Enhancing Structural Anomaly Detection Using a Bounded Autoregressive Component

Zhanwen Xin
Department of Civil, Geologic and Mining Engineering
Polytechnique Montreal
Montreal, Canada
zhanwen.xin@ploymtl.ca

James-A. Goulet
Department of Civil, Geologic and Mining Engineering
Polytechnique Montreal
Montreal, Canada
james.goulet@polymtl.ca

Abstract

Structural Health Monitoring has the potential to enhance the safety and serviceability of our aging infrastructures by detecting anomalies at an early stage. Bayesian Dynamic Linear Models (BDLM) have been shown to be effective at detecting anomalies by extracting structural patterns and latent variables from complex and noisy time series. However, the autoregressive component modelling the stationary prediction errors in most BDLM has a tendency to wrongfully capture patterns that should be attributed to anomalies, and thus hinders their detectability. This paper proposes a new bounded autoregressive (`BAR`) component, which imposes constraints on the autoregressive latent process with a new mixture Rectified Linear activation Unit. The `BAR` component is probabilistically verified on synthetic data using a new $F1t$ metric, and is validated using real observations collected on a bridge and on a dam located in Canada. The experimental results demonstrate that the `BAR` model surpasses the performance of the existing autoregressive component with (1) an improved accuracy at estimating hidden states, (2) an early detection of anomalies, (3) a capacity to detect smaller anomaly magnitudes, and (4) the ability to control the tradeoff between the anomaly detectability and the false alarm rate.

1 Introduction

Structural Health Monitoring (SHM) plays a key role in ensuring the safety and serviceability for a growing number of aging infrastructures [1, 2]. Two primary methodologies, physics-based and data-driven, are typically used in SHM. The physics-based approaches rely on the prediction from a numerical model (e.g., Finite Element) calibrated by measurements obtained on scaled [3] or real [4] structures. Data-driven methods leverage statistical or machine learning methods [5, 6] to analyze the data from simulations [7], visual inspections [8] or using sensors [9, 10], without requiring a physics-based model of the structure's behaviour. Data-driven approaches are typically more scalable [11], and they are capable of handling complex relations between structural responses and environmental conditions [12]. A main challenge is their limited interpretability as they rely on patterns and correlations, without explicitly incorporating physical principles.

Early anomaly detection and the prevention of deterioration are the backbone of SHM owing to their potential benefits in enhancing both the safety and economic viability through predictive maintenance. In data science, an *anomaly* is defined as unexpected data, which could be a single observation or a complex pattern, significantly different from the majority of instances [13]. In the case of SHM, anomalies are constrained by the physics of the deterioration processes, where they usually develop gradually over time such that it is difficult to identify them visually. Anomaly detection has been the focus of recent SHM research which has explored a range of machine learning algorithms [14]. Among them, Principal Component Analysis (PCA) [15, 16], Autoencoder [17, 18], Kalman Filter (KF) [10], Long-Short-Term Memory (LSTM) Network [19, 20], Generative Adversarial

Networks (GAN) [18], attention mechanism [21], reinforcement learning [22, 23], etc., have been used for applications on laboratory-scale and real-life case studies. However, a common factor is that these studies mainly focus on time series with large anomalies showing obvious pattern changes, whereas their performance is evaluated in a deterministic setup. The probabilistic evaluation of the anomaly detectability and the false alarm rate of a model under different types of anomalies is key for scaling anomaly detection approaches to a large number of applications, yet it is seldom investigated. Khazaeli et al. [22] have already started exploring these aspects by trading off the detectability of anomalies with the false alarm rate by using different reward functions in a reinforcement learning framework.

Anomaly detection involves epistemic and aleatory uncertainties, which may originate from the model structure selection, parameters estimation, missing dependencies between observations, model prediction errors, etc. Bayesian Dynamic Linear Models (BDLM) consider both types of uncertainties while using the Switching Kalman Filter (SKF) to probabilistically detect anomalies in time series [24, 25, 26] and it can analytically infer interpretable hidden states in a Bayesian and online way. In the BDLM-SKF approach proposed by Nguyen and Goulet [10], the probability of occurrence of an anomaly is represented by a posterior predictive probability for different regimes based on a Gaussian mixture reduction approach [27]. This model incorporates a stationary autoregressive (`AR`) model to capture residuals that cannot be explained by the irreversible baseline components and the reversible periodic ones [28]. It has been observed that when the autoregressive coefficient approaches one, the `AR` component becomes increasingly prone to capturing trends that should be attributed to the baseline or periodic components. This thus hinders the regime switching capacity because the pattern from a non-stationary regime may be hidden under the `AR` component due to its undesired trend-capturing property.

