post-tensioned concrete box-girder bridge in Switzerland, offers a benchmark campaign with intentionally induced progressive damage under carefully monitored conditions. The combination of these two case studies enables complementary validation: robustness under realistic operational complexity and sensitivity to well-documented damage evolution in the presence of strong EOV [8, 17, 20].

Accordingly, the workflow adopted in this thesis can be summarized as:

1. acquisition (or availability) of vibration-derived features (modal quantities),
2. data cleaning and standardization using a baseline (healthy) segment,
3. regime identification to mitigate EOV by comparing observations against an appropriate reference,
4. novelty scoring and thresholding to trigger early-warning indications.

4.1. First Case Study: KW51 Steel Railway Bridge (Belgium)

To assess the capability of the proposed methodology, it is applied to the dynamic response of the KW51 steel bowstring railway bridge. This railway structure forms part of line L36N between Leuven and Brussels in Belgium and spans the Leuven–Dijle canal with an overall length of 115 m and a width of 12.4 m. An overview of the bridge is presented in Figure 4.1. Since 2 October 2018, the bridge has been instrumented with a permanent SHM system comprising vibration and environmental sensors, from which acceleration time histories and associated environmental quantities are continuously recorded, in line with modern SHM practice for railway bridges [24, 28, 29].

In the present study, the analysis is restricted to modal-frequency data obtained between 2 October 2018 and 15 May 2019, a period during which the bridge remained in its pre-retrofitted configuration. This allows the influence of environmental variability on the identified modal properties to be examined in isolation. Although the full KW51 campaign includes a retrofitting phase, this thesis uses only the pre-intervention segment for baseline modeling and robustness assessment; the intervention period is not used for training in the analyses reported here [28, 29].

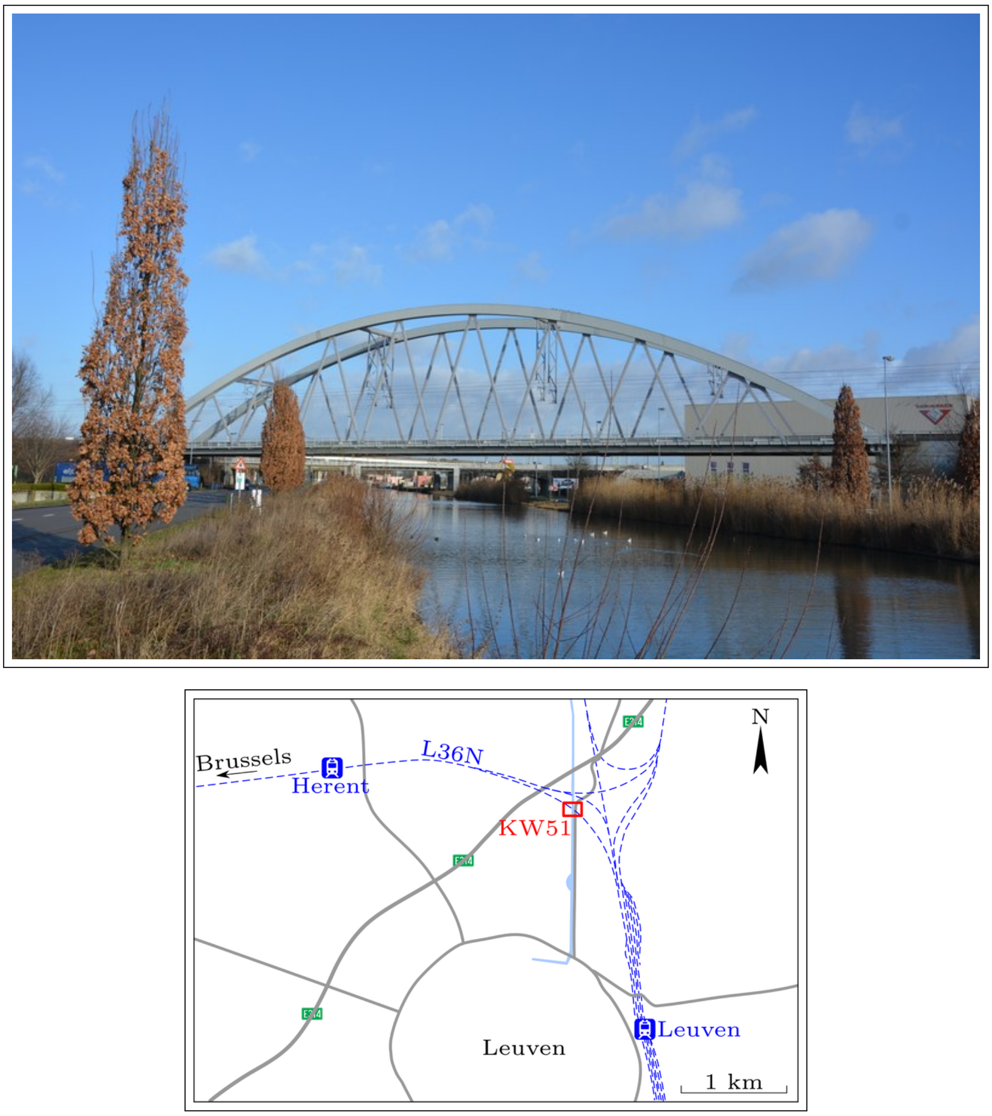


Figure 4.1: KW51 steel bowstring railway bridge (top) and schematic situation sketch (bottom) [24].

4.1.1. Modal Frequencies of KW51

An OMA procedure developed for KW51 is used to extract the modal properties of the bridge, including its natural frequencies [29]. The procedure initially identified 14 modes; however, several of these exhibited substantial gaps in the time series, rendering them unsuitable for use in data-driven models owing to the limited number of complete observations, a common challenge in continuous SHM datasets. The study, therefore, concentrates on vertical modes 6, 10, 12, and 13, which contained comparatively few missing values. After removing the remaining incomplete entries, a dataset of 2,688 modal-frequency estimates in the undamaged state was obtained. Here, one sample corresponds to one OMA-based modal-frequency estimate obtained from a fixed-length time window of the continuous monitoring data. The temporal evolution of these modes is depicted in Figure 4.2, where pronounced jumps are observed between samples 1,345 and 2,017. While these discontinuities may be associated with sub-zero air temperatures, additional patterns indicate that other environmental mechanisms also affect the modal frequencies, consistent with earlier observations on EOV effects in bridges. A correlation study reported for KW51 revealed only weak linear dependence between air temperature and the identified frequencies, reinforcing the presence of additional variability sources [28, 29]. These findings motivate the present work, which aims to mitigate environmental influences and derive normalized dynamic features for structural assessment using advanced unsupervised learning techniques.

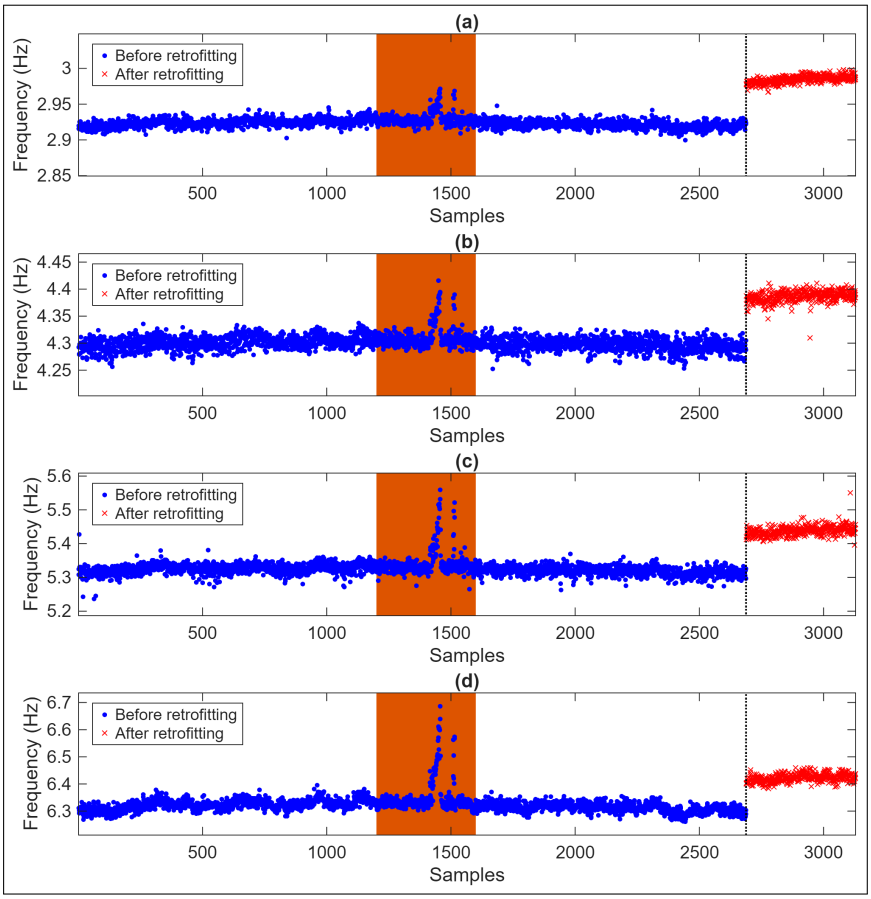


Figure 4.2: Eigenfrequencies of the KW51 bridge for the selected modes.

4.2. Second Case Study: Z24 Concrete Box-Girder Bridge (Switzerland)

The Z24 bridge is one of the most extensively studied benchmark structures in the SHM community and remains a key reference for vibration based damage detection. The bridge was located in the Canton of Bern (Switzerland), carrying a road over the River Emme near the village of Koppigen. Before its demolition in the late 1990s, an intensive monitoring and controlled damage campaign was performed, resulting in a publicly documented dataset that captures both long term environmental effects and a complete sequence of artificial damage scenarios [8, 23, 38].

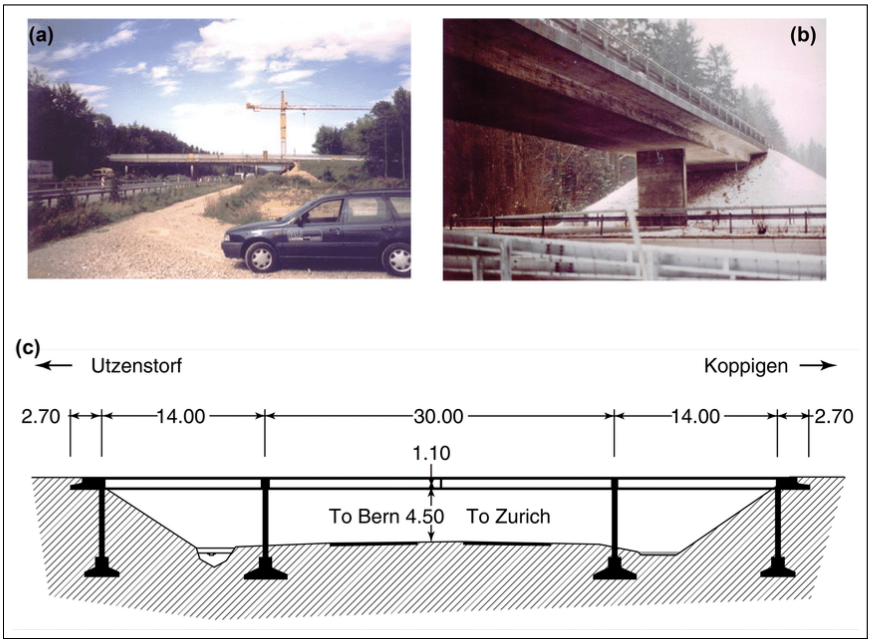


Figure 4.3: Z24 bridge: (a) general view, (b) close view of the deck, and (c) one of the piers [23].

Z24 is a simply supported, post-tensioned concrete box-girder with two internal cells. The main span is approximately 30 m, complemented by shorter approach spans of about 14 m. The deck width is roughly 7.4 m and the structural depth is around 1.2 m. The superstructure rests on concrete abutments and piers via concrete hinges and elastomeric bearings, resulting in a statically determinate system that nevertheless exhibits complex dynamic behavior due to the interaction between the girder, supports, and foundation soil. The cross-section consists of a two-cell box with internal webs, transverse diaphragms, and an asphalt wearing course. Temperature gradients through the cross-section, together with traffic loads and support movements, lead to significant variability in the measured dynamic properties. Progressive damage scenarios were introduced in a controlled, stepwise manner over roughly one month, shortly before the bridge was demolished, as summarized in Table 4.1. The selected scenarios were designed to reflect practical safety-related deterioration mechanisms and other damage patterns commonly observed in service [23, 38].

Table 4.1: Progressive damage scenarios applied to the Z24 bridge [23].

No.	Date (1998)	Scenario / description
1	04 Aug	1st reference measurement (undamaged condition).
2	09 Aug	2nd reference measurement (after installing lowering system).
3	10 Aug	Lowering of pier by 20 mm (settlement of subsoil / erosion).
4	12 Aug	Lowering of pier by 40 mm (settlement of subsoil / erosion).
5	17 Aug	Lowering of pier by 80 mm (settlement of subsoil / erosion).
6	18 Aug	Lowering of pier by 95 mm (settlement of subsoil / erosion).
7	19 Aug	Lifting of pier (tilt of foundation).
8	20 Aug	3rd reference measurement (after lifting back to initial condition).
9	25 Aug	Spalling of concrete at soffit (12 m^2): vehicle impact / carbonization / reinforcement corrosion.
10	26 Aug	Spalling of concrete at soffit (24 m^2): vehicle impact / carbonization / reinforcement corrosion.
11	27 Aug	Landslide of 1 m at abutment (heavy rainfall / erosion).
12	31 Aug	Failure of concrete hinge (chloride attack / corrosion).
13	02 Sep	Failure of 2 anchor heads (corrosion / overstress).
14	03 Sep	Failure of 4 anchor heads (corrosion / overstress).
15	07 Sep	Rupture of 2 out of 16 tendons (forgotten injection / chloride influence).
16	08 Sep	Rupture of 4 out of 16 tendons (forgotten injection / chloride influence).
17	09 Sep	Rupture of 6 out of 16 tendons (forgotten injection / chloride influence).

4.2.1. Modal Frequencies of Z24

Low-temperature periods introduce a pronounced source of variability in the measured eigenfrequencies of the Z24 bridge. Around samples 400–1600, corresponding to air temperatures below 0 °C, the nominally intact frequencies exhibit abrupt fluctuations (see Figure 4.4). Such behavior can compromise the reliability of vibration based assessments and lead to misinterpretation in bridge condition evaluation. Consequently, freezing conditions should be explicitly accounted for, either by excluding these periods from baseline learning or by applying robust normalization methods that reduce their influence on the modal features.

In the Z24 benchmark, each sample corresponds to one measurement instance in the published monitoring campaign, for which modal frequencies were identified and released as part of the benchmark dataset [23, 38].

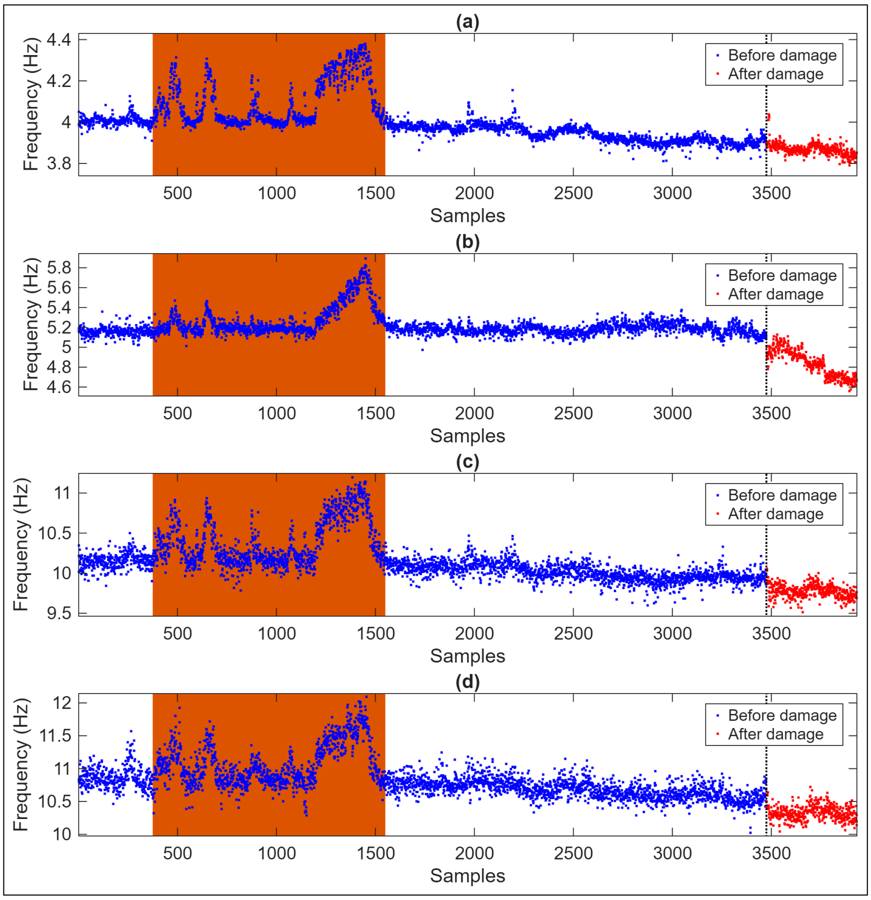


Figure 4.4: Eigenfrequencies of the Z24 bridge for the selected modes.

