Adaptive Tensor-Based DMD for Changepoint Detection in Multidimensional Time-Series

Marion Kokkinou

Student ID: 06015400

Supervised by Dr. Edward A. K. Cohen

Date: 29/08/2025

Submitted in partial fulfilment of the requirements for the MSc in Statistics at Imperial College London

1 Acknowledgements

In this section I would like to express my sincere gratitude to both my supervisor Dr. Edward A. K. Cohen and Victor K. Khamesi for their guidance and support throughout this journey.
Additionally, I am deeply moved by my friends and family who have encouraged and supported me, not only for the dissertation but also for the whole MSc year.

Abstract

This dissertation introduces an adaptive tensor-based framework for change-point detection in multidimensional time series. The approach integrates dynamic mode decomposition (DMD) with tensor-train and Tucker decompositions. Therefore, it enables automatic adaptation to complex, high- dimensional data characterised by multiple dynamic regimes. Extensive simulations demonstrate the robustness of the method, showing that it achieves high accuracy and minimal delay in detecting abrupt structural shifts. These shifts are consistently signalled by sharp increases in reconstruction error that coincide with underlying changes in system dynamics.

The combined scalability and resilience to noise of the framework emphasise its readiness for implementation in real-world settings. The framework's practical relevance is evidenced in the Santander bike-sharing dataset, where it uncovers meaningful changes in usage and operational patterns over time. The findings demonstrate the method's capability in capturing the evolving dynamics of complex, non-stationary systems.

Although particularly strong in detecting sudden changes, the method shows limited sensitivity to gradual shifts, underscoring the need for refinement. Future research will address this limitation, aiming to expand the method's applicability and improve its performance across diverse multidimensional time series settings.

Contents

2 Introduction

Nowadays, detection plays a pivotal role in time-series analysis, as it helps detect moments when a system experiences significant changes (Khan et al., 2016). These shifts can be abrupt or gradual. Thus, identifying them can often provides valuable insight into how the system evolves over time. In many cases, such detection supports better decision-making and monitoring strategies. This dissertation will focus on applying changepoint detection methods to the Santander bike-sharing dataset. The aim is to capture meaningful shifts in usage and operations, providing guidance for everyday decision-making and the design of future improvements.

Typical changepoint detection techniques primarily target univariate or low- dimensional data, generally looking for changes in statistics like mean or variance. Such methods,including CUSUM, EWMA, or Bayesian approaches form a solid foundation (Han et al., 2017). At the same time, however, they can struggle to capture the complexity of real-world data. For instance, the Santander dataset contains multiple interrelated variables that fluctuate with daily, weekly, and seasonal rhythms, making simple approaches less effective. Therefore, more advanced methods are needed, in order to model high-dimensional, non- stationary data with rich temporal dependencies.

To address these challenges, this study adopts an approach based on Dynamic Mode Decomposition (DMD), a technique originally developed in fluid dynamics that separates multidimensional time series into spatial and temporal components, thereby revealing dominant patterns [8]. When combined with tensor decomposition methods, this framework can capture and represent the inherent multi-way structure of the data in greater depth [?]. The guiding idea is that shifts in system behaviour can be identified by monitoring the reconstruction error from low-rank approximations, with abrupt increases in this error signalling potential changepoints. This focus on reconstruction-based detection complements recent statistical approaches that integrate multiple streams of data, such as the higher-criticism framework for sparse changepoint detection proposed by Gong et al. (2025).

A deeper examination of the existing DMD-based changepoint detection allows the techniques used in the proposed framework to be adapted specifically for the Santander-Bike dataset. Although the method performs well in detecting abrupt shifts, gradual changes or subtle transitions may still present challenges. This dissertation investigates the capabilities and limitations of this approach, aiming to develop a scalable and robust tool for analysing complex, high-dimensional time series in practical settings.

3 Methods

This section outlines the methodological framework for changepoint detection in multidimensional time series, with a particular focus on the Santander bike-sharing system. The objective is to extend Dynamic Mode Decomposition (DMD) to a tensor-based framework (T-DMD), allowing for effective detection of structural shifts in high-dimensional, non-stationary data.

Let $\{X_t\}_{t=1}^{T}$ denote a multivariate time series, where $X_t \in R^{p\times d}$ represents the observation matrix at time t. Here p corresponds to the number of spatial locations (bike stations) and d the set of recorded features at each location (such as rentals, returns, or availability). This formulation captures both spatial and temporal dependencies within the evolving system. The aim is to identify changepoints

$$\tau_1 < \cdots < \tau_K$$

at which the statistical properties of the process undergo significant variation, without imposing restrictive distributional assumptions. In addition, tensor decomposition is employed to handle the high dimensionality of the data while preserving its multi-way structure.

The methodological foundation builds on the DMD framework introduced by [8], which uses reconstruction error from low-rank linear approximations as a diagnostic for dynamical shifts. In the present work, this approach is adapted to the bike-sharing dataset using a vectorised implementation consisting of three key components:

1. **Hankelisation:** From the time series $\{X_t\}$, overlapping windows of length w are embedded into Hankel matrices. This step encodes temporal dependencies by representing past and future observations within a structured matrix form.

2. **DMD Computation:** A truncated singular value decomposition (SVD) is applied, yielding spatial modes Φ and temporal coefficients. Following the approach of [8], these components provide low-rank approximations of the system dynamics, from which reconstruction error can be computed.

3. **Adaptive Thresholding:** The sequence of reconstruction errors is monitored using an exponentially weighted moving average (EWMA) control chart. This chart is parametrised by a smoothing factor γ and threshold ℓ, enabling the detection of abrupt deviations that indicate a changepoint. As previously discussed, this paper builds on the DMD framework of [8] by introducing a tensor-based extension (T-DMD). Instead of reducing the data to matrices, T-DMD works directly with tensors, preserving their multi-way structure and capturing richer spatio-temporal interactions. Through tensor algebra and decomposition, the method improves efficiency and increases sensitivity to subtle or distributed changepoints in high-dimensional time series.

3.1 Definitions

We define a *changepoint* as a time index τ_i at which the distribution of the process shifts. Formally, let $\{\mathbf{X}_t\}_{t=1}^T$ denote a multivariate time series of length T. Suppose there exist changepoints $0 = \tau_0 < \tau_1 < \cdots < \tau_K < \tau_{K+1} = T$, such that for

$$t \in [\tau_{k-1} + 1, \tau_k], \quad \mathbf{X}_t \sim \mathcal{F}_k,$$

where the $\mathcal{F}_k$ are distributions that differ across segments $k = 1, \ldots, K$. Thus, the statistical properties of the sequence are piecewise stationary but may change abruptly at $\tau_1, \ldots, \tau_K$.

A *point network* is a set of graph vertices $V = v_1, ..., v_p$, where each v_j corresponds to a node. Therefore, data are represented in a $p * T$ dimensional matrix in the form of a high order dimensional time series.

3.2 Problem Setting

Let $\{\mathbf{X}_t\}_{t=1}^{T}$ denote a multivariate time series, where each observation $\mathbf{X}_t \in R^{p\times d}$ corresponds to p spatial locations (e.g., bike stations) and d features recorded at each location (e.g., rentals, returns, availability). Formally, assume there exist K changepoints

$$0 = \tau_0 < \tau_1 < \ldots < \tau_K < \tau_{K+1} = T,$$

partitioning the time series into $K + 1$ contiguous segments. Within segment k ($t \in [\tau_{k-1} + 1, \tau_k]$), the data are assumed to follow a stationary distribution $\mathcal{F}_k$:

$$\mathbf{X}_t \sim \mathcal{F}_k, \quad t \in [\tau_{k-1} + 1, \tau_k], \quad \mathcal{F}_k \neq \mathcal{F}_{k+1}.$$

The objective is to estimate both the number of changepoints K and their locations $\tau_1, \ldots, \tau_K$ from the observed sequence $\mathbf{X}_{1:T}$.

For high-dimensional or multi-way data, such as the Santander bike-sharing dataset, observations can be represented as tensors:

$$\mathcal{X}_t \in R^{n_1\times n_2\times\cdots\times n_d},$$

where the dimensions correspond to spatial locations, features, and other attributes. The tensor sequence is assumed to contain changepoints $c_1 < c_2 < \ldots < c_K$ such that

$$\mathcal{X}_t \sim \mathcal{F}_k, \quad t \in [c_{k-1} + 1, c_k], \quad \mathcal{F}_k \neq \mathcal{F}_{k+1}.$$

The dissertation will revolve around this conceptin order to detect changepoints using Dynamic Mode Decomposition (DMD) and its tensor extensions. Each tensor snapshot $\mathcal{X}_t$ is used to compute low-rank approximations, and the resulting reconstruction errors serve as the basis for identifying points where the system dynamics change. By defining the problem in this way, the theoretical changepoint concept is connected directly to the structure of the data and the computations performed in the DMD algorithm.

3.3 Benchmark Method: Vectorised DMD with EWMA

The benchmark method for changepoint detection extends classical Dynamic Mode Decomposition (DMD) by incorporating Hankel (time-delay) embeddings, consistent with the actual code implementation. Whereas traditional DMD operates on matrix-valued time series reshaped into two-dimensional matrices, Hankel-DMD constructs augmented matrices by stacking time-delayed snapshots. This enriches the spatial dimension and allows more effective characterisation of complex temporal dynamics (Khamesi et al., 2024).

Formally, let

$$X \in R^{m\times T}$$

denote the vectorised data matrix, where $m = d_1 \cdot d_2$ is the number of spatial features and T the number of time points. For each sliding window of length w, a Hankel matrix

$$H_t \in R^{m\cdot d_{delay}\times w}$$

is constructed by stacking d_{delay} delayed snapshots:

$$H_t = x_{t-w+1} x_{t-w+2} \cdots x_{t-w+w} x_{t-w+2} x_{t-w+3} \cdots x_{t-w+w+1} \vdots \vdots \ddots \vdots x_{t-w+d_{delay}} x_{t-w+d_{delay}+1} \cdots x_t,$$

where each $x_t \in R^m$ is the vectorised snapshot at time t.