In this paper, we propose a new *bounded autoregressive* (`BAR`) component, which imposes constraints on the estimation of an `AR` hidden state while remaining compatible with the BDLM framework (BDLM-BAR). For that purpose, a *mixture Rectified Linear activation Unit* (mReLU) is developed to approximate the distribution of the `BAR` hidden state, and a modified $F1$ score is put forward for the probabilistic evaluation of the new `BAR` component. In general, the existing BDLM-SKF method and our current extension are intended to allow infrastructure owners to monitor their structures' health conditions using data measured by large-scale sensors without human interventions. The `BAR` component proposed in this article can further improves the hidden states' accuracy, reduces the false alarm rate, detects smaller anomalies and detect them earlier as demonstrated by our experiments, where the BDLM-BAR is verified on synthetic data, as well as validated using real observations collected on a bridge and on a dam located in Canada.

This paper is structured as follow. Section 2 discusses the characteristics of time series anomaly detection in terms of anomaly type and evaluation metric. Section 3 presents the existing BDLM framework with emphasis on its residual components. Section 4 proposes the formulation for the mReLU method and the `BAR` component, followed by three sets of experimental results on synthetic data, as well as data measured on a bridge and on a dam in Section 5. Finally, conclusions on the performance of the newly devised `BAR` component are drawn in Section 6.

2 Time series anomaly detection in SHM

This section presents six common types of anomalies in time series including shifts in local acceleration, trend or level for describing the deterioration processes in infrastructure. The criteria for detecting these anomalies is a confusion matrix adapted to the context of SHM and a new metric named $F1t$, which expands on the existing $F1$ score by considering the delay in the detection time.

2.1 Anomaly types

Anomalies in time series are typically categorized into *point-based* and *sequence-based* [29]. Point-based anomalies refer to a single data point that strongly deviates from a sequence within a time series. It can be further divided into a point anomaly which is an outlier compared to the complete time series, and a contextual anomaly as an outlier with respect to a subset of data within a time series. Unlike the point-based anomalies, sequence-based anomalies are defined as a sequence of data that do not follow a typical pattern from the past. Goswami [30] et al. categorized some of the sequence-based anomalies including noise anomalies where the noise pattern changes.

In SHM, the deterioration processes taking place on structures, such as corrosion, crack opening,

foundation settlements, etc, typically span over long periods from months to years so that point-based anomalies are not representative of their timescale. Furthermore, the point-based and noise anomalies are often associated with sensor malfunction or external environment changes, instead of a structural deterioration. To detect a state change in slow-deteriorating structures, this study focuses on the anomalies involving changes of time-series baselines [10] such as (1) a *local acceleration* (LA) anomaly changes time series from a constant speed regime to a constant acceleration one; (2) a *local trend* (LT) anomaly shifts the time series from a constant speed to a new speed; and (3) a *local level* (LL) anomaly corresponds to a jump from one local level to another. Note that these three anomalies suppose that the normal regime is having either a constant or zero speed. The early detection of these irreversible baseline changes could enable triggering alarms signalling the process of structural deteriorations before they can be visually observed. Figure 1 shows a visual representation of the above-mentioned anomalies.