Dataset Summary as Used in This Thesis

4.2.2. Dataset Summary

This thesis evaluates the proposed workflow on two benchmark bridge-monitoring datasets, selected to represent complementary conditions. The first dataset concerns the KW51 steel bowstring railway bridge in Belgium. In this study, only the pre-intervention (undamaged) segment is used, and the input features consist of the modal frequencies of vertical modes 6, 10, 12, and 13. After removing entries affected by missing values, the resulting dataset contains 2,688 samples, all of which are treated as healthy observations. This segment is characterized by strong EOV, making it suitable for assessing regime separation and false-alarm control under realistic long term monitoring conditions [28, 29].

The second dataset is the Z24 post-tensioned concrete box-girder bridge in Switzerland, which combines long term EOV with a well-documented, progressive damage campaign.

The features used here are the modal frequencies of four selected modes, yielding a total of 3,932 samples. In the thesis implementation, a baseline period prior to the damage scenarios is defined as the healthy reference set, comprising 3,475 healthy samples, while the remaining samples correspond to stages affected by the controlled damage progression [23, 38].

From a methodological perspective, the two case studies serve distinct but complementary purposes. The KW51 dataset primarily stress-tests the robustness of EOV mitigation and anomaly detection under real operational complexity—long term environmental trends, transient train passages, missing data, and sensor noise—but without controlled damage scenarios in the selected segment, so the emphasis is on maintaining stability and avoiding over-reaction to non-damage variability [12, 44]. In contrast, the Z24 benchmark enables a more quantitative evaluation because the timing and progression of the induced damage are documented with high confidence, allowing detection timing and false-alarm/missed-detection behavior to be benchmarked directly against the reported scenarios [23, 38, 43]. Taken together, this dual-case design also probes generalisability across different materials (steel versus concrete), structural forms, and climate/loading conditions, which is essential for deploying data-driven SHM strategies beyond a single bridge type [10, 15]. These two case studies, therefore, provide a coherent experimental framework for validating the proposed unsupervised statistical-learning approach in the subsequent chapters.

5 | Proposed TSC–CAMSD Approach

5.1. Introduction

The main difficulty in vibration based SHM is that changes in measured features do not arise exclusively from structural damage. EOV can induce shifts in modal parameters that resemble damage, potentially leading to false alarms or missed detections if the variability is not treated explicitly [7, 38]. This challenge is particularly relevant for bridge monitoring, where operational loads and ambient conditions evolve continuously, and the available damage-labeled data are typically limited [8, 9].

To address these issues, this chapter introduces the proposed TSC–CAMSD approach, an unsupervised pipeline that combines: (i) a regime-identification step based on spectral clustering of baseline data TSC, and (ii) a regime-conditioned novelty index based on an adaptive, conditioning-guided Mahalanobis squared distance CAMSD [41, 43, 45]. The clustering step aims to separate operating conditions and reduce the confounding effect of EOV, while the novelty detector provides a scalar index that can be thresholded to trigger an early warning [13]. The approach is intended to be applicable to both case studies considered in this thesis (Z24 and KW51) without requiring damage-labeled training data.

5.2. Overview of the Proposed Approach

The proposed TSC–CAMSD method consists of two main steps:

1. **Training feature partitioning (TSC):** baseline features are clustered into distinct regimes. This step groups samples that share similar operating conditions (e.g., temperature/traffic-driven states), so that subsequent decisions are made relative to an appropriate local reference rather than a single global baseline [7, 41].
2. **Novelty detection (CAMSD):** Computed for new observations with respect to a

learned reference set. This produces an anomaly score that increases when a sample deviates from healthy baseline behavior [13, 43].

Algorithm 5.1 TSC–CAMSD

Input: baseline features $X \in \mathbb{R}^{p \times n_{\text{Tr}}}$; monitoring features $Z \in \mathbb{R}^{p \times n_{\text{Te}}}$ (first n_{Vd} healthy); target FPR α (%).

Output: novelty scores (d_1, d_2, DI), threshold τ, alarms.

(I) TSC on baseline (regime discovery)

Standardize $X^\top$ and compute pairwise distances D^2; set $\sigma_0 \leftarrow \text{median}(\sqrt{D^2_{ij}})$.

Grid-search (k, σ); build affinity $S = \exp(-D^2/(2\sigma^2))$ and run spectral clustering on S.

Select $(k^\star, \sigma^\star)$ by maximizing mean silhouette on the embedding; obtain labels `idx`.

Form clusters $\{\texttt{Cluster}\{r\}\}$ and set regime reference `refPts` $\leftarrow \texttt{Cluster}\{1\}^\top$.

(II) CAMSD training (baseline self-scoring)

Compute $d_1 \leftarrow \texttt{CAMSD}(X^\top, X^\top, \texttt{opts})$ (with conditioning-adaptive k and regularized QR).

Set $\tau_{\text{Tr}} \leftarrow Q_{1-\alpha/100}(d_1)$.

(III) CAMSD monitoring (validation-calibrated threshold)

Compute $d_2 \leftarrow \texttt{CAMSD}(Z^\top, \texttt{refPts}, \texttt{opts})$.

Calibrate $\tau \leftarrow Q_{1-\alpha/100}\big(d_2(1{:}n_{\text{Vd}})\big)$ (healthy-only).

(IV) Fusion and decision

$DI \leftarrow [d_1; d_2]; \qquad \text{alarm}(t) \leftarrow \mathbb{1}\{DI(t) > \tau\}$.

Algorithm 5.2 CAMSD

Input: `testPts` $\in \mathbb{R}^{m \times p}$, `refPts` $\in \mathbb{R}^{n \times p}$, options `opts` (incl. $k_{\min}, k_{\max}$, z-score, `minRcond`, auto-ridge).

Output: scores `d2` $\in \mathbb{R}^m$ and diagnostics (`usedK`, `rcondR`, λ).

(1) Standardize fit (μ, σ) on `refPts`; compute scaled `refZ`, `testZ`.

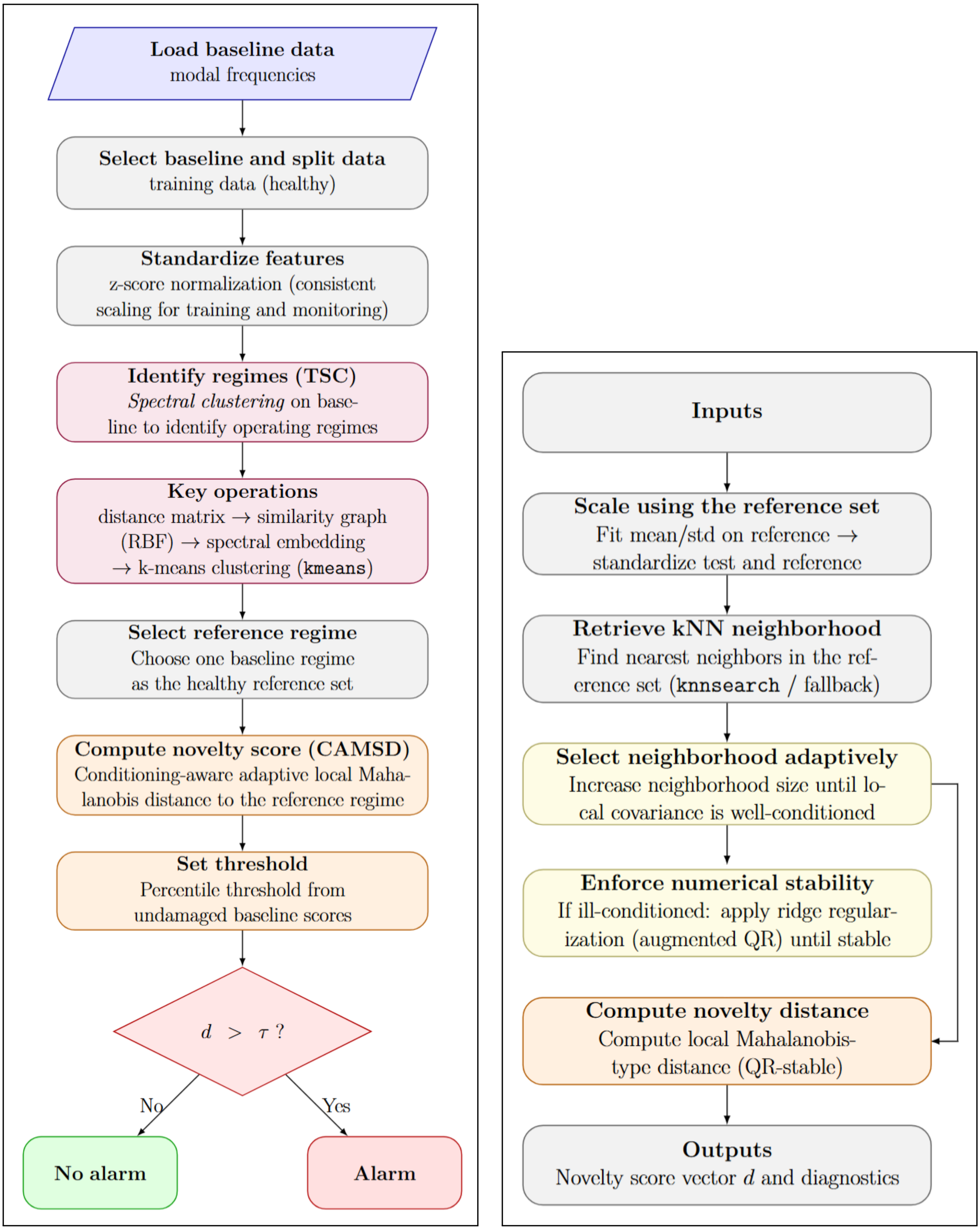
(2) Retrieve candidates for each test point, find $k_{\max}$ nearest neighbors in `refZ` (e.g., `knnsearch`).

(3) Select local k (conditioning-guided) increase k from $k_{\min}$ until $\text{rcond}(R) \geq$ `minRcond` for the (regularized) QR factor R of the centered neighborhood; store `usedK` and λ.

(4) Score with neighbors of size $\hat{k}$:

$$\delta = (\texttt{testZ} - \mu_{\text{loc}})^\top, \quad u = R^\top \backslash \delta, \quad \texttt{d2} = (\hat{k} - 1)\, u^\top u.$$

Return `d2` and diagnostics.



(a) Simplified TSC–CAMSD workflow. (b) Simplified CAMSD internals.

Figure 5.1: Overview of the proposed TSC–CAMSD framework. Left: global workflow from baseline clustering to novelty scoring. Right: internal CAMSD scoring mechanism with conditioning control and adaptive regularization.

5.3. Explanation of the Functioning Process

(A) Training feature partitioning by TSC

The first component partitions baseline features into a small number of representative regimes. In bridge monitoring, this partitioning is useful because the "healthy" feature space is rarely unimodal: EOV often produces multiple clusters or manifolds corresponding to different operating states [7]. By performing clustering on baseline data, the method reduces the risk of comparing new samples against an inappropriate reference, which is a frequent cause of false positives in long term monitoring [45].

Let $\{\mathbf{x}_i\}_{i=1}^{n}$ be standardized baseline feature vectors with $\mathbf{x}_i \in \mathbb{R}^p$. TSC constructs a similarity graph using a Gaussian (RBF) kernel,

$$w_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{2\sigma^2}\right), \tag{5.1}$$

where σ controls the similarity scale [54]. Let $\mathbf{W} = [w_{ij}]$ be the similarity matrix and $\mathbf{D} = \mathrm{diag}(d_i)$ the degree matrix with $d_i = \sum_{j=1}^{n} w_{ij}$. The symmetric normalized graph Laplacian is defined as

$$\mathbf{L}_{\mathrm{sym}} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}. \tag{5.2}$$

TSC computes the K eigenvectors associated with the smallest eigenvalues of $\mathbf{L}_{\mathrm{sym}}$, stacks them into an embedding matrix $\mathbf{V} \in \mathbb{R}^{n\times K}$, and clusters the rows of $\mathbf{V}$ (e.g., by K-means) to obtain regime labels for baseline samples [54]. This is advantageous in SHM because temperature- and load-driven feature trajectories can be curved or intertwined, where purely centroid-based clustering in the original feature space may be ineffective [38]. Implementation note (reproducibility): The hyperparameters K and σ (or any alternative graph construction parameters) must be reported explicitly, including the selection criterion, tested ranges, and final values used in each case study.

(B) Novelty Detection by CAMSD

After regime partitioning, novelty detection is carried out using CAMSD. The CAMSD index is built on the squared Mahalanobis distance concept, but it is adapted to practical SHM conditions by using local neighborhoods and conditioning checks [43]. For each query sample, CAMSD selects a local set of baseline neighbours and computes a covariance-adjusted distance that accounts for feature scaling and correlation. The neighborhood size is selected adaptively within a prescribed range, and numerical stability is enforced by checking the conditioning of the local covariance estimate and applying regularization when needed [43].

Let $\mathbf{z} \in \mathbb{R}^p$ be a standardized query feature vector. For a given neighborhood size k, define a local baseline set $\mathcal{N}_k(\mathbf{z})$ (e.g., the k nearest neighbours in the selected reference set). Let $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ denote the sample mean and covariance estimated from $\mathcal{N}_k(\mathbf{z})$. CAMSD computes

$$d^2(\mathbf{z}) = (\mathbf{z} - \boldsymbol{\mu}_k)^\top \boldsymbol{\Sigma}_{k,\lambda}^{-1} (\mathbf{z} - \boldsymbol{\mu}_k), \tag{5.3}$$

where $\boldsymbol{\Sigma}_{k,\lambda}$ is a regularized covariance (to protect inversion under ill-conditioning), for instance

$$\boldsymbol{\Sigma}_{k,\lambda} = \boldsymbol{\Sigma}_k + \lambda \mathbf{I}, \tag{5.4}$$

with $\lambda \geq 0$ selected (or increased) when the conditioning of $\boldsymbol{\Sigma}_k$ is insufficient [43]. Large values of $d^2(\mathbf{z})$ indicate a deviation from the baseline pattern and are interpreted as potential novelty. Compared to a single global covariance model, the local and adaptive nature of CAMSD improves robustness when baseline data are heterogeneous or exhibit multiple regimes due to EOV [14, 43].

Implementation note (Reproducibility); The neighborhood range $k \in [k_{\min}, k_{\max}]$, the conditioning criterion (e.g., minimum reciprocal condition number), and the regularization policy (initial λ and auto-regularization rule) must be stated clearly and matched to the implementation.

5.4. Unsupervised operation vs. label-informed tuning

In principle, the proposed TSC–CAMSD framework is fully *unsupervised*, because regime discovery (TSC) and novelty scoring (CAMSD) rely only on unlabeled feature observations, and the decision threshold can be calibrated using a healthy-only segment. However, in this thesis the algorithm is also evaluated under a *label-informed tuning* setting to enable fair benchmarking and to explore the sensitivity of performance to key hyperparameters (e.g., $k_{\min}$, $k_{\max}$, σ, and α). In particular, when damage timing (or an event onset) is available, these parameters are selected by minimizing missed detections (FNR) while constraining false alarms (FPR) on healthy data; this procedure uses the labels only for *model selection*, not for training the representation or the novelty index. After tuning, the final configuration is fixed and applied in a strict hold-out manner to compute reported metrics (FPR, FNR, and detection delay). To avoid ambiguity, throughout the remainder of the thesis we distinguish between: (i) a *deployment mode* where hyperparameters are chosen *a priori* and only healthy data are used for calibration, and (ii) a *benchmarking mode* where labels are used exclusively to select hyperparameters and quantify performance, without altering the unsupervised nature of the underlying detection mechanism.

5.5. Thresholding and Decision Rule

Threshold estimation is a critical component of unsupervised novelty detection because it defines how anomaly scores are mapped to final decisions [9]. In the proposed approach, the threshold is determined from baseline anomaly scores using a non-parametric, quantile-based rule. This avoids imposing a strict parametric distribution on the scores and supports practical use under EOV [13].

Let $\{d_i\}$ denote anomaly scores computed on baseline (undamaged) data, and let α be a target FPR. The decision threshold τ is defined as

$$\tau = Q_{1-\alpha}\big(\{d_i\}\big),$$

where $Q_{1-\alpha}(\cdot)$ denotes the empirical $(1-\alpha)$-quantile [13]. The resulting decision rule is

$$d > \tau \;\Rightarrow\; \text{alarm (abnormal)}, \qquad d \leq \tau \;\Rightarrow\; \text{normal}.$$

This thresholding strategy is straightforward to implement and can be calibrated to meet a desired rate of false alarms on baseline data.