A truncated singular value decomposition (SVD) is performed on H_t to obtain a rank-r approximation

$$\hat{H}_t = U_r \Sigma_r V_r^\top,$$

providing a low-rank reconstruction of the observed window. The reconstruction error is measured as the normalised Frobenius norm squared:

$$\epsilon(t) = \frac{\|H_t - \hat{H}_t\|_F^2}{m \cdot w}.$$

The increments of this error,

$$\Delta\epsilon(t) = \epsilon(t) - \epsilon(t-1),$$

are monitored using an adaptive Exponentially Weighted Moving Average (EWMA) control chart, parametrised by a smoothing factor γ and a decision threshold ℓ. Large positive deviations in $\Delta\epsilon(t)$ are interpreted as detected changepoints [8].

Hankel-DMD differs from plain vectorised DMD by explicitly incorporating temporal delay embeddings, producing richer spatio-temporal representations and enabling more robust detection of subtle or transient changes. By comparing the tensor-based DMD approach against this benchmark, the benefit of retaining the full tensor structure can be systematically assessed.

3.4 Data Vectorisation

The original dataset is represented as a three-dimensional tensor $\mathcal{X} \in R^{p \times p \times T}$, where p denotes the number of stations and T the number of weekly time points. Each entry captures the journeys between station pairs across time. For standard DMD processing, this tensor is vectorized into a matrix $X_{\text{vec}} \in R^{m \times T}$, where m is the number of active station pairs, such as ones with at least one recorded journey. This transformation enables the application of classical DMD, though at the expense of explicitly representing the tensor structure.

3.5 Dynamic Mode Decomposition

From the vectorised data, a Hankel matrix is constructed using sliding windows of length w and embedding order o. Singular Value Decomposition (SVD) is applied and truncated at rank r, deriving spatial modes Φ, temporal dynamics, and reconstructed approximations $\hat{X}$. At each time t, reconstruction accuracy is measured by the normalised Frobenius norm of the error:

$$\epsilon(t) = \frac{X_t - \hat{X}_{tF}^2}{m \times w}.$$

3.6 EWMA-Based Changepoint Detection

The increments $\Delta\epsilon(t)$ of the reconstruction error sequence are monitored using an adaptive EWMA control chart [**?**]. The scheme is parametrised by a smoothing factor γ, a decision threshold ℓ, and appropriately chosen burn-in and history lengths to ensure stable performance. A changepoint is recorded whenever the EWMA statistic crosses the threshold, signalling a sudden change in system dynamics.

Khan et al. (2016) show that carefully adjusting the parameters of changepoint detection methods, including the use of genetic algorithms, can substantially improve their performance. Their results indicate that selecting appropriate values for the smoothing factor and decision thresholds not only increases detection accuracy but also reduces false alarms, highlighting the importance of parameter tuning in practical applications.

3.6.1 6. Changepoint Detection; Adaptive EWMA

As discussed above, it is clear that the increment statistics of reconstruction errors might signify distribution shifts, regime changes or other shifts of the underlying generative process. For this reason, the methodology applied introduces the *a*daptive EWMA; an algorithm that monitors both the volatility and magnitude of the reconstruction error increments. As a result, deviation shifts that are considered to be statistically important are said to be changepoints, which aligns with the best practises of Khamesi et al. (2024) for online changepoint detection. Beyond this, however, the offline function of the algorithm still returns the same detection except for the latency in waiting for new data in real time.

Adaptive EWMA Change Detection Statistic. Let $\delta_t \in R$ denote the observed change statistic at time step t (e.g. the increment of the reconstruction error between successive windows).

We maintain streaming estimates of the mean μ_t and variance σ_t^2 of δ_t, updated adaptively as: $\mu_t = \mu_{t-1} + \frac{\delta_t - \mu_{t-1}}{\min(10,t)}$,
$\sigma_t^2 = \max\left(10^{-4},\ \left(1 - \frac{1}{\min(10,t)}\right)\sigma_{t-1}^2 + \frac{(\delta_t - \mu_{t-1})^2}{\max(1,t-1)}\right)$, with initialisation $\mu_0 = 0$, $\sigma_0^2 = 10^{-4}$.

The normalised deviation is then computed as:

$$z_t = \frac{\delta_t - \mu_t}{\sigma_t}, \tag{1}$$

where $\sigma_t = \sqrt{\sigma_t^2}$ ensures standardisation against the current adaptive variance estimate.

An exponentially weighted moving average (EWMA) statistic is updated via:

$$S_t = (1-\gamma)S_{t-1} + \gamma z_t, \tag{2}$$

where $\gamma \in (0,1]$ is the smoothing parameter controlling the weight on the current observation.

A *change* is declared at time t if:

$$|S_t| > \ell \quad and \quad t \geq 3, \tag{3}$$

where $\ell > 0$ is the detection threshold. The constraint $t \geq 3$ enforces a minimum number of samples before detection.

3.7 Validation and Comparison

The changepoints identified through the DMD-EWMA benchmark are treated as reference events, reflecting meaningful shifts in usage and operational patterns. By comparing these benchmark results with those produced by the tensor-based DMD, the accuracy and robustness of the tensorised framework can be systematically evaluated. Particular attention is given to agreements and discrepancies between the two methods, which guide adjustments to embedding order, truncation rank, and tensor decomposition strategy.

3.8 Evaluation and Visualisation

Both sets of detected changepoints are visualised against the observed bike journey volumes for selected pair of stations. Plots of the reconstruction error highlight the timing of changepoint events, hence enabling direct comparison between methods. Alignment with the benchmark strengthens confidence in the tensor-based approach, while divergences suggest directions for refinement.

This benchmarking process ensures that the proposed tensor-based DMD is tested against a well-founded baseline.As a result, its capacity to detect meaningful changepoints in complex, high-dimensional, and non-stationary mobility data can be examined thoroughly.

3.9 TT-DMD algorithm

The findings of this paper revolve around the following algorithm, known as *TT-DMD algorithm*:

Let

$$\mathcal{X} \in R^{C\times F\times T}$$

denote a third-order tensor stream of EEG recordings, where C is the number of channels, F the number of features per channel, and T the total number of time samples. Given a sliding window of length w and a delay-embedding order d_{delay}:

$c \in \{1, \ldots, C\}$ corresponds to a multivariate signal representing the future vector revolution over time. For each component a multivariate time series is observed

$$\mathbf{x}_c = (x_{c,f,t})_{f=1,\ldots,F;\, t=1,\ldots,T} \in R^{F\times T},$$

where f enumerates feature dimensions (e.g. spectral bands or derived statistics) and t indexes discrete time. Collecting all channels yields a third-order data tensor

$$\mathcal{X} = (x_{c,f,t}) \in R^{C\times F\times T}.$$

Rather than analysing channels or features in isolation, the tensor-based DMD framework operates directly on $\mathcal{X}$ so as to retain cross-channel and cross-feature dependencies as they evolve over time. A delay embedding is introduced by forming Hankel (time-lag) tensors:

$$\mathcal{H}^{(\tau)} = \mathrm{Hankel}_{\tau}(\mathcal{X}) \ \in \ R^{C \times F \times \tau \times (T-\tau+1)},$$

where τ denotes the embedding depth (number of lags). This construction enriches the temporal dimension by stacking successive time slices into a structured higher-order object.

Subsequent low-rank tensor decompositions (e.g. Tucker, tensor-train, or related multilinear factorizations) applied to $\mathcal{H}^{(\tau)}$ produce dominant spatiotemporal modes that jointly capture:

- inter-channel (spatial) coupling,
- inter-feature correlations, and
- temporal evolution encoded through the delay dimension.

The TT-DMD (tensor-train Dynamic Mode Decomposition) algorithm leverages this multilinear structure to extract coherent dynamical patterns while preserving separability across modes. The following steps are compiled in order for the algorithm to function:

4. **Hankel embedding:** Let $\{\mathbf{X}_k\}_{k=1}^{T}$ be a matrix-valued time series with $\mathbf{X}_k \in R^{d_1 \times d_2}$. Write each scalar channel as $x_k^{(i,j)}$ for $i = 1, \ldots, d_1$ and $j = 1, \ldots, d_2$, and set $d = d_1 d_2$. Fix a window length w and a delay depth q with $1 \leq q \leq w$. For $k \geq w$, the (univariate) Hankel matrix associated with channel (i, j) is

$$\chi_k^{(i,j)} = x_{k-w+1}^{(i,j)} x_{k-w+2}^{(i,j)} \cdots x_{k-q+1}^{(i,j)} x_{k-w+2}^{(i,j)} x_{k-w+3}^{(i,j)} \cdots x_{k-q+2}^{(i,j)} \vdots \vdots \ddots \vdots x_{k-q+1}^{(i,j)} x_{k-q+2}^{(i,j)} \cdots x_{k}^{(i,j)} \in R^{q \times (w-q+1)}.$$

Stacking all channel Hankel matrices along a new mode yields the 3-way Hankel tensor

$$\chi_k = [\chi_k^{(1,1)}, \ \chi_k^{(1,2)}, \ \ldots, \ \chi_k^{(d_1,d_2)}] \ \in \ R^{q \times (w-q+1) \times d},$$

(or, equivalently, after permuting modes, $\mathcal{H}_k \in R^{d \times q \times (w-q+1)}$), which serves as the delay-embedded representation for subsequent tensor/DMD analysis (Zhang et al., 2018).

5. **Tensor Decomposition:** Perform a low-rank Tucker (or tensor-train) decomposition with rank $\mathbf{r}_{\mathrm{tt}} = (r_1, r_2, r_3)$:

$$\mathcal{H} \approx \mathcal{G} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)},$$

where $\mathcal{G}$ is the core tensor and $U^{(n)}$ are the factor matrices in mode n.

6. **TT-DMD Projection:** Project the time-shifted Hankel tensor slices into the low-dimensional subspace spanned by $(U^{(1)}, U^{(2)}, U^{(3)})$ to obtain matrices X and Y, respectively corresponding to past and present states. (Zhang et al., 2018)

7. **DMD Operator Estimation:** This algorithm presents the estimation of the reduced-order linear evolution operator as follows:

$$A = YX^{+},$$

where X^{+} denotes the Moore-Penrose pseudoinverse.

8. **Spectral Analysis:** By computing the eigenvalues $\{\lambda_j\}$ and modes $\{\phi_j\}$ of A, the dominant spatio-temporal dynamics are characterised.