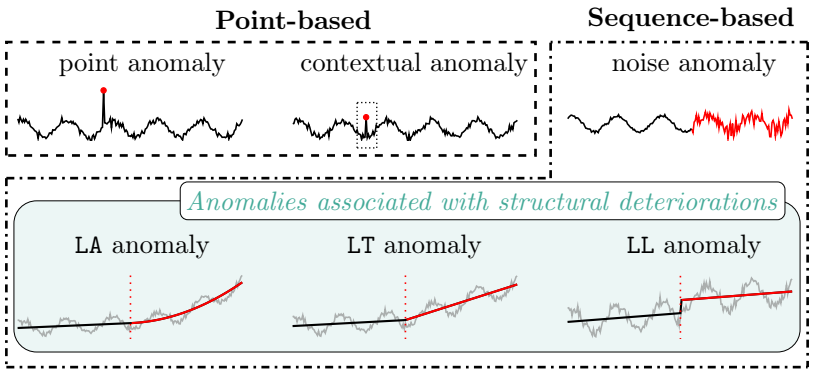


Figure 1: Visual representation for six types of anomalies in time series. The point, contextual, and noise anomalies are summarized in the literatures [29, 30]. Baseline-related anomalies including LA, LT, LL are the focus of our research as they are associated with potential structural deteriorations progressively developing over months or even years.

2.2 Evaluation metrics

Detecting anomalies on time series is analogous to a classification problem as both tasks consist in predicting the correct label for a given observation. Metrics commonly used for evaluating classification models, such as the confusion matrix and the $F1$ score [31], can be adapted to the anomaly detection task.

A typical confusion matrix adapted to the context of SHM is presented in Table 1. Given an observation, when the model prediction matches the ground truth, the corresponding prediction is evaluated as either a *true positive* (TP) or *true negative* (TN) depending on whether an anomaly is present or not. Conversely, if the model prediction is incorrect, this prediction is labeled as either a *false negative* (FN) or *false positive* (FP). To evaluate the overall accuracy while considering the TP, FP and FN, we can rely on the $F1$ score that is defined following

$$F1 = \frac{2\,TP}{2\,TP + FP + FN}.$$

Table 1: A confusion matrix in classification problems and its analogue in anomaly detection. True positives (TP) are the instance that alarm is triggered with the presence of abnormal data. False positives (FP) are the false alarms in anomaly detection. False negatives (FN) correspond to the missed alarms. True negatives (TN) indicate that no alarm is triggered given that there is no anomaly.

		ground truth	
		+ (abnormal)	− (normal)
model prediction	+ (alarm)	TP	FP
	− (no alarm)	FN	TN

In the context of time series anomaly detection, it is not feasible to evaluate the anomaly detectability pointwise because it would be biased by the length of anomalies, where models gain more reward for detecting long anomalies rather than short ones [32]. Take the baseline-related anomalies in Figure 1 as an example, after an anomaly is introduced in a time series where hidden states are shifted, all the following observations are labeled as abnormal. A model could then go through more TP instances when applied to a longer abnormal segment and thus wrongfully achieve a higher $F1$ score. As a result, by using the existing $F1$ definition, the performance of the anomaly detection model is not only a function of its detectability, but it is also biased by the length of abnormal segment.
Instead of updating the confusion matrix for each observation, we propose to assess it on a *time-series-wise* basis, whose concept is demonstrated in Figure 2. Given a time series with an anomaly, e.g., a `LA` anomaly, if a model triggers an alarm before the anomaly starts, its prediction will be labeled as a FP and the FP count is incremented by 1. On the other hand, if an alarm is correctly triggered within the abnormal region without any false alarm prior to it, the TP count is incremented by 1. Yet, if no alarm is triggered throughout an entire time series subjected to an anomaly, or if an alarm is triggered after a pre-defined detection window, the model performance will be assessed as a missed alarm with 1 added to the FN count. The evaluation on a time series is terminated as soon as any of the FP, TP or FN count is updated.



(a) False alarm (b) True alarm (c) Missed alarm

Figure 2: Time-series-wise TP, FP and FN definition. FP is incremented by 1 when an alarm is triggered before the occurrence of anomalies. TP is incremented by 1 when an alarm is correctly triggered in the abnormal segment. FN is increased by 1 if an alarm is missed within the detection window. When any of these three scenarios happens, analysis on this time series terminates and confusion matrix is updated.