5.6. Novelty Detection Protocol

Novelty detection with TSC–CAMSD follows a baseline–monitoring structure typical of unsupervised SHM. During the baseline stage, the clustering model is learned, and CAMSD scores are computed to define a reference distribution for threshold estimation. During monitoring, novelty scores are computed for each new sample relative to a baseline reference set [9, 45].

Regime Conditioning and Reference-set Choice. In the general TSC–CAMSD framework, each new sample can be associated with the most compatible baseline regime (e.g., via nearest regime centroid in the spectral embedding or by maximum similarity), and CAMSD can then be computed using regime-specific reference points. In the numerical experiments reported in this thesis, CAMSD is evaluated with respect to a selected baseline reference set derived from the TSC partitioning, which provides a practical regime-conditioned baseline while keeping the monitoring stage computationally simple.

This regime-aware evaluation is important in bridge monitoring because it reduces the chance that normal regime shifts are interpreted as damage. At the same time, CAMSD remains sensitive to deviations that are not explainable by baseline variability, thereby supporting early warning when the structure departs from its normal patterns [16, 43].

5.7. Results Comparison Metrics

To compare the performance of the proposed method against alternative approaches, results are summarized using detection metrics derived from thresholded anomaly scores [9]. When damage timing/labels are available (e.g., Z24), the evaluation includes:

- False positive rate (FPR): fraction of baseline (undamaged) samples incorrectly flagged as alarms.
- False negative rate (FNR): fraction of damaged samples not flagged as alarms.
- Detection delay: the novelty index rises and remains sustained above the threshold for an extended period.

For datasets without damage labels (e.g., KW51 during the selected period), the evaluation focuses on baseline stability (low false-alarm behavior), consistency with known environmental trends, and qualitative interpretation of detected anomalies. Quantitative comparisons and final performance values are reported in Chapter 6.

6 | Results Validation and Comparative Assessment

6.1. Experimental Protocol and Reproducibility Details

This chapter validates the proposed unsupervised regime-partitioning and novelty-detection workflow on two bridge case studies (KW51 and Z24). All clustering results are obtained on the same multivariate feature representation used throughout the thesis, consisting of $v = 4$ modal-frequency features (four selected modes per bridge). To avoid leakage, all preprocessing statistics are computed on the training (baseline) subset only, and then applied to validation/test data.

Table 6.1: Data partitions and target FPR used in this thesis.

KW51 Dataset

Quantity	Value
Total samples n_{Fr}	3129
Modes n_{Mode}	4
Training samples n_{Tr}	2016
Healthy samples n_{He}	2688
Validation n_{Vd}	672
Testing samples n_{Te}	1113
Target FPR α	0.5%

Z24 Dataset

Quantity	Value
Total samples n_{Fr}	3932
Modes n_{Mode}	4
Training samples n_{Tr}	3145
Healthy samples n_{He}	3475
Validation n_{Vd}	330
Testing samples n_{Te}	787
Target FPR α	0.5%

Feature Standardization and Baseline-only Training

Let $\mathbf{x}_t \in \mathbb{R}^v$ denote the feature vector at sample t. Standardization is performed using the training-set mean $\boldsymbol{\mu}_{\mathrm{tr}}$ and standard deviation $\boldsymbol{\sigma}_{\mathrm{tr}}$:

$$\tilde{\mathbf{x}}_t = \frac{\mathbf{x}_t - \boldsymbol{\mu}_{\mathrm{tr}}}{\boldsymbol{\sigma}_{\mathrm{tr}}}.$$

All clustering models are fitted on standardized baseline (undamaged) training data, and the resulting partitions are then inspected on the full record to assess operating-regime structure. For fair comparison across clustering families, the dominant regime split is reported consistently (two-regime partitions are shown in the figures), while method-specific behavior (e.g., DBSCAN noise points, GMM soft membership) is discussed explicitly.

Hyperparameters and Model-selection Reporting

For transparency, all clustering methods are reported together with the hyperparameter that most strongly governs their behavior. For K-means, GMMs, and spectral clustering, the primary modeling choice is the number of clusters K (in this chapter, the figures show the dominant split as $K = 2$ for consistent comparison), together with standard robustness settings such as multiple random initializations. For hierarchical clustering, the outcome is determined by the selected distance metric and linkage rule; the displayed partition corresponds to a coarse dendrogram cut chosen to represent the dominant high-level split. For DBSCAN, clustering is controlled by the neighborhood radius ε and the minimum neighbor count MinPts, which define density reachability and determine whether sparse observations are assigned to a secondary cluster or labeled as noise. Finally, for spectral clustering, in addition to K, the similarity-graph construction and kernel scale are explicitly stated because they directly control graph connectivity and therefore the resulting partition [53, 54]. These hyperparameter values are reported exactly as used in the MATLAB implementation to ensure full reproducibility.

6.2. Experimental Assessment on the KW51 Bridge Dataset

This section reports the clustering results obtained for the KW51 bridge case study using different unsupervised techniques. The objective is to assess how each method partitions the feature space into operating regimes, and how clearly the resulting partitions separate distinct behavioral patterns in the measured features.

6.2.1. Data Clustering for the KW51 Bridge Using K-means

Figure 6.1 shows that K-means produces a hard partition of the data into a small number of regimes. One regime captures the dominant baseline behavior, while the other concentrates samples associated with pronounced deviations from that baseline. This behavior is consistent with the centroid-based nature of K-means: samples are assigned to the nearest representative centre, which typically yields compact partitions and can be sensitive when regimes are not well represented by spherical geometry. Colours indicate cluster assignments learned on baseline training data and shown across the full record to illustrate regime consistency.

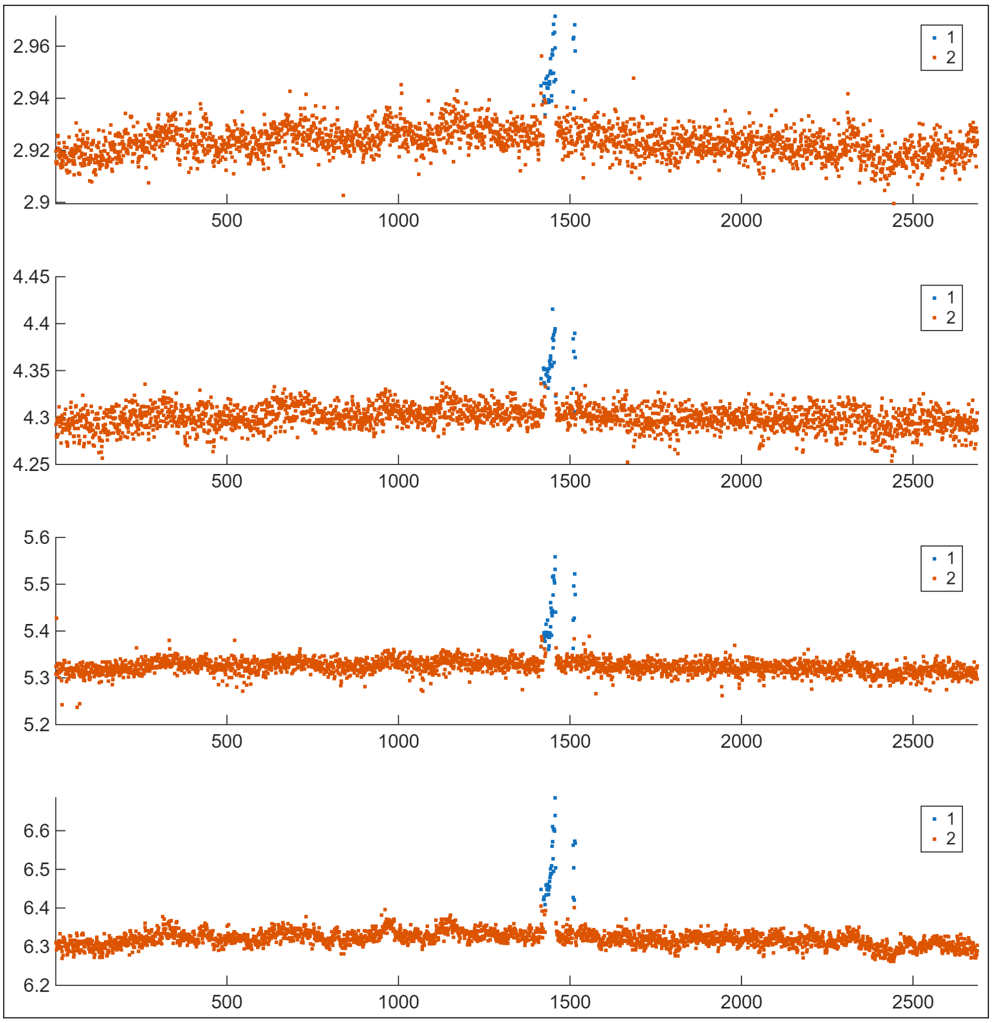


Figure 6.1: K-means clustering for KW51.

6.2.2. Data Clustering for the KW51 Bridge Using Hierarchical Clustering

The hierarchical clustering result in Figure 6.2 indicates a two-regime structure comparable to the K-means partition, but obtained through successive merges governed by the linkage rule rather than by centroid updates. The dendrogram in Figure 6.3 visually supports this separation by showing a clear high-level split into two main branches, suggesting that the dataset contains two dominant groups at a coarse resolution. This is useful for regime exploration in SHM because the cut level can be adjusted to inspect clustering structure across multiple resolutions. The dominant high-level split supports a two-regime interpretation at coarse resolution and the shown partition corresponds to the dominant high-level split observed in the dendrogram.

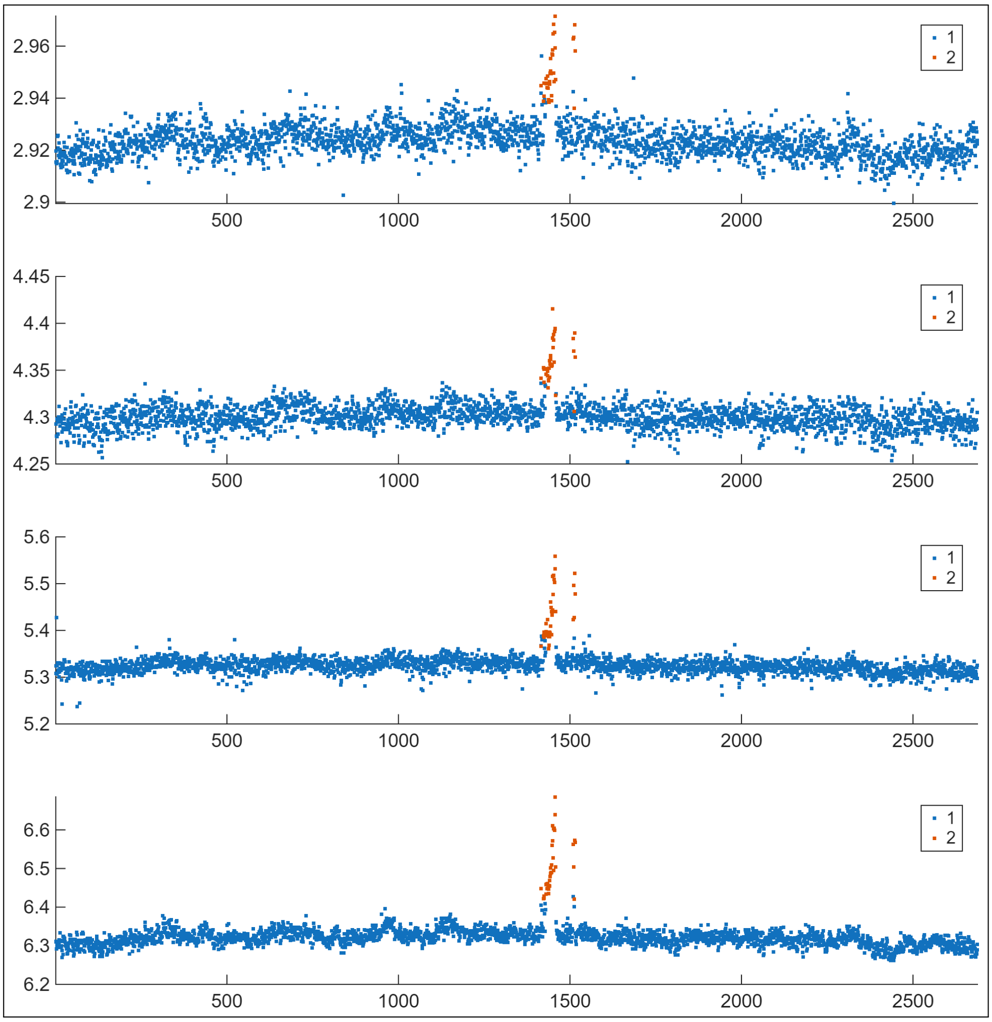


Figure 6.2: Hierarchical clustering for KW51

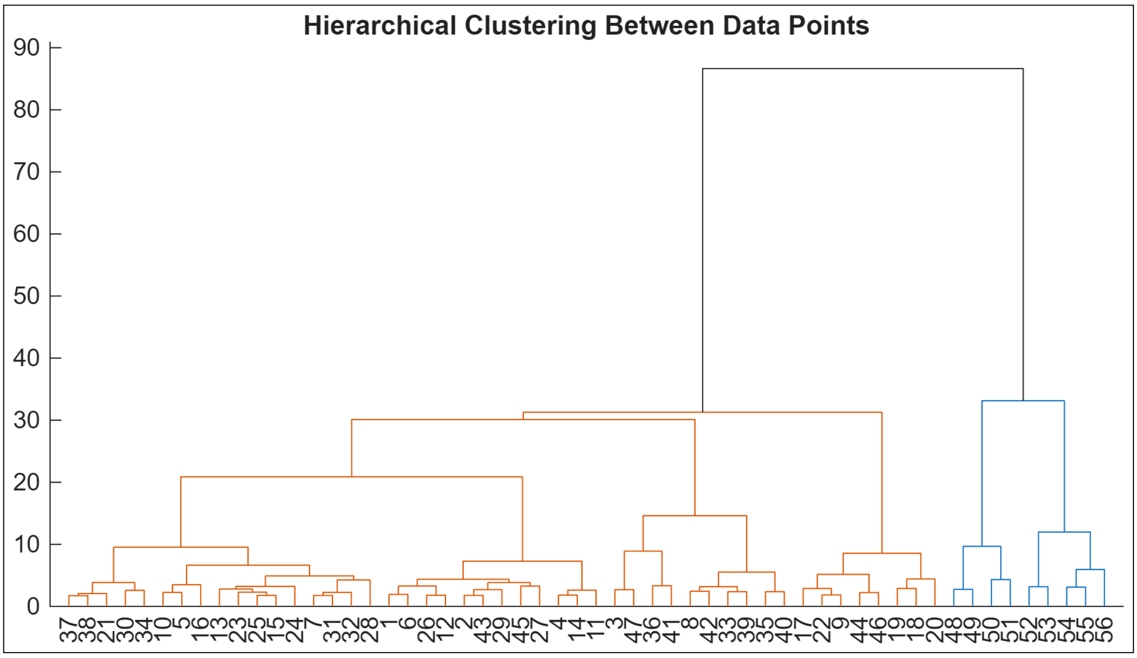


Figure 6.3: Dendrogram for KW51 hierarchical clustering.

6.2.3. Data Clustering for the KW51 Bridge Using DBSCAN

DBSCAN (Figure 6.4) identifies dense regions as clusters while treating low-density points as noise. This behavior is particularly relevant for SHM, where isolated samples may correspond to transients, sensor artifacts, or abnormal events. In the presented result, the method yields a dominant cluster representing the main regime and a set of points that are either assigned to a secondary cluster or marked as noise, concentrated around the most distinct deviations. This highlights DBSCAN's practical advantage: it does not force every sample into a cluster, which can reduce regime contamination when outliers are present. Points labeled as noise represent low-density deviations that are not forced into a regime.