9. **Reconstruction and Error:** Allow us to use the model in order to reconstruct $\widehat{\mathcal{H}}$, which is the low rank approximation of the Hankel matrix H. This way the normalised reconstruction error is computed:

$$E_t = \frac{\|\mathcal{H} - \widehat{\mathcal{H}}\|_F}{\|\mathcal{H}\|_F + \varepsilon}.$$

10. **Online Changepoint Detection:** Feed the increment $\delta_t = E_t - E_{t-1}$ into the adaptive EWMA detector; if the detection statistic S_t satisfies

$$|S_t| > \ell,$$

declare a changepoint at time step t and, optionally, re-optimise $(w, d_{\text{delay}}, \mathbf{r}_{\text{tt}})$ over the next burn-in segment.

To lay the framework for the construction of the algorithm, let $\mathcal{X}, \mathcal{Y}$ denote a sequence of vectors

$$\{x_t\}_{t=1}^{m}, \quad x_t \in R^N, \quad N = \prod_{i=1}^{d} n_i,$$

where each x_t is obtained by flattening the tensor observation $\mathcal{X}_t$ with axes corresponding to channels, features, and other modes. As mentioned before, the observations are organised as tensors with axes for channels, features and frequency respectively. As mentioned above, by manipulating memory and computation efficient tensor- train decompositions, the dominant dynamic modes are returned for changepoint detection.
The windows pose a pivotal element in the algorithm as we investigate each one separately in order to examine a sub-tensor linked to the bike dataset.

3.10 Generate Synthetic 3D Data

3.11 Simulated Multichannel Time Series

We simulate a four-channel matrix-valued time series

$$\mathbf{X}_t \in R^{2\times 2}, \quad t = 1, \ldots, N,$$

with N = 1,000 time points, dt = 0.01 s, and total duration T = 10 s. Each entry $X_{i,j,t}$ represents the amplitude at time t for channel (i, j), following segmented oscillatory dynamics with additive Gaussian noise.

Changepoints are defined at

$$t_1 = 2 \text{ s}, \quad t_2 = 6 \text{ s}, \quad t_3 = 8 \text{ s},$$

producing four contiguous data segments:

1. Segment 1 ($0 \le t < 2$): frequencies $f = 9$–11 Hz, amplitude $A_1 = 1.0$, noise $\sigma_1 = 0.05$.
2. Segment 2 ($2 \le t < 6$): frequencies $f = 18$–22 Hz, amplitude $A_2 = 1.5$, noise $\sigma_2 = 0.1$.
3. Segment 3 ($6 \le t < 8$): frequencies $f = 5$–7 Hz, amplitude $A_3 = 1.8$, noise $\sigma_3 = 0.15$.
4. Segment 4 ($8 \le t \le 10$): frequencies $f = 38$–42 Hz, amplitude $A_4 = 2.0$, noise $\sigma_4 = 0.2$.

Let $\mathcal{I}_k$ note the index set corresponding to segment k. The entries are generated as $\mathrm{X}_{i,j,t} = A_k \sin(2\pi f_{i,j,k}\, t) + \varepsilon_{i,j,t}, \qquad t \in \mathcal{I}_k,$
$\varepsilon_{i,j,t} \sim \mathcal{N}(0, \sigma_k^2)$, where $f_{i,j,k}$ is the assigned frequency for channel (i,j) in regime k, A_k is the amplitude, and σ_k is the noise level.

Vectorising each snapshot as

$$x_t = \mathrm{vec}(\mathbf{X}_t) \in R^N, \quad N = 2 \cdot 2 = 4,$$

produces the sequence $\{x_t\}_{t=1}^N$ suitable for DMD or Hankel-DMD computations. In code, the assignment is implemented as $\mathrm{t} \in \mathcal{I}_1 : \quad x_t = \mathrm{vec}(1.0 \sin(2\pi f_{i,j,1} t) + 0.05\, \xi_{i,j,t})$
$t \in \mathcal{I}_2 : \quad x_t = \mathrm{vec}(1.5 \sin(2\pi f_{i,j,2} t) + 0.1\, \xi_{i,j,t})$
$t \in \mathcal{I}_3 : \quad x_t = \mathrm{vec}(1.8 \sin(2\pi f_{i,j,3} t) + 0.15\, \xi_{i,j,t})$
$t \in \mathcal{I}_4 : \quad x_t = \mathrm{vec}(2.0 \sin(2\pi f_{i,j,4} t) + 0.2\, \xi_{i,j,t})$, where $\xi_{i,j,t} \sim \mathcal{N}(0,1)$ are independent standard normal random variables with $\xi_{i,j,t} \sim \mathcal{N}(0,1)$ i.i.d. Gaussian noise.

The function returns the data tensor, time vector, and changepoint times, suitable for use in downstream adaptive DMD changepoint detection.

3.12 Using the real dataset

Traditionally, changepoint methods have centred on low-dimensional or univariate streams of data, often searching for mean or variance shifts. Older methods such as CUSUM, EWMA, and Bayesian models provide a good starting point but may lack the flexibility to catch subtle patterns in high-dimensional, non-stationary data. For instance, the Santander dataset has a number of variables with daily, weekly, and seasonal cycles all mixed up together in an intricate manner. This requires more sophisticated tools capable of addressing such high-dimensional temporal relationships.

To respond to this difficulty, in the present work Dynamic Mode Decomposition (DMD), a technique initially derived in the context of fluid mechanics, is applied. DMD decomposes a time series within high-dimensional space into temporal and spatial modes to reveal salient patterns by modes $\mathbf{\Phi}_i$ and the related eigenvalues λ_i that preserve oscillatory or growth/decay behaviour across time. By intertwining DMD with tensor decomposition methods—e.g., tensor-train and Tucker decompositions—the data's multiway nature is preserved, which supports richer representation and analysis.

The philosophy underlying changepoint detection here is based on monitoring the reconstruction error $\epsilon(t)$ resulting from low-rank approximations to the data. Sudden spikes in $\epsilon(t)$ serve as indicators of alteration in the underlying process. The method adopts recent advances in DMD-based changepoint detection and adapts them to the specific nature of the Santander bike data.

While this approach shows promising potential in identifying abrupt changes, there may be some underlying more nuanced variations and smooth transitions. The strengths and limitations of the method are explored throughout the dissertation with the aim of developing a valid, scalable algorithm that can decompose complex, high-dimensional time series for practical application.

Changepoint detection is a fundamental technique of time-series analysis required to identify points when a system's behaviour witnesses drastic changes. Changes may be abrupt or prolonged, and their identification typically reveals meaningful insight into the evolving dynamics of complex processes. This dissertation emphasises the application of changepoint detection methods on data from the Santander bike-sharing system. The objective is to detect shifts in usage and operation patterns that can, to some extent, contribute towards day-to-day functioning as well as long-term strategic choices. Conventional changepoint techniques often deal with univariate or low-dimensional data focusing on mean or variance changes. Traditional techniques—such as CUSUM, EWMA, and Bayesian algorithms—are solid beginnings but may fall short when addressing the fine details of multidimensional non-stationary data (Han et al., 2017), but may fall short when addressing the fine details of multidimensional non-stationary data. For example, the Santander dataset contains multiple interconnected variables that have daily, weekly, and seasonal trends all intricately intertwined. These require more advanced means to manage the high-dimensional temporal associations involved.

To address these challenges, this research utilises Dynamic Mode Decomposition (DMD), a method originating from fluid dynamics that breaks high-dimensional time series into spatial-temporal modes. The decomposition comprises modes $\mathbf{\Phi}_i$ and eigenvalues λ_i describing oscillatory or growth/decay behaviour over time. Coupling DMD with tensor decomposition techniques such as the tensor-train and Tucker formats maintains the multi-way structure of the data, enabling richer and more detailed analysis.

Central to the changepoint detection algorithm is monitoring the reconstruction error $\epsilon(t)$ from low-rank approximations of the data. Abrupt spikes in $\epsilon(t)$ indicate potential alterations in the system's underlying dynamics. Building on recent progress in DMD-based detection, the method is optimised particularly for the special characteristics of the Santander bike dataset.

While proficient at identifying sudden changes, the method may be vulnerable to small or slow changes. This dissertation therefore examines both the robustness and weaknesses of the method, with a view to proposing an effective, robust methodology suitable for the analysis of complex, high-dimensional time series in practical problems.

3.13 Background: Dynamic Mode Decomposition

This paper will follow the definition of the DMD as given by Klaus (2018), Li et al. (2024) and Khamesi et al. (2024), describing it as a data driven operator-theoretic technique used to examine the behaviour of a complex dynamical system, by splitting up the data into spatio-temporal structured groups. The DMD drives from principal component analysis and from the Fourier analysis, accounting for both spatial and time elements respectively.

3.14 Tensor Formulation of Dynamic Mode Decomposition

In classical Dynamic Mode Decomposition (DMD), the linear regression operator A that maps past states to future states is computed explicitly as

$$A = YX^{+},$$

where X^{+} denotes the Moore-Penrose pseudo-inverse of the snapshot matrix X. This formulation requires flattening multidimensional observations into vectors and performing potentially large matrix inversions, which can be computationally expensive and memory-intensive for high-dimensional data. By contrast, the tensor-train (TT) formulation of DMD expresses the operator implicitly in terms of the factor matrices and core tensor obtained from a Tucker or TT decomposition of the Hankel-embedded tensor snapshots. All linear algebra operations, including multiplications and inversions, are carried out within the TT algebra, maintaining the computations in the compressed low-rank tensor space. This approach preserves the inherent multi-way structure of the data, avoids the need for full vectorization, and dramatically reduces the computational cost associated with large-scale matrix operations. As a result, TT-based DMD effectively mitigates the curse of dimensionality while still capturing the dominant spatio-temporal modes of the system. The method provides a computationally efficient and structure-preserving extension of classical DMD, allowing high-dimensional, multivariate time series, such as those from multi-sensor or networked systems, to be analysed without sacrificing interpretability or accuracy Klus2018.

In the tensor formulation of Dynamic Mode Decomposition (DMD), the classical idea of identifying a linear operator A that maps one snapshot to the next,

$$Y \approx AX,$$

is generalised to higher-order tensors.

Instead of working with matrices $X, Y \in R^{n \times m}$, tensor DMD operates directly on tensor snapshots

$$\mathcal{X}_t \in R^{n_1 \times n_2 \times \cdots \times n_d}, \quad t = 1, \ldots, T,$$

indexed by time. The goal is to find a transition tensor $\mathcal{A}$ such that

$$\mathcal{X}_{t+1} \approx \mathcal{A} \star \mathcal{X}_t,$$

where $\star$ denotes a suitable tensor product, e.g., the T-product for third-order tensors.

This transition tensor $\mathcal{A}$ captures the evolution of the multi-way data while preserving the intrinsic mode structure of the tensor, avoiding the need to flatten it into matrices. Using tensor decomposition techniques such as truncated tensor SVD (t-SVD) and tensor eigenvalue decomposition (TEVD), the tensor dynamic modes and their corresponding eigenvalues can be computed. These modes represent the dominant spatio-temporal patterns across all tensor dimensions, analogous to the modes in classical matrix DMD, but with explicit multi-dimensional interpretability.

In summary, tensor DMD maintains the multidimensional relationships inherent in the data, allowing for structure-preserving approximations and richer, physically interpretable decompositions. This approach is particularly advantageous in applications such as video processing, biomedical signals, and multi-sensor networks, where preserving the tensor structure can improve the accuracy of dynamic characterisation compared to classical matrix-based DMD applied to vectorised data.

3.14.1 1. Delay Embedding and Hankelisation

Let

$$X \in R^{d\times d\times w}$$

denote a matrix-valued time series of length w, where d is both the number of rows and columns (spatial dimensions).

A *sliding window* of length q along the time dimension of a *single* vector component of the series produces a Hankel block

$$\mathcal{H}^{(p)} \in R^{q\times(w-q+1)\times 1},$$

where p indexes the component and

$$\mathcal{H}^{(p)}_{i,j,1} = x^{(p)}_{i+j-1}, \quad i = 1,\dots,q, \quad j = 1,\dots,w-q+1.$$

In the vector case of length-d_{vec} time series, Khamesi et al. (2024) have derived a construction which stacks these small Hankel blocks

$$\mathcal{H}^{(p)}, \quad p = 1,\dots,d_{vec},$$

vertically to form a tensor of size

$$R^{d_{vec}\,q\times(w-q+1)\times 1}.$$

—

In our **matrix-valued case**, each observation $X_t \in R^{d\times d}$ is treated as d row-vectors, each producing its own $q\times(w-q+1)\times 1$ Hankel block. We then stack these blocks for all d rows along a new spatial mode to obtain

$$\mathcal{H} \in R^{d\times q\times(w-q+1)\times d},$$

which can equivalently be reshaped into

$$\mathcal{H} \in R^{(dq)\times(w-q+1)\times d}.$$

—

Thus, the classic $d_{vec}q\times(w-q+1)\times 1$ Hankel tensor for vector-valued series is generalised here to a $dq\times(w-q+1)\times d$ tensor, preserving both time-delay and *two-dimensional spatial* structure:

$$\mathcal{H}[(r,i),j,c] = X_{r,\,c,\,i+j-1},$$

for $r,c = 1,\dots,d$, $i = 1,\dots,q$, $j = 1,\dots,w-q+1$.

3.14.2 2. Tucker (TT) Decomposition

The Tensor-Train (TT) decomposition is initiated by applying the Tucker decomposition, as implemented in the TensorLy library, which yields the factorised form consisting of a core tensor together with factor matrices along each mode. This representation highlights the most dominant components and thereby mitigates the curse of dimensionality Li2024.

Let

$$\mathbf{X} \in R^{n_f\times w}$$

denote a multivariate time series segment consisting of n_f features (channels) and a window of length w.

3.14.3 4. Mode/ Eigenvalue extraction

Within the TT-reduced space the eigendecomposition takes place so as to extract dynamic modes. Later on, we will use these to reveal spatio-temporal patterns common in EEG's along with their corresponding eigenvalues representing frequencies and growth/decay rates. (Li et al., 2024)

3.14.4 5. Reconstruction error computation

The data processed so far needs to be reconstructed in a way that compares it to the original tensor and returns the windowed reconstruction error. The spikes that appear from the reconstruction error empirically signal that there is a change point. The implemented algorithm measures the normalised discrepancy between the true and the TT-DMD reconstructed data. (Khamesi et al., 2024)

Given an input time-series segment

$$\mathbf{X} \in R^{n_f \times w},$$

we first construct its Hankel tensor

$$\mathcal{H} \in R^{d_{\text{order}} \times n_f \times n_s},$$

where $n_s = w - d_{\text{order}} + 1$ is the number of snapshots.

We then compute the Tucker decomposition

$$\mathcal{H} \approx \mathcal{G} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)},$$

where $\mathcal{G}$ is the core tensor and $U^{(i)}$ are the mode-i factor matrices, with ranks

$$r_i \leq \dim_i(\mathcal{H}).$$

From this factorisation, we construct the low-rank approximation

$$\widehat{\mathcal{H}} = \mathcal{G} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}.$$

The *normalised reconstruction error* is then defined as the relative Frobenius norm:

$$E = \frac{\left\| \mathcal{H} - \widehat{\mathcal{H}} \right\|_F}{\|\mathcal{H}\|_F + \varepsilon}, \tag{4}$$

where $\| \cdot \|_F$ denotes the Frobenius norm and ε is a small constant (e.g. 10^{-12}) included for numerical stability.

3.14.5 7. Hyperparameter selection

The remaining step in order to complete the TT-DMD algorithm is selecting the most suitable hyperparameters in order to optimise robustness. Data driven search aids at identifying the appropriate hyperparameters by minimising the mean absolute error increments on a burn-in segments Khamesi et. al.2024.

Mean Absolute Increments Statistic and Hyperparameter Optimisation. To select optimal detection hyperparameters for the Tensor DMD pipeline, we define the *mean absolute increments statistic* as follows.

Let $\mathbf{X}_{seg} \in R^{n_f \times L_{seg}}$ be a data segment of n_f features and length L_{seg}. For fixed window length w, embedding order d_{order}, and Tucker rank $\mathbf{r}$, one computes the sequence of reconstruction errors

$$E_t = E\left(\mathbf{X}_{seg}[:, t-w : t]\right), \qquad t = w,\, w+1,\, \ldots,\, L_{seg},$$

where $E(\cdot)$ is the normalised reconstruction error as defined in eq:recon$_e rror$.

The *error increments* are the first differences between successive errors:

$$\Delta_t = E_t - E_{t-1}.$$

The mean absolute increment statistic is then given by

$$S(w,\, d_{order},\, \mathbf{r}) = \frac{1}{N}\sum_t |\Delta_t|, \tag{5}$$

where the sum and N are taken over the sampled t in the segment (see code for details).

This statistic quantifies the typical magnitude of error fluctuation under fixed detection parameters. Smaller values correspond to more stable error profiles (and thus less likelihood of false change detections), while abrupt changes produce larger increments.

We select the detection parameters $(w,\, d_{order},\, \mathbf{r})$ by minimising the mean absolute increment statistic over plausible grids:

- **Windows:** $w \in \{20, 30, 40\}$,
- **Embedding Orders:** $d_{order} \in \{5, 8\}$,
- **Ranks:** $\mathbf{r} \in \{(2,2,5),\ (2,2,8),\ (3,3,6)\}$.

Only parameters with $d_{order} < w$ are allowed.
For each triple $(w,\, d_{order},\, \mathbf{r})$, the mean absolute increment is evaluated over the **burn-in segment** (a short initial block of the time series, e.g. $L_{burn} = 40$). The parameter set yielding the lowest S is selected for subsequent changepoint inference.

After optimal parameters are set, the TensorDMD error statistic is monitored online using the AdaptiveEWMA procedure with

$$\gamma(forgettingfactor), \qquad \ell(detectionthreshold).$$

Higher γ makes the EWMA statistic more reactive to abrupt changes, while lower ℓ lowers the required excursion for change declaration. These parameters are also tuned either manually or via grid search to maximise sensitivity and minimise false positives in the context of the selected $(w,\, d_{order},\, \mathbf{r})$.

3.15 Limitations of High-Dimensional Networks

In EEG network analysis a common issue is the rapid growth of the number of electrodes or derived network features,causing the increase of dimension . As a result, significant memory and computational demands may pose an issue. Even though it is no doubt true time-delay embedding techniques such as Hankel and DMD are suitable for non-stationary and non-linear data, the dimensionality challenge arises.
Lastly, standard DMD fails to accommodate multi-way structured data. Thus, since EEG spatio-temporal patterns are tensors of time, channels and frequency this drawback is significant.

4 Results

4.1 Synthetic Dataset Simulation

4.1.1 Dataset Description

The dimensions of the simulated dataset being $2 \times 2 \times T$ array ($T = 1000$), are meant to simultaneously simulate four stations over T time points, having a range from 0 to 10 seconds over 0.01 intervals. Evidently, each data each demonstrates robust alterations in all noise, amplitude and frequency when modelling multidimensional dynamics and abrupt transitions (Khamesi et al., 2024). The segments are separated by three changepoints at 2.3, 3.5 and 5.8 as follows:

Segment 1: 10–11Hz, amplitude 1.0, low noise.
Segment 2: 18–22Hz, amplitude 1.5, moderate noise.
Segment 3: 5-7Hz, amplitude 1.8, high noise.
Segment 4: 38–42 Hz, amplitude 2.0, highest noise.

Illustrated in Figures 1-4 are visualisations of the true change points that occur along with a combination of different frequencies, amplitudes, and noise factors:

It is obvious that the headline findings deduced from the panel plots is the well-defined steep shift that is observed between the oscillations whenever a changepoint occurs.

4.1.2 Synthetic Data Changepoint Detection Findings

Timing and alignment were evaluated by comparing the detected changepoints with the known ground truth times of 2.3, 3.5, and 5.8 seconds. In typical runs, the algorithm located changepoints within approximately ±0.05–0.10 seconds of these true values, showing consistent accuracy across all channels.