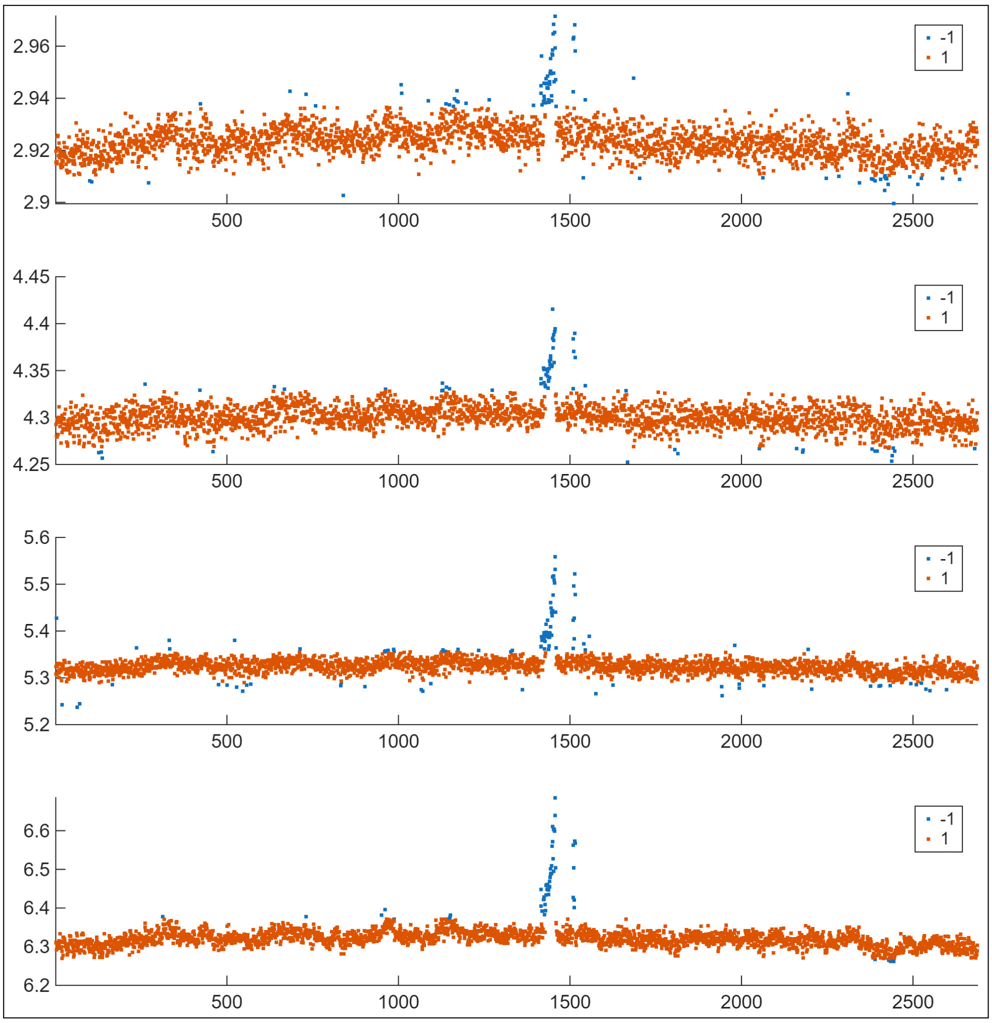


Figure 6.4: DBSCAN clustering for KW51 .

6.2.4. Data Clustering for the KW51 Bridge Using GMMs

GMMs (Figure 6.5) provide a probabilistic partitioning of the feature space, which is beneficial when regime transitions are smooth and strict hard boundaries may be unrealistic. The result indicates two dominant components that explain the observed variability and separate the baseline regime from a distinct set of deviating samples. Compared with purely geometric partitioning, this probabilistic viewpoint supports regime interpretation in terms of component likelihood and soft membership tendencies. Component assignments reflect probabilistic regime membership, reducing hard boundary effects near transitions.

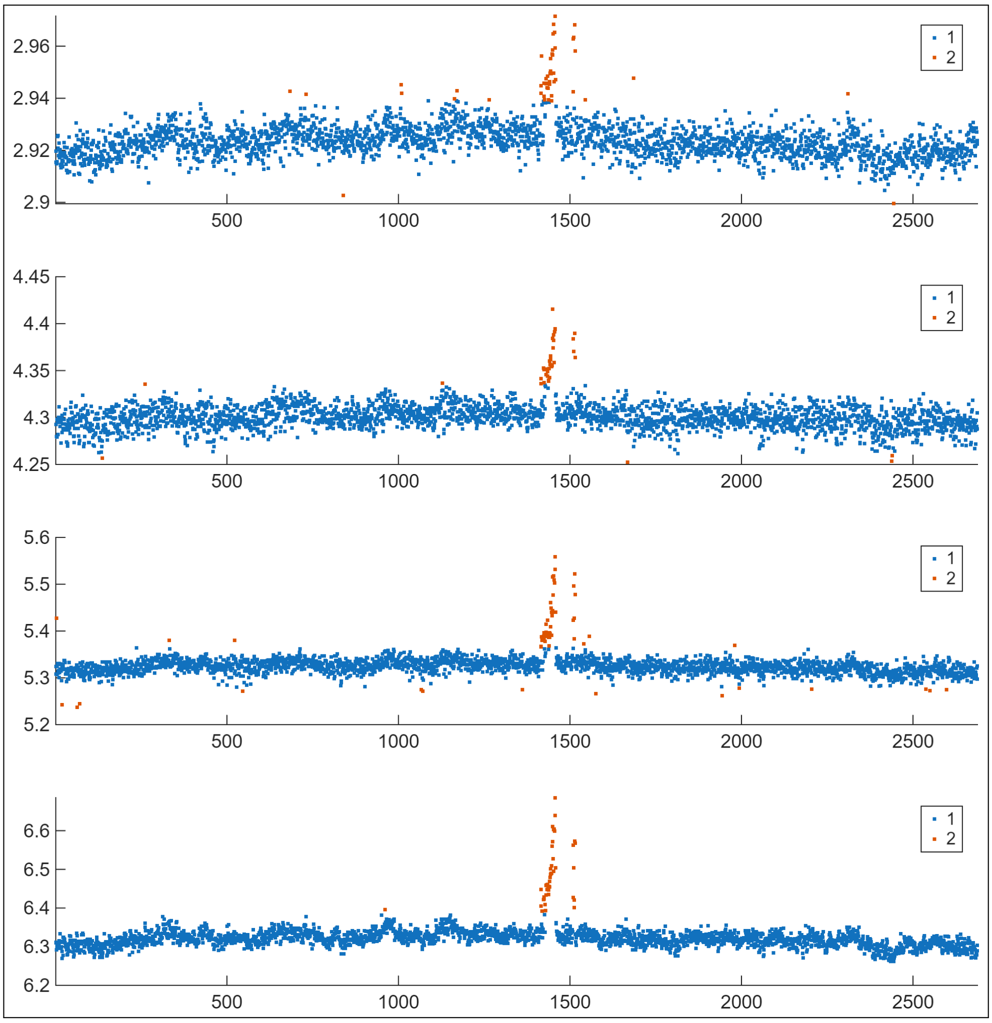


Figure 6.5: GMM regime modeling for KW51 .

6.2.5. Data Clustering for the KW51 Bridge Using SC

SC in Figure 6.6 separates regimes by exploiting the connectivity structure of a similarity graph rather than relying on convex partitions in the original feature space. This is a known advantage when regimes are not well represented by simple centroid geometry or when separations are effectively non-linear [53, 54]. The presented result shows a clear regime split aligned with the most distinct feature deviations, indicating that SC captures the global structure underlying the partition. The regime split is obtained from graph connectivity in a similarity representation [53].

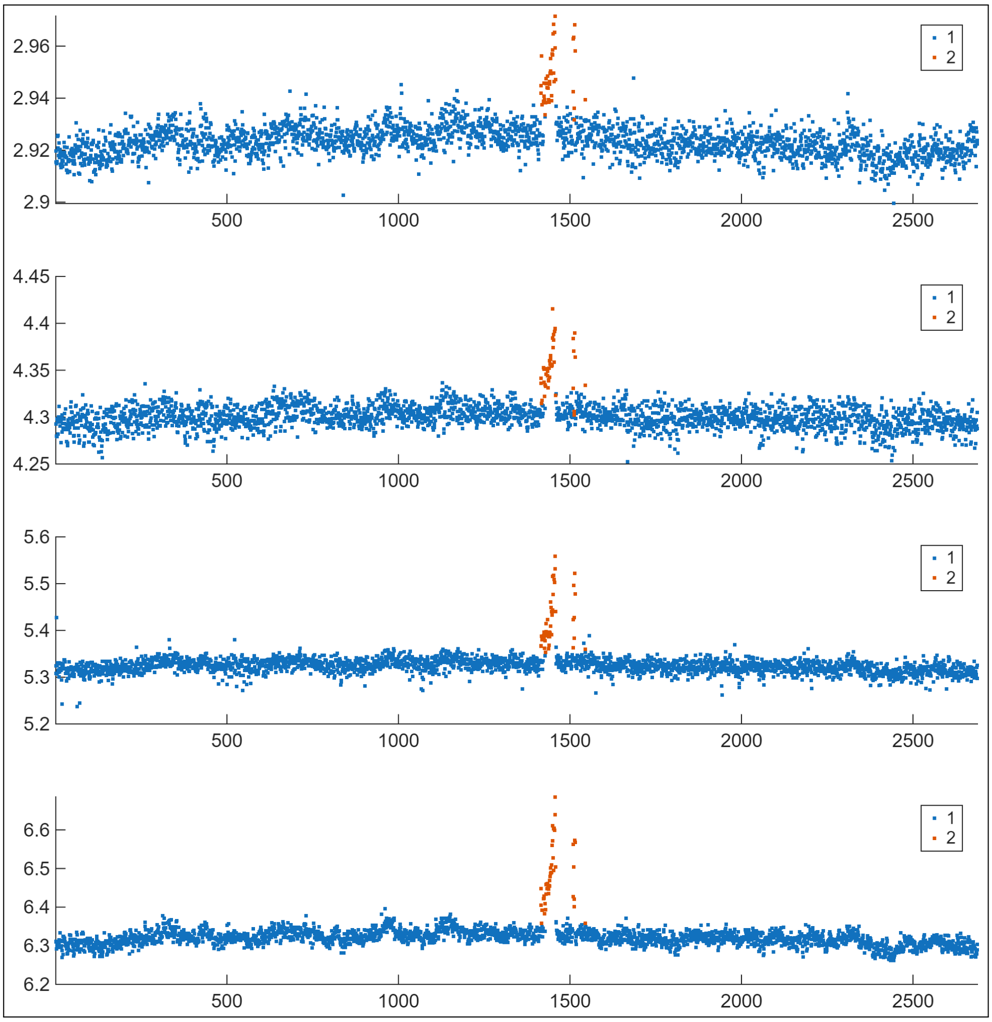


Figure 6.6: SC for KW51.

6.3. Experimental Assessment on the Z24 Bridge Dataset

This section presents clustering results for the Z24 bridge case study, using the same family of unsupervised clustering techniques as in the KW51 assessment. The goal is to evaluate how consistently each method isolates regimes associated with distinct feature behavior, and how robust the partitions appear across the feature set.

6.3.1. Data Clustering for the Z24 Bridge Using K-means

In Figure 6.7, K-means produces two regimes, where one regime accounts for the dominant baseline behavior and the other concentrates the most prominent excursions in the feature trajectories. Colors show the baseline-trained regime partition across the full record. In addition, the separation is visually coherent across the displayed features, indicating that the identified regime split is not driven by a single mode but reflects a broader multivariate change in behavior. This qualitative observation should be complemented by reporting the internal validation score and the exact K used.

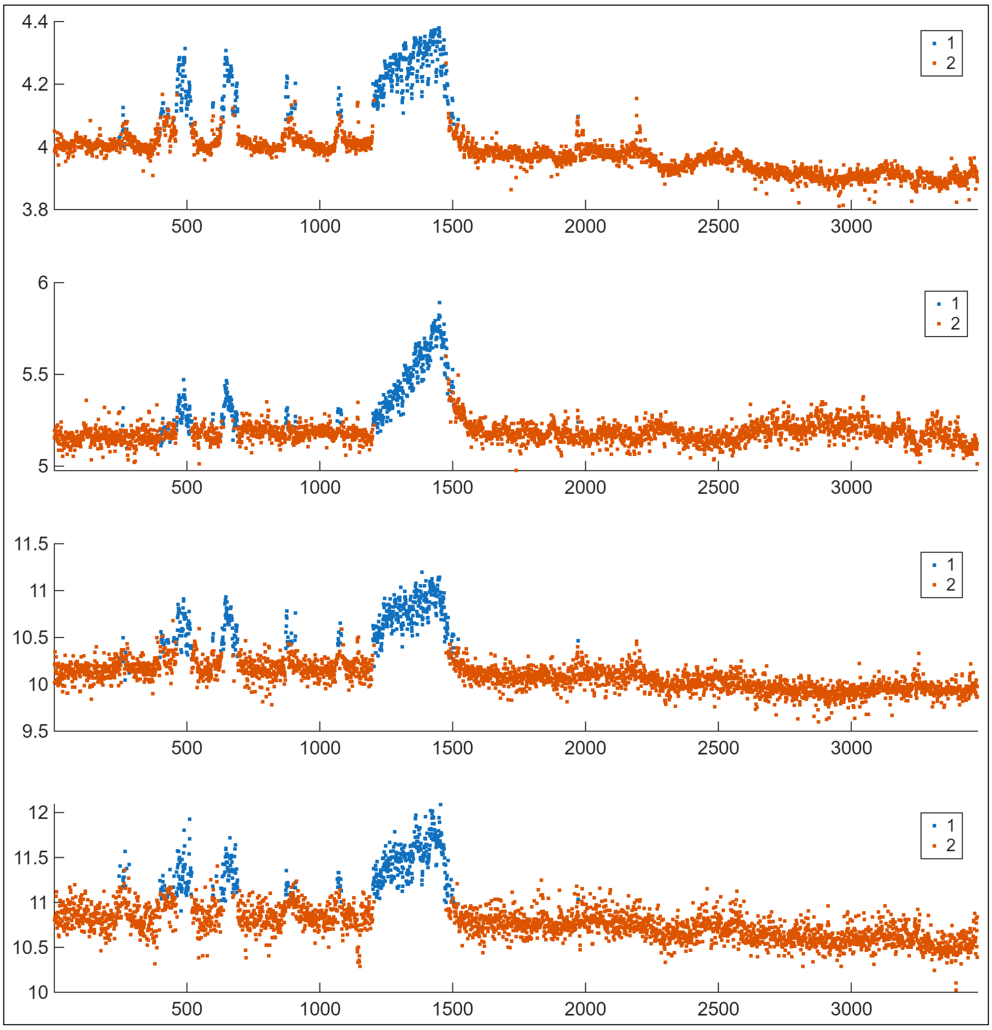


Figure 6.7: K-means clustering for Z24.

6.3.2. Data Clustering for the Z24 Bridge Using Hierarchical Clustering

The hierarchical clustering output in Figure 6.8 yields a regime split comparable to K-means while providing an explicit hierarchy of merges that can be inspected through the dendrogram representation. The dendrogram in Figure 6.9 indicates a dominant high-level split, supporting the interpretation that two principal groups explain much of the baseline variability at a coarse scale. The shown cut corresponds to the dominant high-level split observed in the dendrogram. The dominant split supports a two-regime interpretation at coarse resolution.

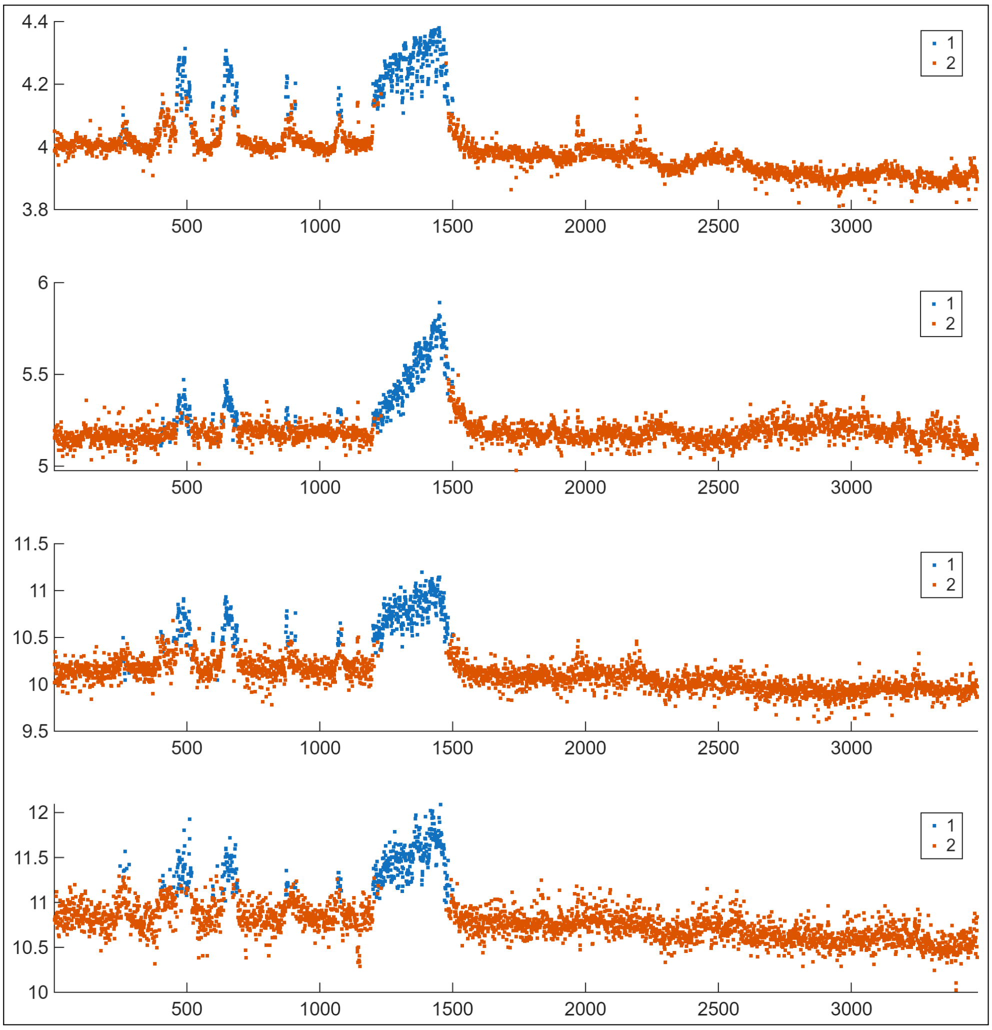


Figure 6.8: Hierarchical clustering for Z24.