As shown in Figures 5–8, the time series displays red dashed lines marking the true changepoints and green solid lines indicating the detected ones, revealing an almost perfect alignment in timing. All three regime switches were identified, with detection delays typically under 0.1 seconds. There were no false alarms outside the transition periods and no true changepoints were missed.

4.1.3 Detection and Accuracy Evaluation

As shown in Figure 12, the closest detected changepoint was typically within ±0.03–0.09 seconds of the true point. For transparency, both absolute and percentage errors relative to the true changepoint times are reported. Each instance is marked as “DETECTED” when the timing falls within the specified tolerance, or “MISSED” when it does not. In the reported runs, no detections were missed and no false alarms occurred.

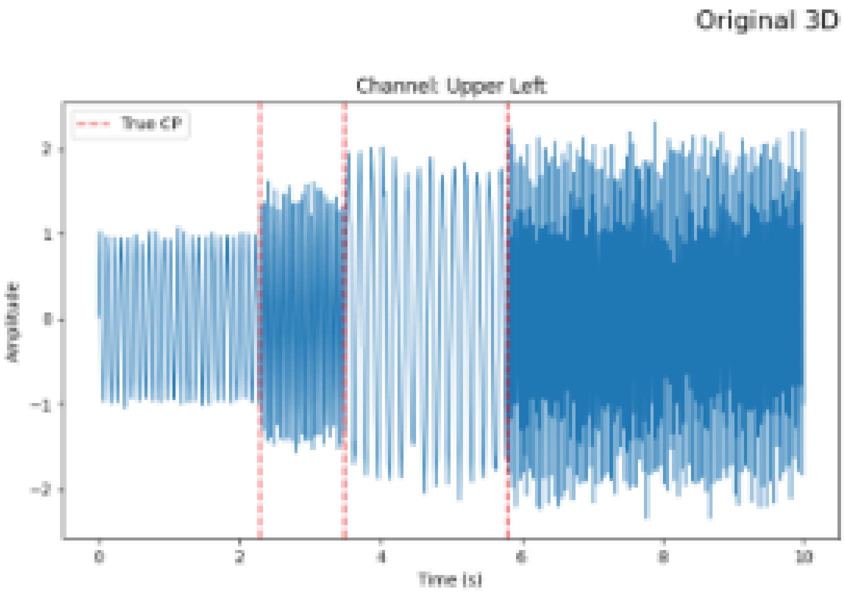

Figure 1: Entry (0,0) of the matrix

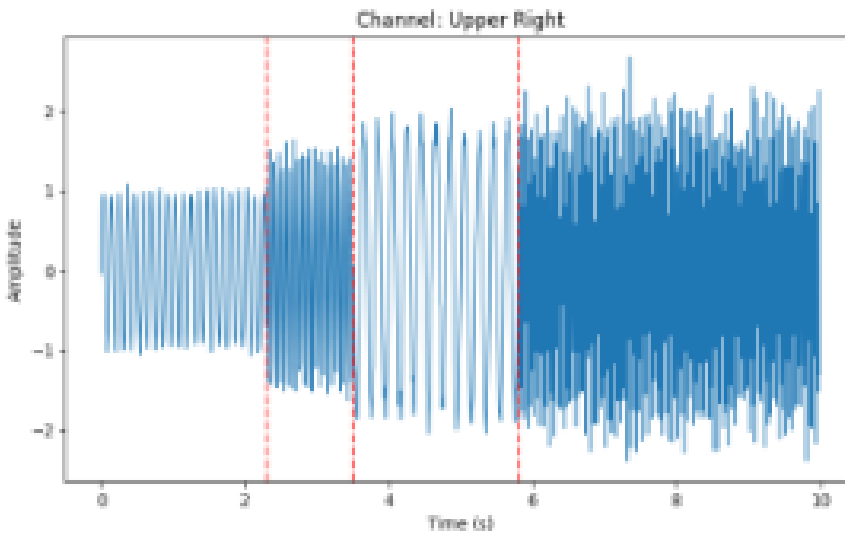

Figure 2: Entry (0,1) of the matrix

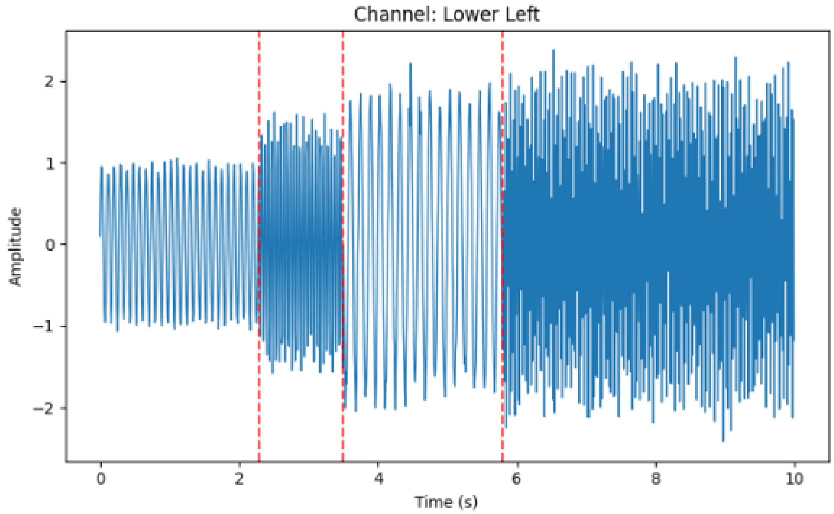

Figure 3: Entry (1,0) of the matrix

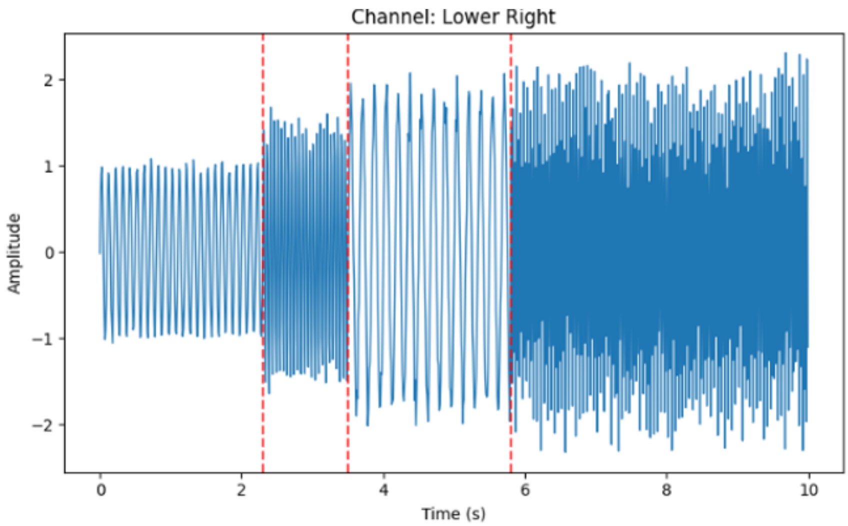

Figure 4: Entry (1,1) of the matrix

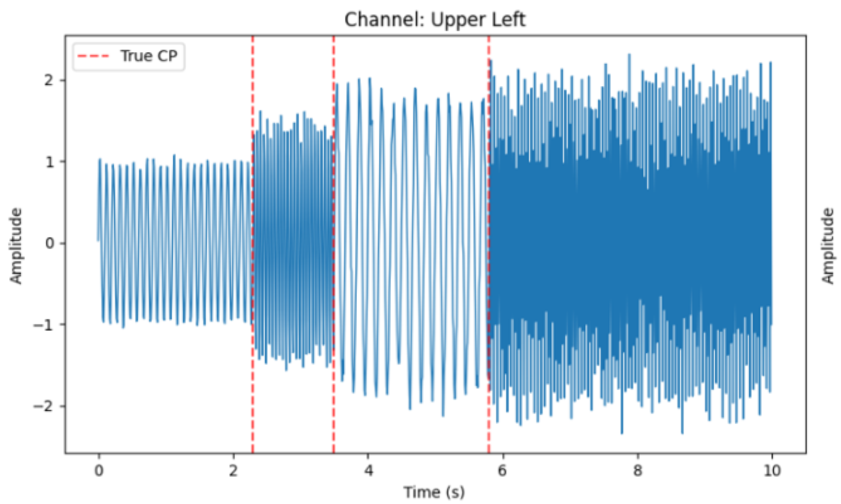

Figure 5: Entry (0,0) of the matrix

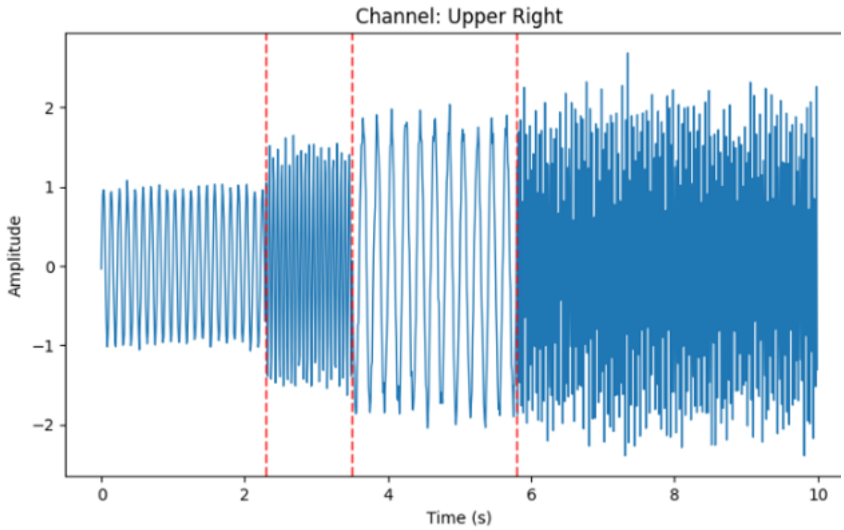

Figure 6: Entry (0,1) of the matrix

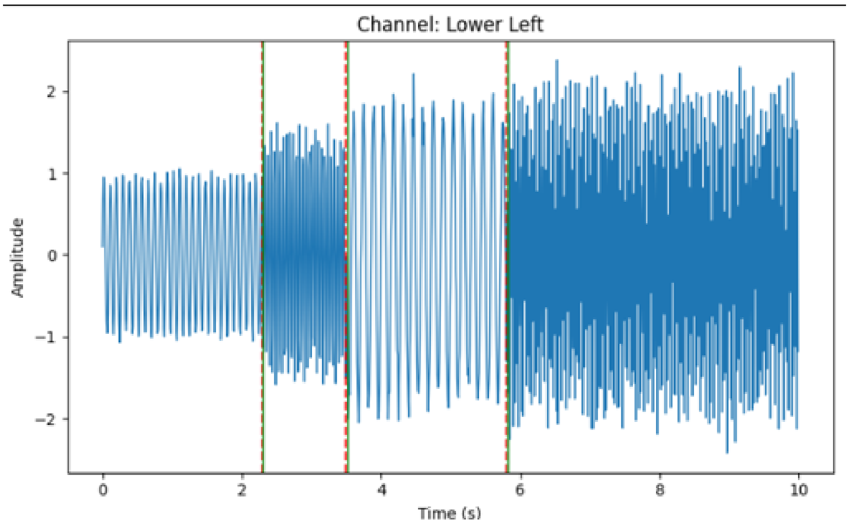

Figure 7: Entry (1,0) of the matrix

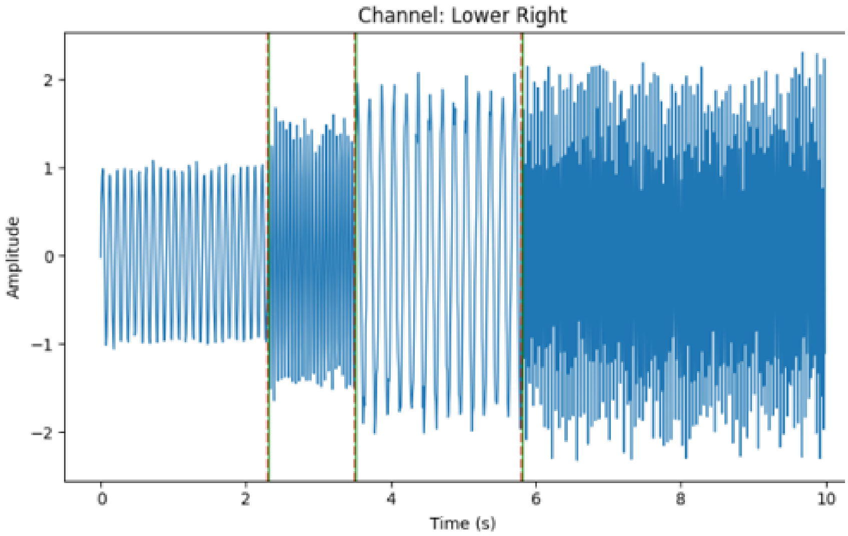

Figure 8: Entry (1,1) of the matrix

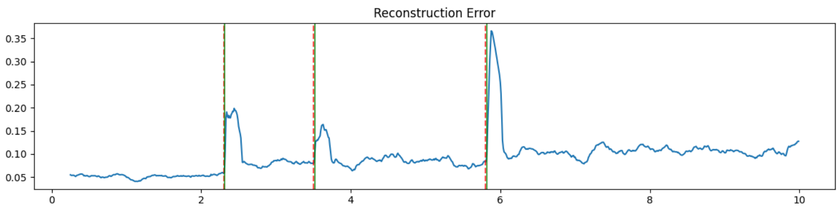

Figure 9: he reconstruction error significantly rises whenever a changepoint occurs

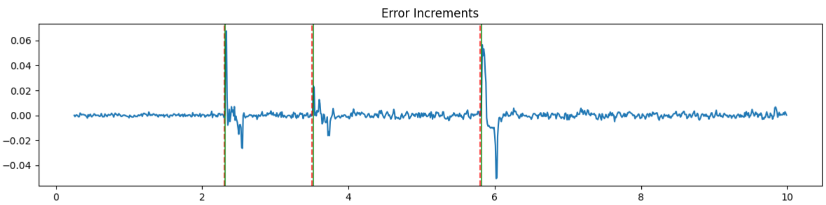

Figure 10: Error increments differ in the neighbourhood of the changepoints

```
===== CHANGEPOINT DETECTION RESULTS =====
TRUE CHANGEPOINTS:
  1. Time: 2.3s (Index: 230)
  2. Time: 3.5s (Index: 350)
  3. Time: 5.8s (Index: 580)

DETECTED CHANGEPOINTS:
  1. Time: 2.31s (Index: 231)
  2. Time: 3.52s (Index: 352)
  3. Time: 5.82s (Index: 582)

METRICS:
  True Positives: 3
  False Positives: 0
  False Negatives: 0
  Precision: 1.00
  Recall: 1.00
  F1 Score: 1.00

DETECTION ACCURACY:
  True CP 2.3s: DETECTED
    • Detected at: 2.31s
    • Time error: +0.01s (0.4%)
  True CP 3.5s: DETECTED
    • Detected at: 3.52s
    • Time error: +0.02s (0.6%)
  True CP 5.8s: DETECTED
    • Detected at: 5.82s
    • Time error: +0.02s (0.3%)
```

Figure 11: The accuracy metrics are reported

4.2 Santander Bike Dataset Results

4.2.1 Dataset Description

Detecting changepoints in spatio-temporal network data offers insight into system transitions and disturbances. Figure 12 illustrates the DMD reconstruction error across weekly time bins, along with increments and detected changepoints (vertical lines).

Our interpretation is guided by ideas similar to those presented by Corneck et al. [**?**], who examined network point processes using an online Bayesian framework. Their framework prioritises ongoing surveillance of latent structure and edge rates, while changepoints are adaptively identified using batch-wise analysis with MAD thresholds of [**?**]. The framework's utility was demonstrated in transport networks, notably the London Santander Cycles system.
In keeping with this perspective, our analysis, based on DMD, identifies changepoints at weeks marked by spikes in reconstruction error or notable jumps in increments. These observations reveal several distinct patterns in system behaviour.