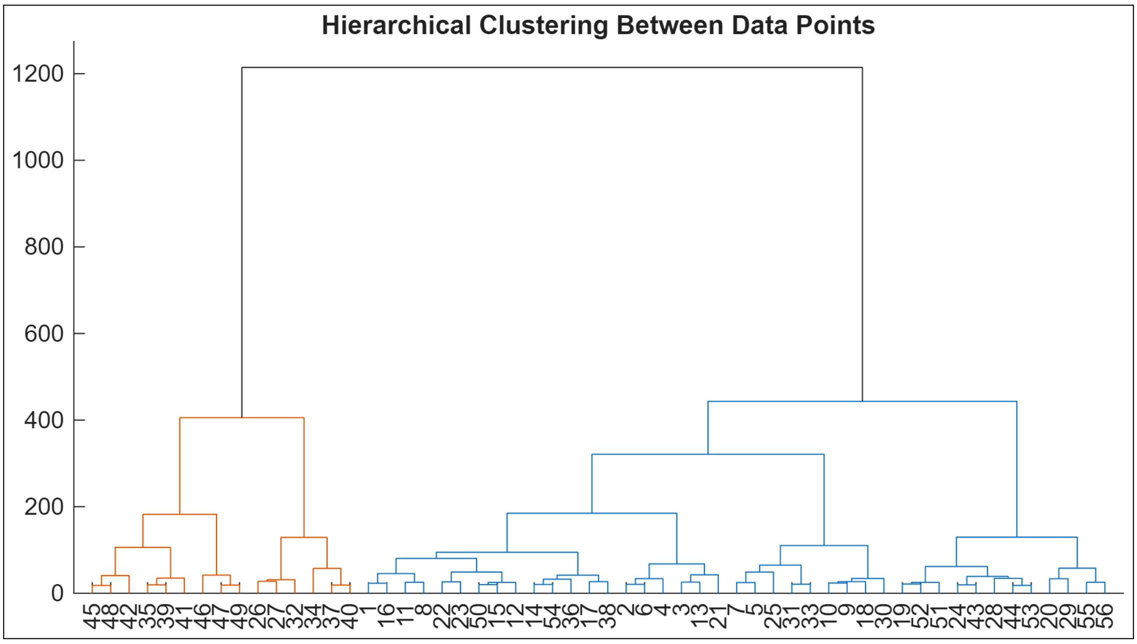


Figure 6.9: Dendrogram for Z24 hierarchical clustering.

6.3.3. Data Clustering for the Z24 Bridge Using DBSCAN

DBSCAN results in Figure 6.10 indicate that most samples belong to a dense baseline cluster, while a comparatively small subset is separated as either a secondary cluster or noise, concentrated around the most distinct deviations. This is consistent with the intended role of DBC in SHM, where rare events may appear as sparse samples relative to a dense healthy regime. Noise points represent sparse deviations separated from the dense baseline regime.

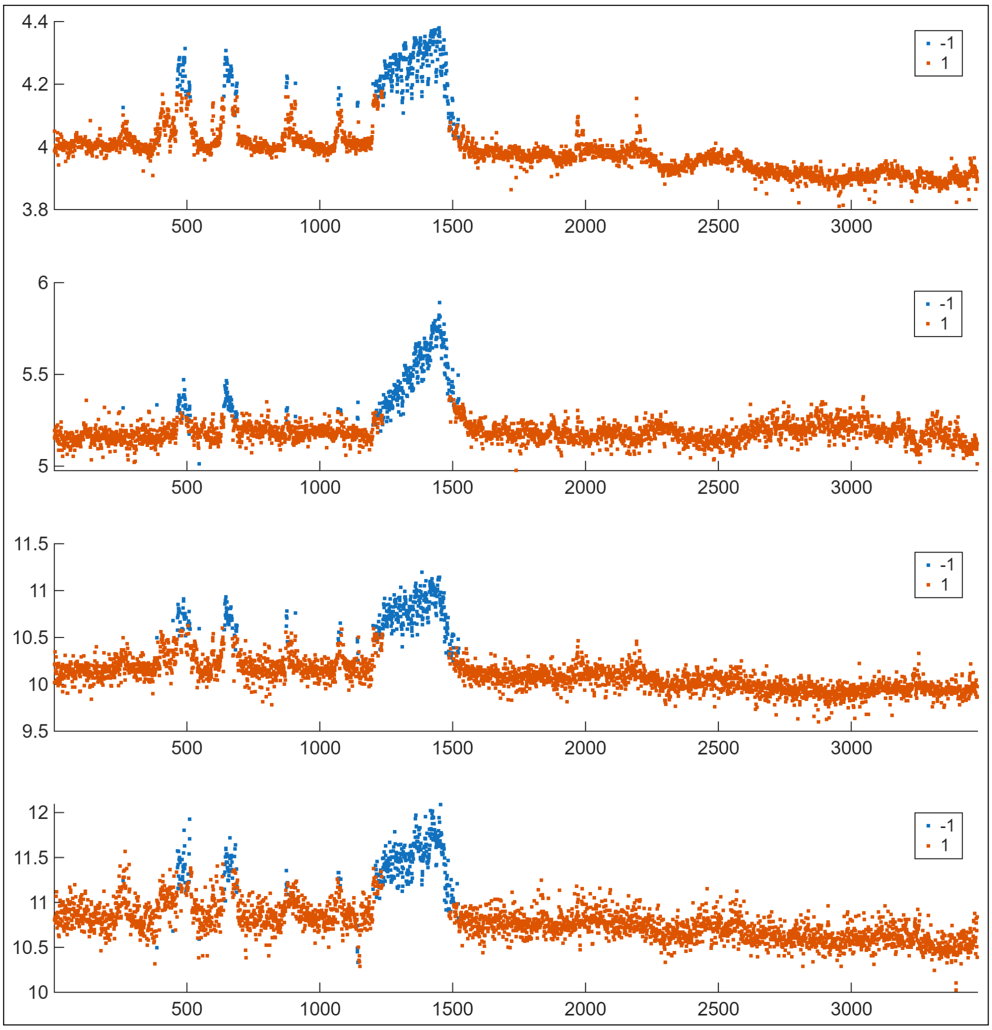


Figure 6.10: DBSCAN clustering for Z24.

6.3.4. Data Clustering for the Z24 Bridge Using GMMs

The GMMs result in Figure 6.11 shows a two-component partition of the data, separating the dominant baseline regime from a distinct group associated with stronger feature deviations. Because GMMs represent regimes probabilistically, they are well-suited to situations in which regime overlap is expected and strict hard boundaries could lead to unstable assignments near transitions.

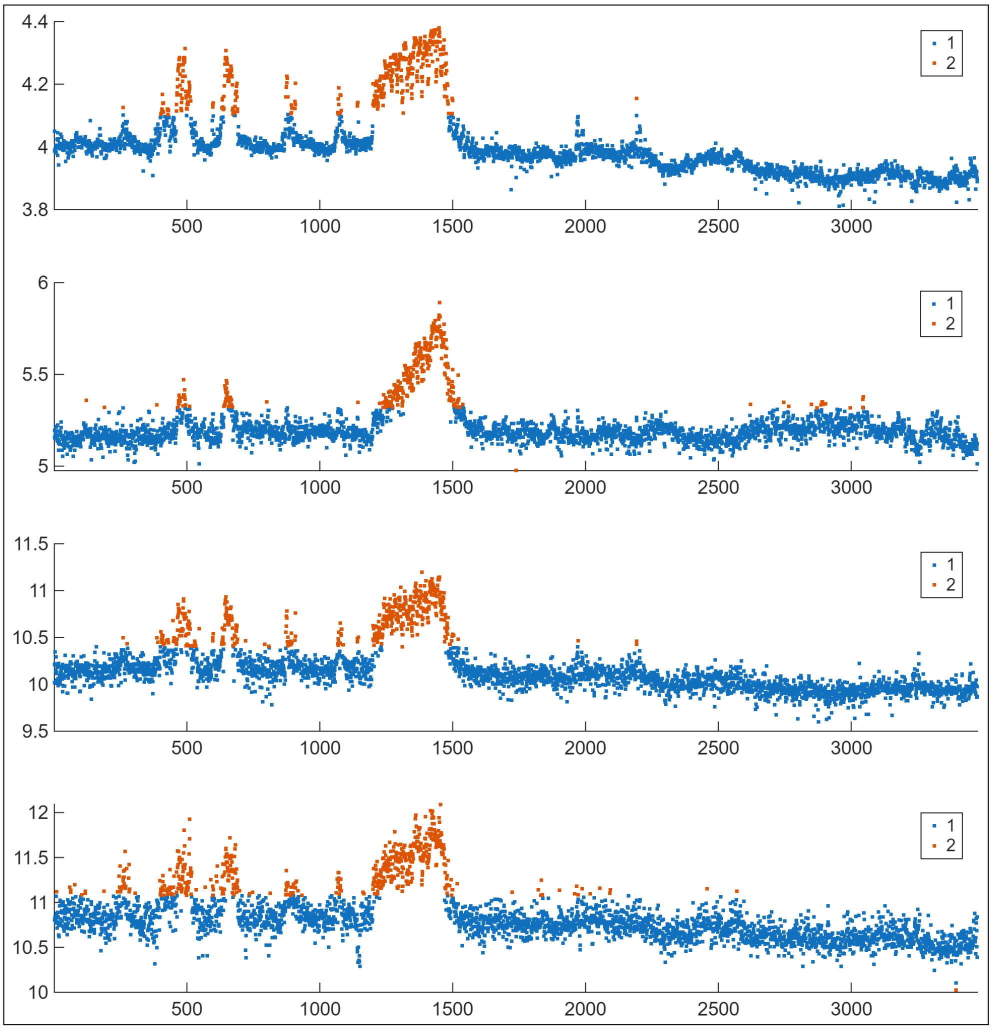


Figure 6.11: GMM regime modeling for Z24.

6.3.5. Data Clustering for the Z24 bridge Using SC

SC in Figure 6.12 provides a clear separation between regimes by exploiting global similarity structure and graph connectivity, which can better capture non-linear separations than centroid-based methods in challenging feature spaces [54]. The resulting regime assignment is visually consistent with the main deviations observed in the feature trajectories, supporting the use of SC as a regime-partitioning step prior to novelty detection. Regimes are separated in a graph-based embedding derived from sample similarities [53].

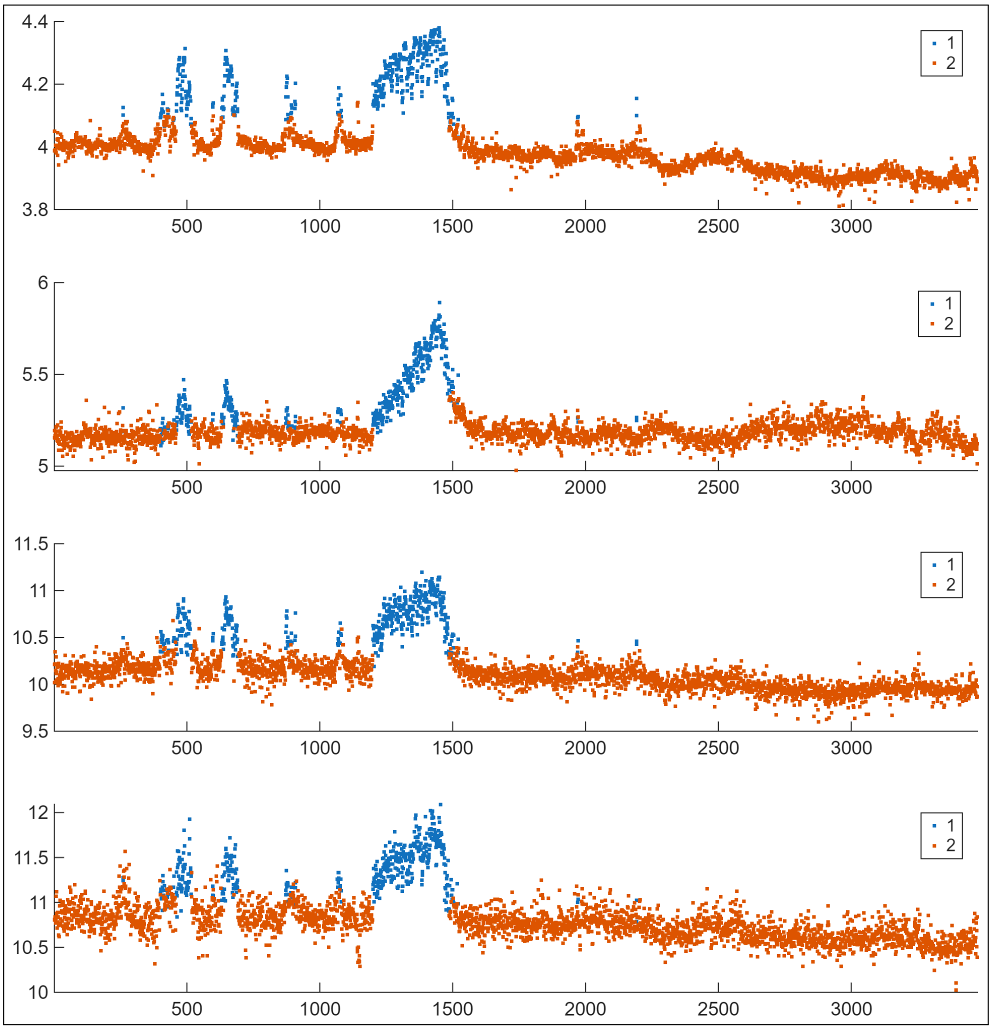


Figure 6.12: SC for Z24.

6.4. Anomaly Detection Using TSC–CAMSD

This section evaluates the damage alarming capability of the proposed TSC–CAMSD framework by analyzing the evolution of the novelty index over time and comparing it against a data-driven decision threshold. In the presented plots, the vertical dashed lines indicate the separation between the training portion and the subsequent undamaged (healthy) validation portion, while the horizontal dashed line represents the threshold boundary used to trigger an alarm. The threshold τ is computed from baseline novelty scores using an empirical quantile rule targeting a specified FPR α:

$$\tau = Q_{1-\alpha}\big(\{d_i\}_{\text{train}}\big).$$

Samples exceeding the threshold during the normal condition correspond to false alarms (false positives). Samples remaining below the threshold after the onset of an abnormal condition correspond to missed detections (false negatives).

6.4.1. Anomaly Detection Results for the KW51 Bridge

Figure 6.13 illustrates the TSC–CAMSD novelty index δ for KW51 (logarithmic scale). All samples in the training portion (black "+" markers) lie below the threshold, indicating that the baseline behavior is largely learned and preserved. The undamaged validation portion (blue "*" markers) exhibits a comparable level of δ, without any number of isolated exceedances that can be interpreted as false alarms should they exceed the threshold. At the onset of the post-event or damaged test portion (red pentagram markers), the novelty index increases abruptly by orders of magnitude and remains well above the threshold, yielding a clear separation between undamaged and damaged conditions and an essentially immediate detection after the change point. The two vertical dashed lines indicate the boundaries between training, undamaged validation, and damaged test segments. The horizontal dashed line denotes the threshold τ, computed as an empirical quantile of the training scores to achieve a target false-positive rate α. Exceedances during undamaged periods, which are not present here, correspond to false alarms, whereas any non-exceedances after the event onset would correspond to missed detections. In the following plots, the novelty index δ corresponds to the CAMSD score, i.e., $\delta := d_2$ as defined in the CAMSD formulation. The horizontal line indicates the decision threshold τ used to flag novelties.

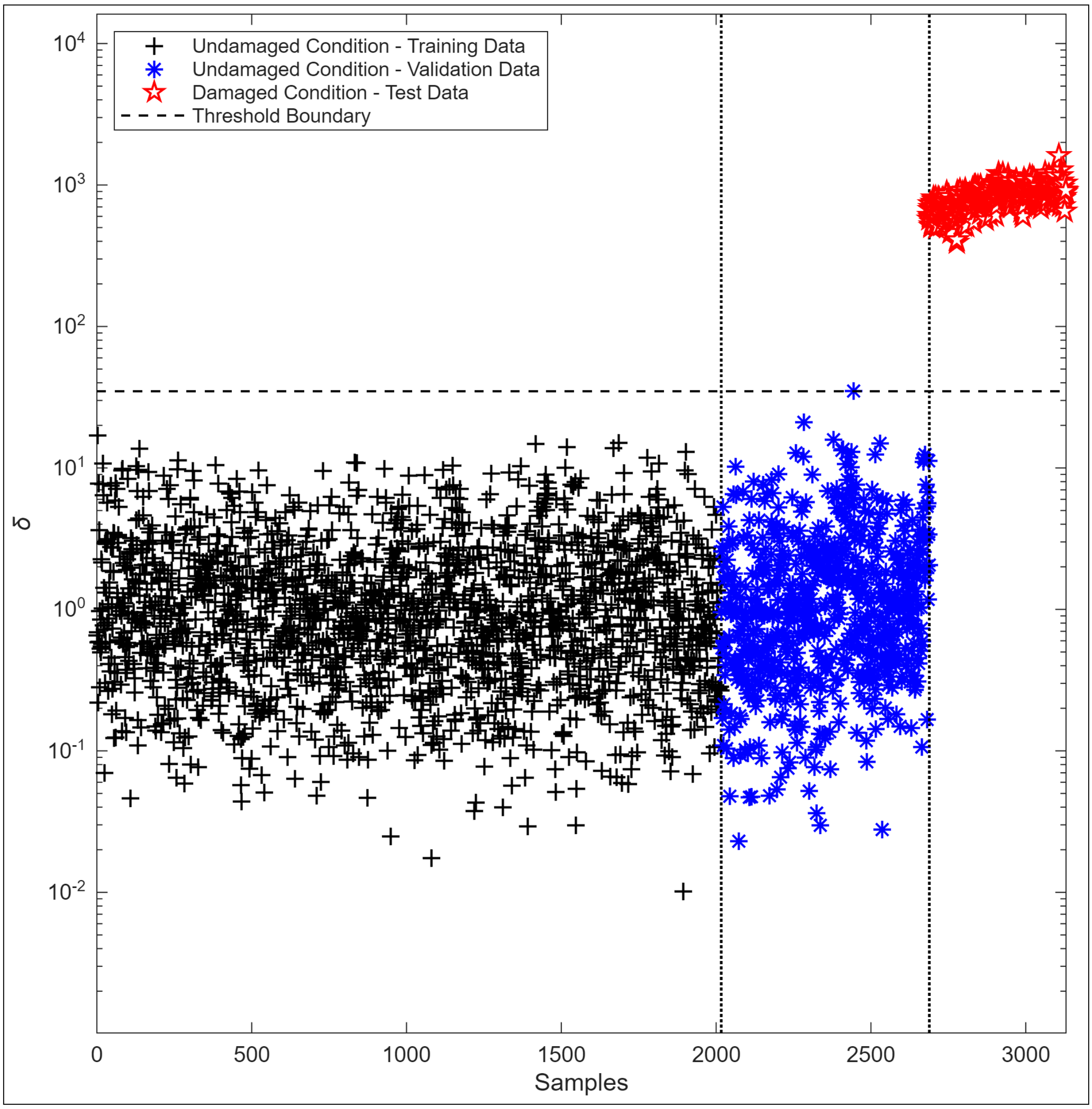


Figure 6.13: TSC–CAMSD novelty index for KW51.

6.4.2. Anomaly Detection Results for the Z24 Bridge

Figure 6.14 presents the novelty index for the Z24 case study. The index stays predominantly below the threshold throughout the training and undamaged validation segments, reflecting a consistent baseline regime despite normal variability. A limited number of exceedances are highlighted as false alarms, indicating that some rare baseline observations are sufficiently far from their local reference to cross the decision boundary. After the transition into the damaged test segment, the novelty index rises sharply and remains substantially above the threshold for an extended period, indicating a persistent deviation from the learned undamaged feature manifold. Only a small number of missed detections are visible around the onset of the abnormal behavior, implying that the method is most conservative precisely at the boundary where the system begins to deviate but has not yet fully separated from the baseline set. Overall, these observations support the suitability of thresholded CAMSD-based indices for early damage alarming, particularly when combined with clustering-based regime handling to mitigate the influence of EOV. The vertical dashed line separates training and undamaged validation; the horizontal dashed line is the threshold τ computed as an empirical quantile of training scores (target FPR α). Exceedances during undamaged periods are false alarms; non-exceedances after damage onset are missed detections. In the following plots, the novelty index δ corresponds to the CAMSD score, i.e., $\delta := d_2$. The horizontal line indicates the decision threshold τ (quantile-calibrated on undamaged data to achieve the target false-positive rate).

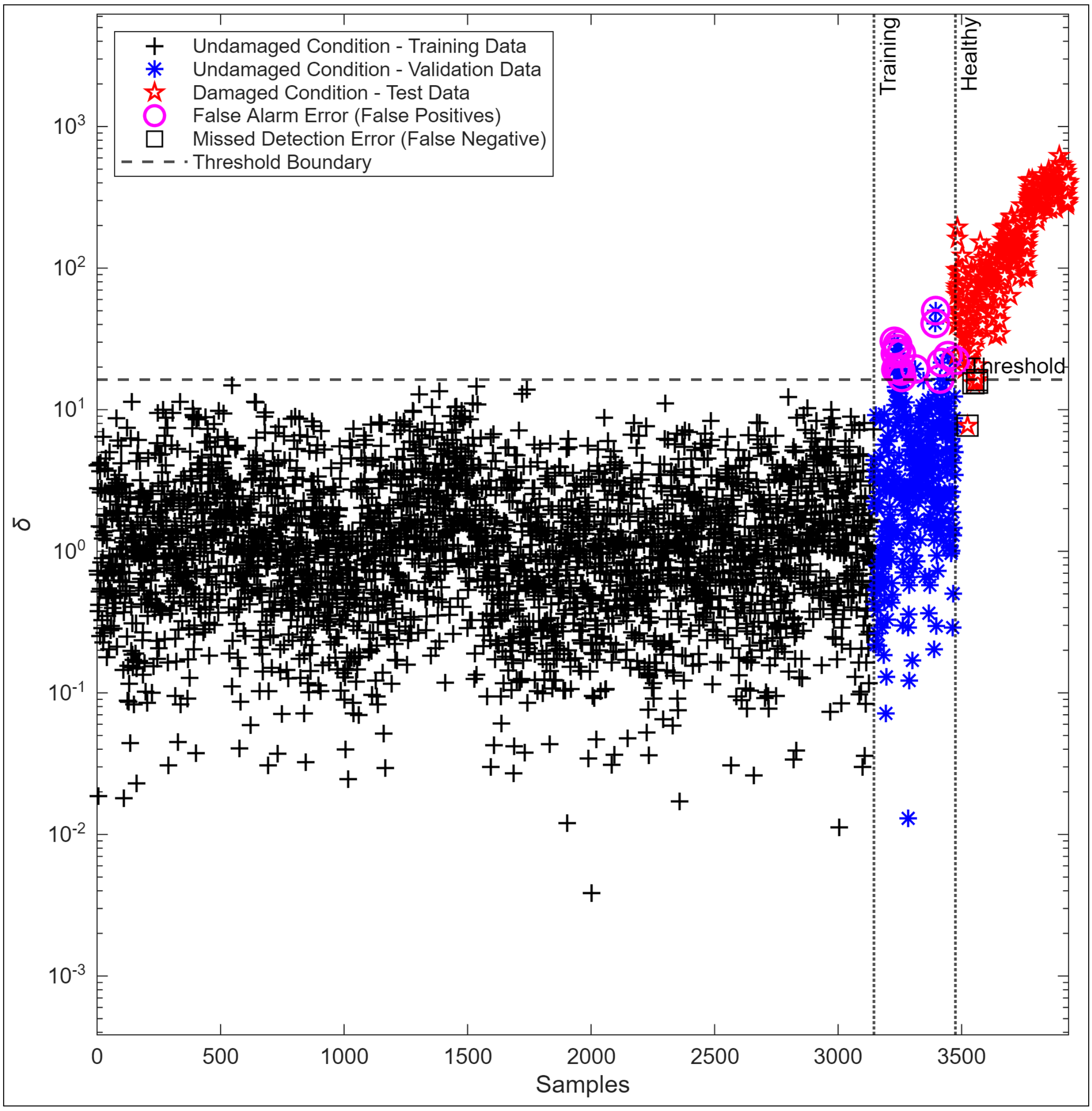


Figure 6.14: TSC–CAMSD novelty index for Z24.

6.5. Quantitative Performance Summary

To complement the qualitative plots, performance should be reported using: (i) FPR on undamaged data, (ii) FNR on damaged or post-event data when labels are available. For KW51, the event time corresponds to the documented structural intervention period. Meanwhile for Z24, damage timing is known. Tables 6.2 and 6.3 put in evidence such quantitative summary respectively.

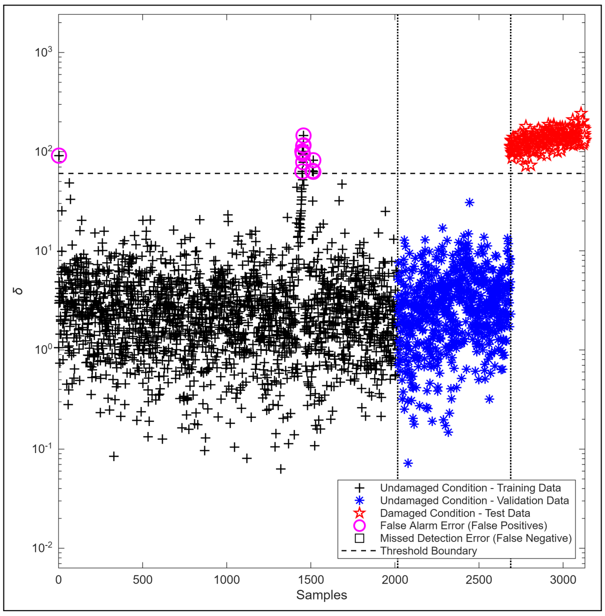
Method	**FPR (%)**	**FNR (%)**
Raw data (no clustering)	0.49	0.00
MSD (no clustering)	0.48	99.55
CAMSD (no clustering)	0.48	0.00
TSC–CAMSD	0.00	0.00

Table 6.2: Quantitative summary of novelty-detection performance for KW51 bridge.

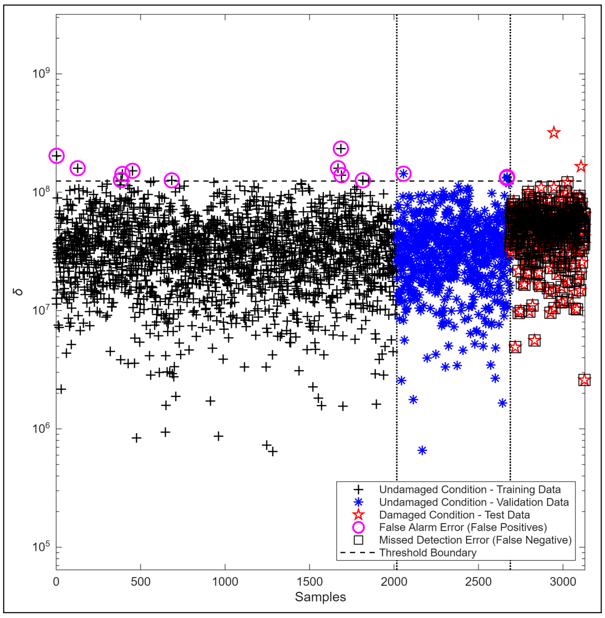
Method	**FPR (%)**	**FNR (%)**
Raw data (no clustering)	0.49	77.24
MSD (no clustering)	0.49	98.69
CAMSD (no clustering)	0.49	1.31
TSC–CAMSD	0.49	0.88

Table 6.3: Quantitative summary of novelty-detection performance for Z24 bridge.

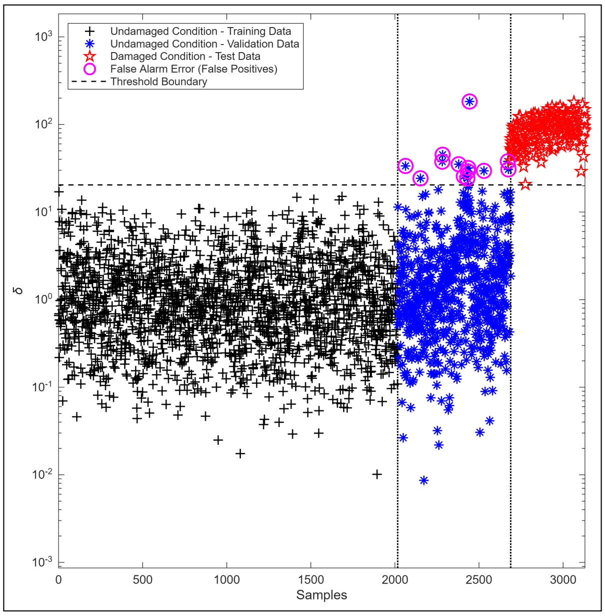
Key takeaways from the comparative results. Across both case studies (KW51 and Z24), the results indicate that combining regime awareness (TSC) with conditioning-guided local novelty scoring (CAMSD) yields a more stable separation between undamaged and damaged regimes under environmental and operational variability: clustering the baseline reduces heterogeneity and makes the reference more appropriate, while adaptive neighborhood selection and conditioning checks mitigate numerical instabilities in local covariance estimation; finally, quantile-based calibration of the decision threshold provides direct control on false alarms, improving interpretability.



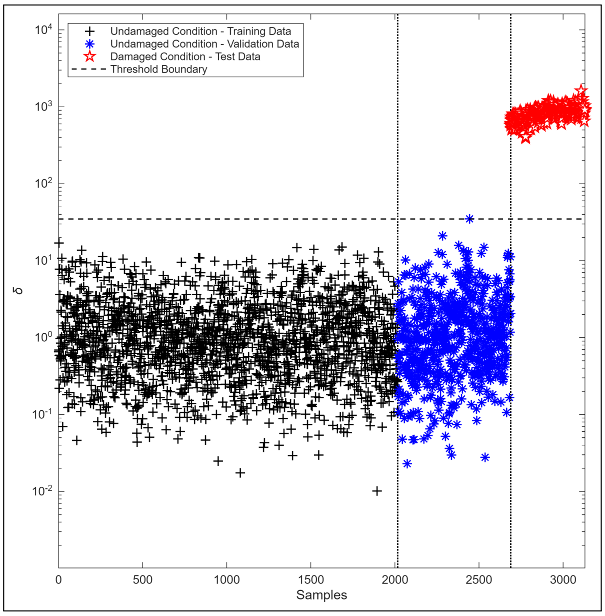
(a) Raw data (no clustering)



(b) MSD (no clustering)

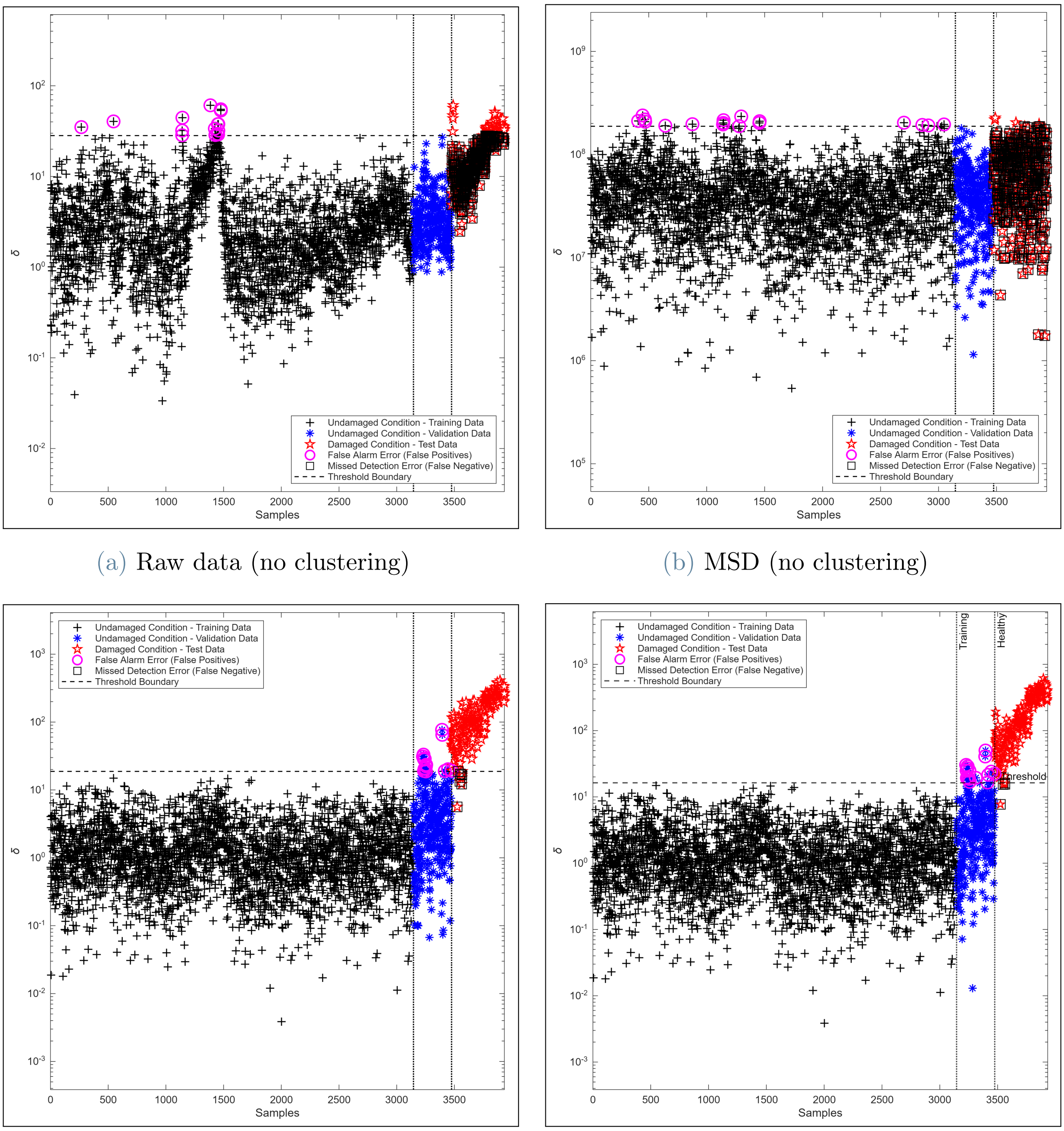


(c) CAMSD (no clustering)



(d) TSC–CAMSD

Figure 6.15: Visual summary of novelty-detection performance for KW51 bridge.