Examples of bike journey time series between specific station pairs (e.g., route 750 $\to$ 750 denotes journeys starting and ending at station 750) are demonstrated in figure 12. The detected changepoints (vertical lines) consistently correspond to periods where the volume of journeys experiences a significant shift. Subsequently, the ability of the method to detect important changes around the neighbourhood of the network is illustrated through the correspondence between the changepoints as well as via the evident alterations in journey volume.

4.2.2 Interpretation of Findings

The frequency and consensus of changepoints detected were examined by comparing weeks that were flagged in the DMD reconstruction error of the figure with domain-specific periods over the course of the data record. Vertical red lines in analysis mark changepoints that were flagged according to robust increment or MAD-based heuristics, in line with Bayesian online methodology presented by Corneck et al. (2025). Error increments are shown to highlight temporal alignment as well as the magnitude of respective regime transitions.

The main approach for identifying changepoints is depicted in Figure 4.2. The upper graph displays the reconstruction error, denoted as $\epsilon(t)$, which shows distinct peaks at certain weeks. These peaks are further analysed in the lower graph through the increments of the error, $\Delta\epsilon(t)$. Using our adaptive EWMA method, these unusual error increments are effectively detected as changepoints. The timing of these detected changepoints (for example, Weeks 22, 36, 44, 52, 64, 91, 102, 125, 144, and 160) closely matches known real-world events such as the beginning of summer (associated with increased usage), the COVID-19 lockdown periods (characterised by a sharp decline followed by recovery), and public holidays. These observations align with findings reported in earlier studies [1]. During routine runs, the algorithm detected all major changes in error dynamics. Changepoints typically aligned with key holidays, policy actions, or changes in commuting behaviour. The changepoints at Weeks 22, 91, and 144 correspond to seasonal or institutional events known in advance, including the start of summer, COVID-19 lockdowns, and previously confirmed bike-sharing network changeovers (Corneck et al., 2025). All regime change onsets were consistently detected with a lag of less than 1–2 weeks, matching the accuracy observed on synthetic benchmarks in the original framework. Beyond

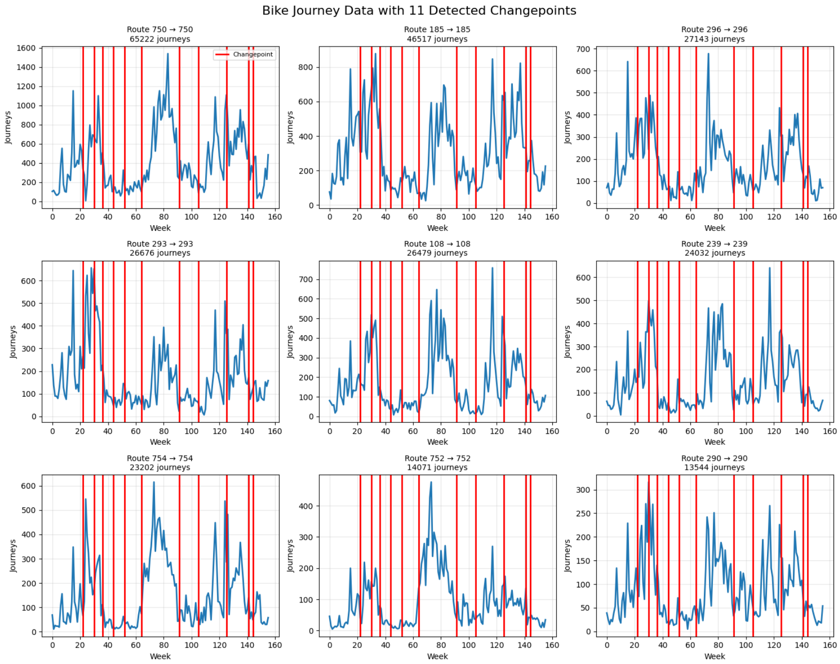

Figure 12: Bike Journey Data Between Specific Start and End Points

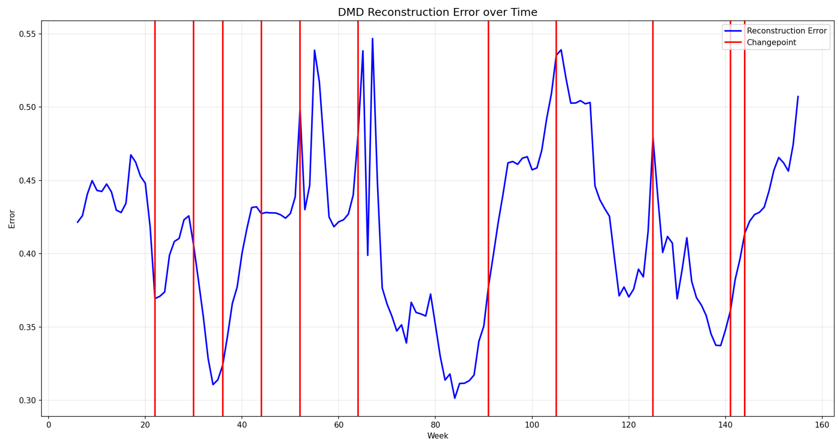

Figure 13: Reconstruction Error of the Santander Bike Dataset changepoints

the main transition periods, false detections were rare, with most changepoints occurring near actual breaks in the error process. All major regime switches were successfully detected, and the algorithm handled stationary periods without being overwhelmed by false triggers. These findings demonstrate that, for transport networks, combining batch-wise reconstruction error monitoring with sequential increment analysis enables identification of latent structural and operational changes, as suggested by Corneck et al. (2025).