(a) Raw data (no clustering)

(b) MSD (no clustering)

(c) CAMSD (no clustering)

(d) TSC–CAMSD

Figure 6.16: Visual summary of novelty-detection performance for Z24 bridge.

7 | Conclusions and Future Developments

7.1. Overview and Research Objective

The primary objective of this research was to develop an unsupervised, non-parametric method for early damage warning in SHM using short term vibration based features, where labeled damage data are typically unavailable. A central challenge in this context is that EOV can alter modal features in ways that may mask or mimic damage, which can reduce the reliability of novelty-based alarming if not treated explicitly.

To address this issue, this thesis proposed the TSC–CAMSD framework, which integrates (i) a TSC and (ii) a CAMSD, followed by threshold-based decision-making for alarm triggering. The combination was designed to produce a stable novelty index under undamaged conditions while remaining sensitive to persistent deviations that are not explainable by baseline variability.

7.2. Main Contributions and Findings

The first contribution of this work is a regime-aware baseline modeling strategy forshort term SHM. By introducing clustering-based partitioning of the baseline data, the proposed approach reduces the risk of comparing new observations against an inappropriate global reference, which is a common mechanism leading to false alarms under an EOV driven regime shifts. This is consistent with the broader understanding that environmental and operational actions can produce modal frequency variability comparable to (or larger than) changes caused by small to moderate damage, motivating regime handling and normalization prior to alarming.

The second contribution is the use of CAMSD as a robust novelty index suitable for limited and heterogeneous baseline sets. CAMSD combines local neighborhood modeling with numerical conditioning checks and regularization safeguards, improving stability when local covariance estimation becomes ill-conditioned. This is particularly relevant in real monitoring data, where missing samples, non-stationary patterns, and regime-dependent scatter can degrade classical distance measures.

The effectiveness of the proposed approach was evaluated on two bridge case studies (KW51 and Z24) using short term vibration features. The novelty index plots show that, during the undamaged portions, the majority of samples remain below the decision threshold, indicating that the baseline behavior is largely learned and maintained. Exceedances during the undamaged portion are limited and appear as isolated events, which aligns with the expected behavior of unsupervised thresholding that controls (but cannot fully eliminate) false alarms under natural variability. After the transition to the damaged test portion, the novelty index rises and remains sustained above the threshold for an extended period, providing a clear separation between undamaged and damaged behavior and supporting early warning decisions. The most challenging region is observed around the onset of abnormal behavior, where a small number of missed detections may occur due to overlap between emerging damage effects and baseline scatter near the decision boundary.

A further observation from these results is that EOV should not be viewed solely as a long term monitoring issue. Even within short term datasets, EOV can introduce noticeable variability in vibration based features, reinforcing the need for regime handling and robust novelty scoring for reliable damage alarming.

7.3. Limitations

Despite the promising performance of TSC–CAMSD, several limitations remain. First, the method relies on the of the available baseline data, and short term monitoring windows may not capture the full range of operational and environmental regimes that the structure can experience. This can reduce generalization when future measurements occur under previously unseen conditions, potentially increasing false alarms or reducing sensitivity. Second, threshold based alarming introduces an inherent trade off between false alarms and missed detections. While the observed results indicate stable behavior under undamaged conditions and strong separation under damaged conditions, the boundary region at the onset of damage remains the most difficult to classify, particularly when the novelty index transitions gradually. Finally, the regime partitioning stage depends on clustering quality; if the operating regimes are weakly separable in the chosen features, regime assignment may become ambiguous and partially reduce the benefits of local referencing.

7.4. Future Work: Decomposition Driven Deep Features under EOV

Future developments can extend the proposed EOV mitigation SHM framework by introducing a signal decomposition front end to generate cleaner, more EOV robust inputs for neural network models, while preserving the current clustering stage and the CAMSD anomaly scoring logic.

- EMD family IMFs: Apply EMD, EEMD, CEEMDAN, or ICEEMDAN to decompose each measured response into intrinsic mode functions (IMFs), then extract per IMF features (e.g., energy, instantaneous frequency bands, entropy, and basic statistics). These features can be used as inputs to the existing clustering step and subsequent CAMSD based novelty detection.
- VMD modes: Use VMD to obtain a fixed number of modes with controlled bandwidths, enabling consistent neural network inputs (same number of channels per sample) across time and varying operating conditions. This consistency is particularly useful for robust learning and inference under EOV.

From a methodological perspective, once a network is trained, the key step is forward simulation (inference): the learned mapping is evaluated on new decomposed feature vectors (IMFs/modes) to produce outputs such as class probabilities, scores, or embeddings that can be monitored for novelty. This forward evaluation stage can be implemented with MATLAB style network simulation functions (e.g., `sim`) or equivalent inference routines in other deep learning frameworks, without changing the underlying theory.

Overall, this thesis demonstrated that combining regime aware partitioning with a robust, adaptive distance based novelty index provides a practical and effective foundation for early warning in short term SHM, while also highlighting the need for further work on generalization, thresholding, and deployment under broader operating conditions.

Bibliography

[1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974. doi: 10.1109/TAC.1974.1100705.

[2] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001. doi: 10.1023/A:1010933404324.

[3] G. Capellari, E. Chatzi, and S. Mariani. Structural health monitoring sensor network optimization through bayesian experimental design. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 4(2):04018016, 2018.

[4] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20:273–297, 1995. doi: 10.1007/BF00994018.

[5] M. H. Daneshvar and H. Sarmadi. Unsupervised learning-based damage assessment of full-scale civil structures under long-term and short-term monitoring. *Engineering Structures*, 256:114059, 2022.

[6] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 1977. doi: 10.1111/j.2517-6161.1977.tb01600.x.

[7] A. Deraemaeker and K. Worden. A comparison of linear approaches to filter out environmental effects in structural health monitoring. *Mechanical Systems and Signal Processing*, 105:1–15, 2018.

[8] S. W. Doebling, C. R. Farrar, M. B. Prime, and D. W. Shevitz. Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review. Technical report, 1996.

[9] A. Entezami, H. Shariatmadar, and A. Karamodin. Data-driven damage diagnosis under environmental and operational variability by novel statistical pattern recognition methods. *Structural Health Monitoring*, 18(5-6):1416–1443, 2019.

[10] A. Entezami, H. Sarmadi, B. Behkamal, and S. Mariani. Big data analytics and struc-

tural health monitoring: a statistical pattern recognition-based approach. *Sensors*, 20(8):2328, 2020.

[11] A. Entezami, H. Shariatmadar, and S. Mariani. Early damage assessment in large-scale structures by innovative statistical pattern recognition methods based on time series modeling and novelty detection. *Advances in Engineering Software*, 150:102923, 2020.

[12] A. Entezami, H. Shariatmadar, and S. Mariani. Fast unsupervised learning methods for structural health monitoring with large vibration data from dense sensor networks. *Structural Health Monitoring*, 19(6):1685–1710, 2020.

[13] A. Entezami, H. Shariatmadar, and C. De Michele. Non-parametric empirical machine learning for short-term and long-term structural health monitoring. *Structural Health Monitoring*, 21(6):2700–2718, 2022.

[14] A. Entezami, H. Sarmadi, and B. Behkamal. Long-term health monitoring of concrete and steel bridges under large and missing data by unsupervised meta learning. *Engineering Structures*, 279:115616, 2023.

[15] A. Entezami, B. Behkamal, and C. De Michele. *Long-Term Structural Health Monitoring by Remote Sensing and Advanced Machine Learning: A Practical Strategy Via Structural Displacements from Synthetic Aperture Radar Images*. Springer Nature, 2024.

[16] A. Entezami, H. Sarmadi, and B. Behkamal. Short-term damage alarming with limited vibration data in bridge structures: A fully non-parametric machine learning technique. *Measurement*, 235:114935, 2024.

[17] A. Entezami, H. Sarmadi, B. Behkamal, and C. De Michele. On continuous health monitoring of bridges under serious environmental variability by an innovative multi-task unsupervised learning method. *Structure and Infrastructure Engineering*, 20(12):1975–1993, 2024.

[18] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96)*, pages 226–231. AAAI Press, 1996. URL `https://dl.acm.org/doi/10.5555/3001460.3001507`.

[19] C. R. Farrar, D. W. Allen, G. Park, S. Ball, and M. P. Masquelier. Coupling sensing

hardware with data interrogation software for structural health monitoring. *Shock and Vibration*, 13(4-5):519–530, 2006.

[20] S. D. Fassois and J. S. Sakellariou. Statistical time series methods for structural health monitoring. In *Encyclopedia of Structural Health Monitoring*, pages 443–472. 2009.

[21] J. A. Hartigan and M. A. Wong. Algorithm AS 136: A K-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(1):100–108, 1979. doi: 10.2307/2346830.

[22] H. Jiang. *Machine Learning Fundamentals: A Concise Introduction*. Cambridge University Press, 2021.

[23] KU Leuven, Department of Civil Engineering, Structural Mechanics Section. Z24 bridge benchmark. URL `https://bwk.kuleuven.be/bwm/z24#references`. Accessed: 2026-02-12.

[24] KU Leuven, Structural Mechanics Section. Kw51. URL `https://bwk.kuleuven.be/bwm/kw51`. Accessed: 2026-02-12.

[25] S. P. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982. doi: 10.1109/TIT.1982.1056489.

[26] J. P. Lynch, H. Sohn, and M. L. Wang, editors. *Sensor Technologies for Civil Infrastructures: Volume 1: Sensing Hardware and Data Collection Methods for Performance Assessment*. Woodhead Publishing, 2022.

[27] J. B. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297. University of California Press, 1967.

[28] K. Maes and G. Lombaert. Monitoring data for railway bridge kw51 in leuven, belgium, before, during, and after retrofitting (version 1.0), 2020. URL `https://doi.org/10.5281/zenodo.3745914`.

[29] K. Maes and G. Lombaert. Monitoring railway bridge kw51 before, during, and after retrofitting. *Journal of Bridge Engineering*, 26(3):1–19, 2021. doi: 10.1061/(ASCE)BE.1943-5592.0001668. URL `https://doi.org/10.1061/(ASCE)BE.1943-5592.0001668`.

[30] C. Malzer and M. Baum. A hybrid approach to hierarchical density-based cluster selection. *arXiv preprint arXiv:1911.02282*, 2019.

[31] L. McInnes and J. Healy. Accelerated hierarchical density clustering. *arXiv preprint arXiv:1705.07321*, 2017.

[32] K. P. Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.

[33] K. P. Murphy. *Probabilistic Machine Learning: An Introduction*. The MIT Press, Cambridge, MA, 2022. ISBN 9780262046824. URL `http://probml.github.io/book1`.

[34] K. P. Murphy. *Probabilistic Machine Learning: Advanced Topics*. The MIT Press, Cambridge, MA, 2023. ISBN 9780262048439. URL `https://probml.github.io/book2`.

[35] F. Murtagh and P. Legendre. Ward's hierarchical agglomerative clustering method: Which algorithms implement ward's criterion? *Journal of Classification*, 31(3):274–295, 2014. doi: 10.1007/s00357-014-9161-z.

[36] A. C. Neves. *Structural health monitoring of bridges: data-based damage detection method using machine learning*. PhD thesis, KTH Royal Institute of Technology, 2020.

[37] A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2002.

[38] B. Peeters and G. De Roeck. One-year monitoring of the z24-bridge: Environmental effects versus damage events. *Earthquake Engineering & Structural Dynamics*, 30(2):149–171, 2001. doi: 10.1002/1096-9845(200102)30:2<149::AID-EQE1>3.0.CO;2-Z.

[39] C. A. Perez-Ramirez, J. P. Amezquita-Sanchez, H. Adeli, M. Valtierra-Rodriguez, R. d. J. Romero-Troncoso, A. Dominguez-Gonzalez, and R. A. Osornio-Rios. Time-frequency techniques for modal parameters identification of civil structures from acquired dynamic signals. *Journal of Vibroengineering*, 18(5):3164–3185, 2016.

[40] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, MA, 2006. ISBN 9780262182539. URL `https://gaussianprocess.org/gpml/`.

[41] C. K. Reddy, editor. *Data Clustering: Algorithms and Applications*. Chapman and Hall/CRC, 2018.

[42] H. Sarmadi and A. Entezami. Application of supervised learning to validation of damage detection. *Archive of Applied Mechanics*, 91(1):393–410, 2021.

[43] H. Sarmadi and A. Karamodin. A novel anomaly detection method based on adaptive

mahalanobis-squared distance and one-class knn rule for structural health monitoring under environmental effects. *Mechanical Systems and Signal Processing*, 140:106495, 2020.

[44] H. Sarmadi, A. Entezami, B. S. Razavi, and K.-V. Yuen. Ensemble learning-based structural health monitoring by mahalanobis distance metrics. *Structural Control and Health Monitoring*, 28(2):e2663, 2021.

[45] H. Sarmadi, A. Entezami, M. Salar, and C. De Michele. Bridge health monitoring in environmental variability by new clustering and threshold estimation methods. *Journal of Civil Structural Health Monitoring*, 11(3):629–644, 2021.

[46] H. Sarmadi, A. Entezami, and F. Magalhães. Unsupervised data normalization for continuous dynamic monitoring by an innovative hybrid feature weighting-selection algorithm and natural nearest neighbor searching. *Structural Health Monitoring*, 22 (6):4005–4026, 2023.

[47] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2): 461–464, 1978. doi: 10.1214/aos/1176344136.

[48] scikit-learn developers. Clustering — scikit-learn documentation, 2025. URL `https://scikit-learn.org/stable/modules/clustering.html`. Accessed: 2026-02-12.

[49] S. M. Shapourian. Unsupervised damage detection framework based on spectral clustering and density peaks for early warning systems. M.sc. thesis, Politecnico di Milano, School of Civil, Environmental and Land Management Engineering (Polo territoriale di Lecco), 2024. M.Sc. programme in Civil Engineering for Risk Mitigation; Advisor: Prof. Alireza Entezami; Academic Year: 2023–2024.

[50] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000. doi: 10.1109/34.868688.

[51] M. D. Todd, G. A. Johnson, and B. L. Althouse. A novel bragg grating sensor interrogation system utilizing a scanning filter, a mach-zehnder interferometer and a 3×3 coupler. *Measurement Science and Technology*, 12(7):771, 2001.

[52] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4): 395–416, 2007. doi: 10.1007/s11222-007-9033-z.

[53] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4): 395–416, 2007. doi: 10.1007/s11222-007-9033-z.

[54] U. Von Luxburg, M. Belkin, and O. Bousquet. Consistency of spectral clustering. *The Annals of Statistics*, pages 555–586, 2008.

[55] J. H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963. doi: 10.1080/01621459.1963.10500845.

A | Appendix A

The related MATLAB code to the Z24 Bridge focuses on implementing an unsupervised learning algorithm for Tuned Spectral Clustering combined with Conditioning-guided Adaptive Mahalanobis Squared Distance aimed at analyzing vibration data to detect anomalies and conduct SHM.

```

%% Z24 Bridge
clc; clear; close all; format long; warning off; rng default;
global nTr nHe nTe nFr nMode alpha

%% Load Datasets
load('Freq_Z24N.mat'); clear X Z i j nHe nTe nTo nTr nm;

[~,nFr] = size(fnew); % Total Number of Frequency Data = 3932
[nMode,~] = size(fnew); % Number of Frequency Modes = 4
nTr = floor(0.80 * nFr); % Number of Training Data = 3145
nHe = 3475; % Number of Healthy Data = 3475
nVd = nHe - nTr; % Number of Validation Data = 330
X = fnew(:,1:nTr); % Training Data = 4x3145
Z = fnew(:,nTr+1:nFr); % Testing Data = 4x787
[~,nTe] = size(Z); % Number of Testing Data = 787
alpha = 0.5; % 0.5% Target False Positive Rate (FPR) on Healthy

%% Plot Original Signals
ylabel = ["  Mode '(a)'  ","  Mode '(b)'  ","  Mode '(c)'  ","  Mode '(d)'  "];
figure ('Name', 'All modes', 'NumberTitle', 'off');
stackedplot(fnew', 'Color', 'b',"Title"," Mode '(a)','(b)','(c)','(d)' ", ...
        "DisplayLabels",ylabel);
xlim([1 nFr]); xlabel('Measurements'); clear ylabel;

Title = {'(a)','(b)','(c)','(d)'};
figure ('Name', 'Z24', 'NumberTitle', 'off')
```

```
for i = 1:nMode
    subplot(4,1,i)
    if i == 1; area([375 375;1550 1550],[min(fnew(1,:))-0.05 ...
         max(fnew(1,:))+0.05;min(fnew(1,:))-0.05 max(fnew(1,:))+0.05]); end
    if i == 2; area([375 375;1550 1550],[min(fnew(2,:))-0.05 ...
         max(fnew(2,:))+0.05;min(fnew(2,:))-0.05 max(fnew(2,:))+0.05]); end
    if i == 3; area([375 375;1550 1550],[min(fnew(3,:))-0.05 ...
         max(fnew(3,:))+0.05;min(fnew(3,:))-0.05 max(fnew(3,:))+0.05]); end
    if i == 4; area([375 375;1550 1550],[min(fnew(4,:))-0.05 ...
         max(fnew(4,:))+0.05;min(fnew(4,:))-0.05 max(fnew(4,:))+0.05]); end

    hold on
    f2=plot(1:nHe,fnew(i,1:nHe),'LineStyle','none','Marker','.', ...
         'MarkerSize',4,'Color','b');
    ax = gca; ax.FontName = 'Arial'; ax.FontSize = 10;

    hold on
    f4=plot(nHe+1:nFr,fnew(i,nHe+1:nFr),'LineStyle','none','Marker','.', ...
         'MarkerSize',4,'Color','r');
    ax = gca; ax.FontName = 'Arial'; ax.FontSize = 10;

    hold on
    line([nHe nHe],[min(min(fnew))-0.05 max(max(fnew))+0.05], ...
         'LineStyle',':','LineWidth',1,'Color','k');
    xlabel('Samples'); ylabel('Frequency (Hz)');
    set(gca,'XLim',[1 nFr]); title(Title{i});
    if i == 1; set(gca,'YLim',[min(fnew(1,:))-0.05 max(fnew(1,:))+0.05]); ...
              legend([f2,f4],'Before damage','After damage', ...
              'Location','northeast'); end
    if i == 2; set(gca,'YLim',[min(fnew(2,:))-0.05 max(fnew(2,:))+0.05]); ...
              legend([f2,f4],'Before damage','After damage', ...
              'Location','northeast'); end
    if i == 3; set(gca,'YLim',[min(fnew(3,:))-0.05 max(fnew(3,:))+0.05]); ...
              legend([f2,f4],'Before damage','After damage', ...
              'Location','northeast'); end
    if i == 4; set(gca,'YLim',[min(fnew(4,:))-0.05 max(fnew(4,:))+0.05]); ...
              legend([f2,f4],'Before damage','After damage', ...
              'Location','northeast'); end
end
```

```matlab

%% Data Normalization + Plot
fnew_Norm = normalize(fnew','zscore',"std");
figure ('Name', 'Normalized Modal Frequencies', 'NumberTitle', 'off');
plot(fnew_Norm , 'LineStyle' , 'none' , 'Marker' , '.' , 'MarkerSize' , 5);
xlim([1 nFr]); xlabel('Measurements');
hold on; line([nTr nTr],[min(min(fnew_Norm)) max(max(fnew_Norm))], ...
        'LineStyle','--','LineWidth',1.2,'Color','k');
hold on; line([nHe nHe],[min(min(fnew_Norm)) max(max(fnew_Norm))], ...
        'LineStyle',':','LineWidth',1.2,'Color','k');
legend("Mode (a)","Mode (b)","Mode (c)","Mode (d)","Training Data", ...
        "Undamaged Condition");

%% Spectral Clustering (tuning) + PCA
X_std = zscore(X.');
D2 = pdist2(X_std, X_std, 'euclidean').^2;
D2(1:size(D2,1)+1:end) = 0;

kList = 2:6;

d = sqrt(D2(:));
d = d(d > 0);
sigma0 = median(d);

sigmaList = sigma0 * [0.25 0.5 1 2 4];

best.score = -Inf;
best.k = NaN;
best.sigma = NaN;
best.idx = [];
best.V = [];
best.eigs = [];

rng default

for k = kList
    for sigma = sigmaList

        S = exp(-D2./(2*sigma^2));
```

```
        S(1:size(S,1)+1:end) = 0;
        S = (S + S.')/2;

        [~, V, D] = spectralcluster(S, k, ...
            'Distance','precomputed', ...
            'LaplacianNormalization','symmetric');

        idx = kmeans(V, k, 'Replicates', 20, 'MaxIter', 2000);

        s = silhouette(V, idx);
        score = mean(s);

        if score > best.score
            best.score = score;
            best.k = k;
            best.sigma = sigma;
            best.idx = idx;
            best.V = V;
            best.eigs = D;
        end
    end
end

fprintf('Best SC params: k=%d, sigma=%.4g, mean silhouette=%.4f\n', ...
            best.k, best.sigma, best.score);

idx = best.idx;

[~, scorePCA] = pca(X_std);

figure('Name', 'SC Results (PCA) - UPDATED', 'NumberTitle', 'off');
gscatter(scorePCA(:,1), scorePCA(:,2), idx);
xlabel('PC1'); ylabel('PC2');
title(sprintf('Spectral Clustering (best k=%d, sigma=%.3g)', ...
                best.k, best.sigma));

figure('Name','Laplacian smallest eigenvalues (best)','NumberTitle','off');
stem(best.eigs, 'filled');
xlabel('Index'); ylabel('Eigenvalue');
```

```
title('k smallest Laplacian eigenvalues (best setting)');

%% Final Cluster
nCluster = max(idx);
Cluster = getcluster(X,idx,max(idx));

for i = 1:nCluster
    [~,nSamCluster(i,1)] = size(Cluster{i});
end

minSamCluster = min(nSamCluster);

figure ('Name', 'Final Cluster', 'NumberTitle', 'off')
for i = 1:nTr
    if idx(i) == 1
        h1=plot(i,idx(i),'LineStyle','none','Marker','.','MarkerSize',8, ...
            'Color','b'); hold on
    else
        h2=plot(i,idx(i),'LineStyle','none','Marker','x','MarkerSize',6, ...
            'Color','m'); hold on
    end
end
xlabel('Training data'); ylabel('Cluster indices'); title('(a)')
legend([h1,h2],'Cluster 1','Cluster 2')
set(gca,'XLim',[0 nTr+1],'YLim',[0.8 nCluster+.2]);
ax = gca; ax.FontName = 'Arial'; ax.FontSize = 14;

%% CAMSD - Training - Input Finder
fprintf('\n- CAMSD - Training - Input Finder \n');

testPts = X';
refPts = X';

y = zeros(nTr,1,'uint8');    % 0=healthy

p = size(refPts,2);
kMinList = unique([max(p+1,20), 25:5:60]);   % 20-60
kMaxList = 60:5:160;                         % 60-160

```

```
best = struct('FNR',inf,'FPR',inf,'kMin',NaN,'kMax',NaN,'thr',NaN);

totalIters = numel(kMinList) * numel(kMaxList);
iter = 0;

h = waitbar(0, 'Grid search: 0%', ...
    'Name','CAMSD Grid Search - Training', ...
    'CreateCancelBtn','setappdata(gcbf,''canceling'',1)');
setappdata(h,'canceling',0);

updateEvery = 1;

for kMin = kMinList
    for kMax = kMaxList
        iter = iter + 1;

        if ~ishandle(h) || getappdata(h,'canceling')
            if ishandle(h), delete(h); end
            error('Canceled by user.');
        end

        opts = struct('kMin',kMin,'kMax',kMax,'doZScore',true);
        [d1,~] = CAMSD(testPts, refPts, opts);

        % threshold from healthy to control FPR
        thr = prctile(d1(y==0), 100-alpha);

        yhat = d1 > thr;

        FPR = mean(yhat(y==0)==1);

        if FPR < best.FPR && FPR <= 2*alpha/100
            best = struct('FPR',FPR,'kMin',kMin,'kMax',kMax,'thr',thr);
        end

        if mod(iter, updateEvery) == 0 || iter == totalIters
            waitbar(iter/totalIters, h, ...
                sprintf('Grid search: %d%% (%d/%d)', ...
                    round(100*iter/totalIters), iter, totalIters));
```

```
            drawnow limitrate;
        end
    end
end

if ishandle(h), delete(h); end

disp(best)

%% CAMSD - Training
fprintf('\n- CAMSD - Training \n');

testPts = X';
refPts = X';

opts = struct();
opts.kMin = best.kMin;
opts.kMax = best.kMax;
opts.doZScore = true;
opts.lambdaInit = 0;          % start unregularized
opts.lambdaAuto = true;       % auto-regularize if needed
opts.minRcond = 1e-5;

[d1, info1] = CAMSD(testPts, refPts, opts);

thr = prctile(d1, 100 * (1 - alpha/100));  % Threshold Boundary

figure('Name', 'CAMSD - d1', 'NumberTitle', 'off');
semilogy(d1); ylim([min(d1) max(d1)]);
yline(thr,'k--','Threshold');

cp = findchangepts(d1, 'Statistic','mean', 'MaxNumChanges', 2);
disp(cp);

%% CAMSD - Testing - Input Finder
fprintf('\n- CAMSD - Testing - Input Finder \n');

testPts = Z';
refPts = Cluster{1,1}';
```

```

y = zeros(nTe,1,'uint8');    % 0=healthy
y(nVd+1:end) = 1;    % 1=damaged

p = size(refPts,2);
kMinList = unique([max(p+1,20), 25:5:60]);    % 20-60
kMaxList = 60:5:160;                           % 60-160

best = struct('FNR',inf,'FPR',inf,'kMin',NaN,'kMax',NaN,'thr',NaN);

totalIters = numel(kMinList) * numel(kMaxList);
iter = 0;

h = waitbar(0, 'Grid search: 0%', ...
    'Name','CAMSD Grid Search - Testing', ...
    'CreateCancelBtn','setappdata(gcbf,''canceling'',1)');
setappdata(h,'canceling',0);

updateEvery = 1;

for kMin = kMinList
    for kMax = kMaxList
        iter = iter + 1;

        if ~ishandle(h) || getappdata(h,'canceling')
            if ishandle(h), delete(h); end
            error('Canceled by user.');
        end

        opts = struct('kMin',kMin,'kMax',kMax,'doZScore',true);
        [d2,~] = CAMSD(testPts, refPts, opts);

        thr  = prctile(d2(y==0), 100-alpha);
        yhat = d2 > thr;

        FPR = mean(yhat(y==0)==1);
        FNR = mean(yhat(y==1)==0);

        if FNR < best.FNR && FPR <= 2*alpha/100
```

```
            best = struct('FNR',FNR,'FPR',FPR,'kMin',kMin, ...
                'kMax',kMax,'thr',thr);
        end

        if mod(iter, updateEvery) == 0 || iter == totalIters
            waitbar(iter/totalIters, h, ...
                sprintf('Grid search: %d%% (%d/%d)', ...
                round(100*iter/totalIters), iter, totalIters));
            drawnow limitrate;
        end
    end
end

if ishandle(h), delete(h); end

disp(best)

%% CAMSD - Testing
fprintf('\n- CAMSD - Testing \n');

testPts = Z';
refPts = Cluster{1,1}';

opts = struct();
opts.kMin = best.kMin;
opts.kMax = best.kMax;
opts.doZScore = true;
opts.lambdaInit = 0; % start unregularized
opts.lambdaAuto = true; % auto-regularize if needed
opts.minRcond = 1e-5;

[d2, info2] = CAMSD(testPts, refPts, opts);

thr = prctile(d2(1:nVd), 100 * (1 - alpha/100)); % Threshold Boundary

figure('Name', 'CAMSD - d2', 'NumberTitle', 'off');
semilogy(d2); ylim([min(d2) max(d2)]); % ,'LineStyle','none','Marker','.'
xline(nVd,'k--','Healthy');
yline(thr,'k--','Threshold');
```

```
cp = findchangepts(d2, 'Statistic','mean', 'MaxNumChanges', 2);
disp(cp);

%%

DI = [d1;d2];
thr = prctile(DI(1:nHe), 100 * (1 - alpha/100));  % Threshold Boundary

falsePositives = sum(DI(1:nHe) > thr);
fpRate = (falsePositives / nHe)*100;
fprintf('False Positives (Undamaged Condition) = %d\n', falsePositives);
fprintf('False Positive Rate (Undamaged Condition) = %.2f%%\n\n', fpRate);

falseNegative = sum(DI(nHe+1:nFr) < thr);
fnRate = (falseNegative / (nFr-nHe) )*100;
fprintf('False Negative (Damaged Condition) = %d\n', falseNegative);
fprintf('False Negative Rate (Damaged Condition) = %.2f%%\n\n', fnRate);


figure('Name', 'Z24_TSC-CAMSD', 'NumberTitle', 'off');

f1=semilogy(1:nTr,DI(1:nTr),'LineStyle','none','Marker','+','MarkerSize',8, ...
    'LineWidth',1,'Color','k','MarkerFaceColor','k');
hold on
xline(nTr,'k:','Training','LineWidth',1.2);
hold on

f2=semilogy(nTr+1:nHe,DI(nTr+1:nHe),'LineStyle','none','Marker','*', ...
    'MarkerSize' ,8,'LineWidth',1,'Color','b','MarkerFaceColor','w');
hold on
xline(nHe,'k:','Healthy','LineWidth',1.2);
hold on

f3=semilogy(nHe+1:nFr,DI(nHe+1:nFr),'LineStyle','none','Marker', ...
    'pentagram','MarkerSize',8,'LineWidth',1,'Color','r', ...
    'MarkerFaceColor','w');
hold on
f4=yline(thr,'k--','Threshold','LineWidth',1.2);
```

```
hold on

set(gca,'XLim',[0 nFr+1],'YLim',[min(DI/10) max(DI*10)]); xlabel('Samples');
ylabel('\delta')

for i = 1:nHe
    if DI(i) > thr

        f5=plot(i,DI(i),'LineStyle','none','Marker','o','MarkerSize',12,...
            'Color',[1 0 1],'LineWidth',1.5);
        hold on
    end
end
for i = nHe+1:nFr
    if DI(i) < thr

        f6=plot(i,DI(i),'LineStyle','none','Marker','square','MarkerSize', ...
            12,'Color',[0 0 0],'LineWidth',1);
        hold on
    end
end

legend( [ f1 , f2 , f3 , f5 , f6 , f4 ] , ...
    'Undamaged Condition - Training Data' , ...
    'Undamaged Condition - Validation Data' ,...
    'Damaged Condition - Test Data' , ...
    'False Alarm Error (False Positives)' , ...
    'Missed Detection Error (False Negative)' , ...
    'Threshold Boundary' , ...
    'Location' , 'northwest' )
ax = gca; ax.FontName = 'Arial'; ax.FontSize = 10;
set(gcf,"Visible","on")

%%

% Export as PNG with fixed size (cm) and DPI
fig = findall(0,'Type','figure','Name','Z24_TSC-CAMSD');
fig = fig(1);            % if multiple matches exist
w_cm = 20; h_cm = 20; dpi = 300;
```

```
set(fig,'Units','centimeters');
fig.Position(3:4) = [w_cm h_cm];
exportgraphics(fig,'Z24_TSC-CAMSD.png','Resolution',dpi);

```

Acknowledgments

I would like to preface by expressing my most sincere gratitude to my supervisors, Prof. Alberto Corigliano and Prof. Alireza Entezami, for their guidance and constant support which allowed me to deepen my understanding in the field of structural health monitoring, and successfully develop an academic thesis independently.

I would also like to thank my family and friends for their unconditional love, support, and trust; their presence and encouragement were essential to everything I have achieved so far.

Last but certainly not least, with the hope of freedom for Iran, I dedicate this work to all who continue to strive for a better future. Long live Iran.