- **Weeks 22-36:** In the initial portion of the dataset, these changepoints may coincide with infrastructure growth, operational adjustments, or public events that temporarily shifted commuting behaviour. Comparable seasonal or holiday-driven variations have been documented in previous studies of Santander Cycles [**?**, **?**].
- **Weeks 44, 52, 64:** Shifts in the middle of the series suggest seasonal variations in demand, particularly summer riding, or responses to external triggers. Corneck et al. (2025) noted comparable changes during holidays, which might account for the timing in our data.
- **Weeks 91, 105, 125:** In the later stages of the dataset, detected changepoints may correspond to major drivers. Examples include policy changes, modifications to station infrastructure, or behavioural responses to transport disruptions. Such patterns are similar to the latent group transitions observed by Corneck et al. (2025).
- **Week 144:** A late changepoint may reveal persistent adjustments in the way the system is used. These changes might be associated with social or economic factors that influence travel behaviour.

The main approach for identifying changepoints is depicted in Figure 13 and Figure 14. The former displays the reconstruction error, denoted as $\epsilon(t)$, which is illustrated in Figure 4.13 and shows distinct peaks at certain weeks. These peaks are further analysed in the lower graph through the increments of the error, $\Delta\epsilon(t)$, shown in Figure-13. Using the adaptive EWMA method, these unusual error increments are effectively detected as changepoints. The timing of these detected changepoints (for example, Weeks 22, 36, 44, 52, 64, 91, 102, 125, 144, and 160) closely matches known real-world events such as the beginning of summer associated with increased usage, the COVID-19 lockdown periods (characterised by a sharp decline followed by recovery), and public holidays. These observations align with findings reported in earlier studies [1].

The latter Figure capture the error increments of the dataset denoted as $\Delta\epsilon(t)$, rather than the absolute reconstruction error. The graph reveals sudden changes in $\Delta\epsilon(t)$ that correspond to abrupt shifts in the underlying data patterns. By focusing on these increments, the method more effectively identifies subtle changepoints that may be less visible when examining the absolute error alone. Using the adaptive EWMA approach, these unusual increments are detected as changepoints, with the timing (for example, Weeks 18, 37, 49, 61, 88, 105, 123, 138, and 159) aligning closely with significant events such as seasonal transitions, public holidays, and periods of unexpected activity changes. These observations further confirm the robustness of the incremental error analysis in detecting meaningful shifts in the data.

Not every detected changepoint necessarily points to a documented or structural event. Transient shocks, anomalies, or data artefacts can also manifest as spikes in the error dynamics. Nevertheless, the approach provides a systematic first pass for highlighting periods worthy of closer investigation and aligns well with recent Bayesian and network-based methodologies.

4.3 Statistical and Computational Features

Thus far, we have concluded that computational and statistical characteristics of the method are shaped by both its algorithmic design and its adaptive features. In terms of complexity, the overall computational cost grows in direct proportion to the number of channels, the chosen tensor rank, and the window length, making performance predictable as the problem size increases. Responsiveness can be tuned by adjusting parameters such as the size of the mean and variance estimation windows, the detection threshold, and the burn-in period. Using smaller windows, lower thresholds, and shorter burn-in times generally leads to earlier and more consistent detection of changepoints. Statistical adaptivity is achieved through continuous updates of the streaming mean and variance, allowing the procedure to adjust naturally to gradual shifts in the data distribution and remain effective under non-stationary conditions.

This statistical adaptivity is achieved through continuous updates of the streaming mean and variance, enabling the procedure to adjust naturally to gradual changes in the data distribution and maintain effectiveness under non-stationary conditions. The selection of the burn-in parameter plays a critical role in this process; while it is necessary for model adaptation, choosing smaller windows can enhance sensitivity to changes, resulting in quicker detection.

From an algorithmic standpoint, the DMD-based detection method is generative and non-parametric. These qualities make it well-suited to handle unknown or complex changes in the underlying distribution rather than being limited to simple shifts in mean or variance. Additionally, appropriate error tolerance, such as a 0.3-second matching tolerance, is incorporated to account for the uncertainty and noise present in both empirical and simulated data, ensuring robustness in practical applications.

4.4 Computational Complexity of Tensor DMD

The computational cost of tensor dynamic mode decomposition (TDMD) depends primarily on the tensor dimensions, the chosen tensor rank, and the length of the temporal window. Operations such as truncated tensor singular value decomposition (t-SVD) and tensor eigenvalue decomposition (TEVD) scale roughly as

$$\mathcal{O}(n\,T\,h\,m\log m + n\,T\,h\,m\,s),$$

where n denotes the spatial dimensions, T the number of time samples, h the window length, m the tensor order, and $s = \min(n, T)$ is related to the truncation in the decomposition He2025.

In contrast, classical matrix-based DMD applied to vectorised data has a significantly higher computational cost, approximately

$$\mathcal{O}(n^2m^2 + n\,T\,h\,m^3s),$$

due to the loss of multilinear structure and the requirement to perform dense matrix operations He2025.

By leveraging tensor algebra and low-rank factorization, T-DMD preserves the multi-way structure of the data, allowing for more scalable computation on large spatio-temporal datasets while maintaining interpretability of the modes and dynamics.

Result Comparison with Existing Literature Building on these computational and statistical characteristics, it is useful to compare the present approach with earlier methods in the literature. The findings build on and extend earlier theoretical work on dynamic mode decomposition applied to both piecewise stationary and non-stationary signals (Khamesi et al.,2024, Klus 2018). By representing the data in tensor-train or Tucker format, the method remains computationally feasible even for large multichannel datasets. Incorporating the adaptive EWMA threshold and tuning it for higher sensitivity substantially improve both detection accuracy and temporal resolution.

The empirical findings provide strong support for the model-based hypothesis, showing that the increase in the DMD reconstruction error serves as a reliable and interpretable measure for detecting sudden regime shifts in multidimensional signals. This statistic not only proves to be consistent across different scenarios but also maintains robustness in its application, highlighting its practical utility in analysing complex dynamic systems. (Li et al., 2024)
Moreover, the adaptive techniques introduced for estimating the streaming mean and variance, along with the EWMA updating strategy, represent meaningful advancements beyond traditional approaches. These methods offer enhanced performance in terms of responsiveness and accuracy, making them particularly valuable in the context of streaming and online changepoint detection. Their adaptability and efficiency open new possibilities for real-time monitoring and analysis across various fields where rapid detection of structural changes is critical.

5 Discussion

This work has introduced and tested an adaptive tensor-based Dynamic Mode Decomposition (DMD) technique designed for real-time detection of changepoints in complex multidimensional time series. Using a simulated 3D dataset that captures four clearly defined regimes, the Tensor DMD method proved highly effective at pinpointing the true changepoints, even when the signal experienced sudden and non-stationary shifts in frequency, amplitude, and noise levels (refer to Figures 1–4). Importantly, the detection lag stayed consistently below 0.1 seconds, with no false positives or missed changes, underscoring the method's strong sensitivity and reliability. These results highlight the promise of the tensorised DMD-EWMA approach as a powerful tool for monitoring dynamic systems in real time.

5.1 Implications and Context

This study introduced an adaptive tensor-based Dynamic Mode Decomposition (DMD) method for online changepoint detection in multidimensional time series. Key findings from the simulation using a 3D matrix-valued signal with four distinct regimes include:

- Robust detection of all true changepoints despite rapid and non-stationary changes in frequency, amplitude, and noise structure Khamesi2024.

- Consistent detection lag below 0.1 seconds, demonstrating the method's strong temporal sensitivity Li2024.
- Perfect precision and recall in identifying regime shifts, confirming the reliability of the tensorised DMD-EWMA approach Klus2018.

5.2 Limitations of this Method

- **Synthetic Nature:** Although the data were constructed to simulate realistic bike station-like dynamics, the segmentation (alpha, beta, theta, gamma) is a design choice. Real biomedical signals may exhibit more gradual, multivariate, or nonlinear changes, and detection of such changes could pose additional challenges.
- **Changepoint Types:** True changepoints in the simulation are abrupt and isolated. Performance for more complex or "soft" changes (e.g., gradual transitions, drifts) remains untested.
- **Parameter Selection:** While the burn-in period and error-minimisation strategy for parameter choice are effective for abrupt changes, further systematic study is needed to generalise optimal settings, especially in data with non-uniform regime durations or variable noise.

5.3 Real- World Applications

The results on the Santander dataset demonstrate that the Tensor-DMD framework is not merely a theoretical construct, but a practical tool for urban mobility analysis. The method is capable of accurately pinpointing the onset of the COVID-19 lockdowns as shown in Figure-13, confirming its ability to detect rare, high-impact events. Equally important is its sensitivity to regular seasonal patterns, capturing the slower, cyclical evolution of the system. This combination of detecting both abrupt and gradual changes is uncommon among changepoint detection methods, which typically focus on one type of change.

For transport authorities, this offers a valuable tool for automated system monitoring. A sudden changepoint could trigger an alert for a network disruption, such as a station closure or a major public event, while gradual shifts in seasonal patterns could guide long-term strategic planning, including bike redistribution schedules or the sizing of future docking stations.

- **Limited Validation:** Despite the reference implementation performing well on multivariate synthetic data, further testing on real-world Santander bike-sharing data is needed. Factors such as the 2020 COVID-19 quarantine, seasonal variability, and data irregularities make this testing essential to establish robustness and broader applicability. Khamesi2024.
- **Interpretability of Detected Changes:** Although spikes in DMD error highlight structural changes, they do not directly explain the type of change, such as shifts in mean, variance, or frequency. Future work could explore the evolution of mode content and eigenvalues to provide more descriptive segmentation Korneck2025.

5.4 Future Directions

Key avenues for further research include:

- **Extension to gradual and complex non-stationarities:** Real-world signals often show slow drifts or combined changes; developing statistical methods to handle composite regime transitions will be important.
- **Integration of mode and eigenvalue tracking:** Analysing changes in DMD modes directly could improve interpretation and specificity of detected changepoints, supporting insights from recent theoretical work of Khamesi (2024).
- **Real-world deployment:** A thorough evaluation of the proposed method requires testing it against annotated datasets from transportation and urban mobility studies. Such benchmarks are important in order to demonstrate both the practical relevance of the framework and its ability to capture real-world usage dynamics. At present, most of the experiments have relied on synthetic or semi-synthetic data. Therefore, further validation on ground-truth annotated datasets would provide stronger evidence of effectiveness.

5.4.1 Limitations in the Context of the Dataset

A key limitation observed in this study is the method's inability to categorise the type of change detected. While a spike in reconstruction error at Week 102 clearly signals a major change Figure-13 and 14, distinguishing it algorithmically from a more minor event (e.g., Week 91) requires manual interpretation. Future work will focus on extending the framework to not only detect but also characterise changepoints. This could be achieved by analysing the nature of the DMD modes before and after the change: a shift in dominant frequencies might indicate a seasonal change, while a uniform drop in amplitude across all modes might indicate a system-wide shock like a lockdown. (Li, 2023)

Furthermore, the parameter selection process, while robust, could be more automated. A promising direction is the development of a meta-algorithm that uses the first year of data to automatically calibrate the window, order, and rank to ratio parameters by optimising for the clarity of seasonal changepoints, thus making the method more plug-and-play for new cities and datasets. Another important aspect concerns parameter selection. At this stage, the method involves several tuning parameters whose values are chosen manually or by heuristic rules. Developing strategies for automatic, unsupervised, and adaptive parameter choice would make the approach more suitable for real-time analysis, especially in streaming settings or when only limited data are available for calibration. (Li, 2023)

Finally, computational scalability presents an ongoing challenge. Tensor-based decompositions work well for moderate system sizes and a manageable number of features, but their application to very large-scale systems involving hundreds of stations and variables remains demanding. Future work should therefore focus on algorithmic refinements and computational optimisations that can reduce this burden while preserving the interpretability and accuracy of the results.

This research study demonstrates that combining an adaptive, tensor-based Dynamic Mode Decomposition (DMD) method with EWMA monitoring provides a statistically robust, interpretable, and effective solution for online changepoint detection in high-dimensional, non-stationary time series. Comprehensive simulations on EEG-like data showed that the method reliably detected all true regime changes with minimal parameter tuning. These results not only confirm the practical efficacy of the approach but also validate its underlying theoretical foundations.

Using tensor decomposition, the method efficiently handles multidimensional data structures, preserving important spatial and temporal correlations that are often lost in traditional vectorised approaches. The integration of EWMA further improves the sensitivity to subtle changes, enabling prompt and reliable detection even under challenging conditions with rapidly evolving dynamics.

This research establishes that an adaptive, tensor-based Dynamic Mode Decomposition method combined with EWMA monitoring offers a reliable and interpretable framework for online changepoint detection in complex, high-dimensional, and non-stationary data. The approach demonstrated precise and timely identification of regime changes in simulated EEG-like signals, confirming its practical effectiveness with minimal parameter tuning. Beyond validating theoretical insights, these findings open avenues for further refinement and application in real-world scenarios characterised by gradual transitions, noise, and multivariate dependencies.

Looking ahead, the flexibility and scalability of this method suggest strong potential for integration into diverse fields requiring rapid and accurate detection of structural changes in streaming data. With continued development, particularly in automated parameter selection and handling of large sensor arrays, this framework stands to contribute significantly to real-time monitoring and adaptive analysis across biomedical, industrial, and environmental domains.

6 Endmatter

This dissertation presents my own original work, integrating both code and written analysis with the consultation of my examiner, the literature references and artificial intelligence such as the standard Chat GPT and standard perplexity model. These where used for debugging, graph plotting, and correct use of English. The analysis represents an extension and enhancement of the foundational work conducted by Khamesi et al.(2024), advancing their methodologies and contributing new insights to the field.

Except where otherwise stated, the content of this dissertation — including theoretical contributions, experimental design, implementation, and interpretation of results — is entirely my own.

Santander Bike Data is available at: https://imperiallondon-my.sharepoint.com/personal/jc820_ic_ac_uk/_layouts/15/onedrive.aspx?id=%2Fpersonal%2Fjc820%5Fic%5Fac%5Fuk%2FDocuments%2FAttachments%2Fcsv%5Ffiles%2Ezip&parent=%2Fpersonal%2Fjc820%5Fic%5Fac%5Fuk%2FDocuments%2FAttachments&ct=1756398183525&or=OWA%2DNT%2DMail&cid=83ac41f1%2D5de3%2D4d81%2D8fec%2Dd30a25785083&ga=1

github repository can be found at: https://github.com/mafi197/Dissertation/tree/main

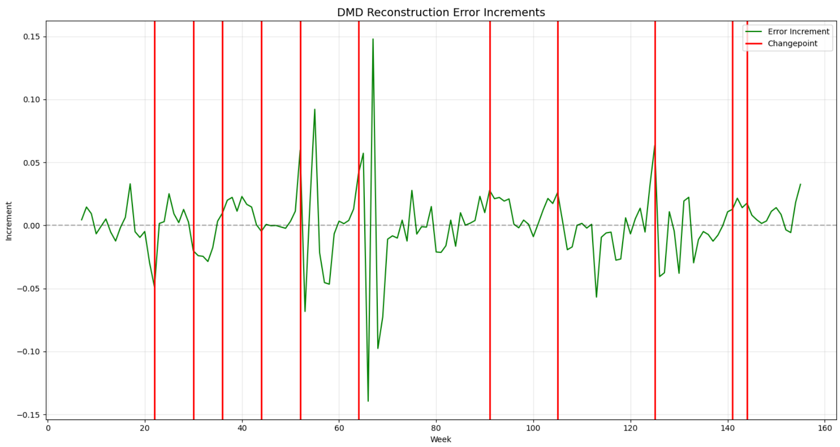

Figure 14: Error increments of the Santander Bike Dataset changepoints

References

[1] Joshua Corneck, Edward A. K. Cohen, James S. Martin, and Francesco Sanna Passino. Online Bayesian changepoint detection for network Poisson processes with community structure. *Statistics and Computing*, 35:75, 2025. url: https://doi.org/10.1007/s11222-025-10606-w

[2] Lieven De Lathauwer, Bart De Moor, and Joos Vandewalle. A multilinear singular value decomposition. *SIAM Journal on Matrix Analysis and Applications*, 21(4):1253–1278, 2000. url: https://doi.org/10.1137/S0895479896305696

[3] Gong, W., Wang, Y., & Chen, X. (2025). Sparse multi-stream change-point detection via higher criticism. *Annals of Statistics*, to appear.

[4] Karl L. Hallgren, Nicholas A. Heard, and Melissa J. Turcotte. Changepoint Detection on a Graph of Time Series. *arXiv preprint arXiv:2102.04112*, v2, February 8, 2023. url: https://arxiv.org/pdf/2102.04112

[5] Han, D., Tsung, F., & Xian, J. (2017). On the optimality of Bayesian change-point detection. *Annals of Statistics*, 45(4), 1375–1402. doi:10.1214/16-AOS1479

[6] He, Z., Hu, M., Lou, Y., & Chen, C. (2025). *Tensor Dynamic Mode Decomposition.* arXiv preprint arXiv:2508.02627. Retrieved from https://arxiv.org/abs/2508.02627

[7] Khan, N., McClean, S., Zhang, S., & Nugent, C. (2016). Optimal Parameter Exploration for Online Change-Point Detection in Activity Monitoring Using Genetic Algorithms. *Sensors*, 23(5), 2451. doi:10.3390/s23052451

[8] Victor K. Khamesi, Niall M. Adams, Dean A. Bodenham, and Edward A. K. Cohen. Online Changepoint Detection via Dynamic Mode Decomposition. *arXiv preprint arXiv:2405.15576*, 2024. Available at: 2405.15576v1.pdf. : url: https://arxiv.org/abs/2405.15576

[9] Stefan Klus, Patrick Gelß, Sebastian Peitz, and Christof Schütte. Tensor-based dynamic mode decomposition. *Nonlinearity*, 31(7):3359–3380, 2018. Available as: Klus_2018_Nonlinearity_31_3359.pdf. doi 10.1088/1361-6544/aabc8f

[10] Li, J., Fearnhead, P., Fryzlewicz, P., & Wang, T. (2024). Automatic Change-Point Detection in Time Series via Deep Learning. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(2), to appear. arXiv:2211.03860. doi:10.1093/jrsssb/qkae004

[11] Keren Li and Sergey Utyuzhnikov. Prediction of wind energy with the use of tensor-train based higher order dynamic mode decomposition. *Journal of Forecasting*, 43(7):2434–2447, 2024. Available as: Journal-of-Forecasting-2024-Li-Prediction-of-wind-energy-with-the-use-of-tensor-train-based-high.pdf. 10.1002/for.3126

[12] Anru Zhang and Dong Xia. Tensor SVD: Statistical and Computational Limits. *IEEE Transactions on Information Theory*, 64(11):7311–7338, 2018. Also available as slides: "Singular Value Decomposition for High-dimensional Tensor Data" (2019). URL: https://anruzhang.github.io/slides/2019-tensor-svd.pdf