With such an evaluation process, the model's performance is independent from the length of abnormal time series because each one only has a maximum increment of 1 in either TP, FP, or FN. Nonetheless, that determination of TP counts ignores the time delay in detecting anomalies because alarms triggered at any time within the anomaly region have the same increment in TP. To factorize the detection time in the anomaly detectability, we propose to use a decay ratio λ that linearly penalizes the $F1$ score with an increase in the *detection time* Δ_t. This new metric is called the *F1 temporal* ($F1t$) score, such that $F1t = \lambda F1$.
Figure 3 shows the relation between λ and Δ_t, which is the average detection time for all the time series. λ is one if a model triggers an alarm before or at the moment when the anomaly starts, while it linearly decreases to zero as the required detection time reaches the end of a pre-defined detection window l_{dw}. This approach is aware that a detection time smaller than zero corresponds to a false alarm. Since false alarms are penalized through the FP in the calculation of $F1$, λ remains one for $\Delta_t < 0$ to avoid a double penalization. A model can achieve a maximum $F1t$ score of 1 if it detects all the anomalies at their exact starting time without triggering any false alarms. On the other hand, any delay in detection, false alarms or missing alarms will lead to a reduction in the $F1t$ score.

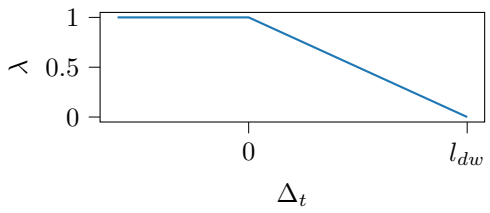


Figure 3: Decreasing λ with longer detection time. l_{dw} is the length of target detection window. λ drops to zero if the required time to detect an anomaly is longer than l_{dw}.

3 Bayesian Dynamic Linear Models (BDLM)

Bayesian Dynamic Linear Models (BDLM) have been shown to be effective for extracting structural patterns and latent variables from complex and noisy time series in SHM [12, 33]. This section focuses on one specific BDLM structure [34], under whose framework the BAR model is developed in Section 4.

3.1 Mathematical formulation of BDLM

The general formulation of BDLM involves a *prediction* and an *update step* [28], which are summarized by the Equations (1–4) with a focus on a pair of consecutive time steps $t-1$ and t under the Markov assumption. The *prediction step* includes a transition model that transfers the knowledge about the hidden states $\boldsymbol{x}$ from $t-1$ to t and an observation model that indicates which one is observable. The transition model is defined by

$$\boldsymbol{x}_t = \mathbf{A}\boldsymbol{x}_{t-1} + \boldsymbol{w}_t, \boldsymbol{W} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \tag{1}$$

where $\mathbf{A}$ is a transition matrix, $\boldsymbol{w}_t$ is a transition error that is a realization for a Gaussian distribution with mean zero and covariance $\mathbf{Q}$. This transition model is used to compute our prior about $\boldsymbol{x}_t$ using all the past data up to $t-1$ such that $f(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}) = \mathcal{N}(\boldsymbol{\mu}^{\boldsymbol{X}}_{t|t-1}, \boldsymbol{\Sigma}^{\boldsymbol{X}}_{t|t-1})$, where $\boldsymbol{y}_{1:t-1}$ are the observations from time 1 up to $t-1$. The observation model is described by

$$\boldsymbol{y}_t = \mathbf{F}\boldsymbol{x}_t + \boldsymbol{v}_t, \boldsymbol{V} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \tag{2}$$

where $\mathbf{F}$ is an observation matrix and $\boldsymbol{v}_t$ is an observation error that is a realization for a zero mean Gaussian distribution with covariance $\mathbf{R}$. Again, this observation model is used to compute our prior about $\boldsymbol{y}_t$ using all the data up to $t-1$ such that $f(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1}) = \mathcal{N}(\boldsymbol{\mu}^{\boldsymbol{Y}}_{t|t-1}, \boldsymbol{\Sigma}^{\boldsymbol{Y}}_{t|t-1})$.

The *update step* computes the posterior knowledge of hidden states at time t conditional on all observations $\boldsymbol{y}_1, \boldsymbol{y}_2, ..., \boldsymbol{y}_t$ following

$$\boldsymbol{\mu}^{\boldsymbol{X}}_{t|t} = \boldsymbol{\mu}^{\boldsymbol{X}}_{t|t-1} + \boldsymbol{\Sigma}^{\boldsymbol{XY}}_{t|t-1}(\boldsymbol{\Sigma}^{\boldsymbol{Y}}_{t|t-1})^{-1}(\boldsymbol{y}_t - \boldsymbol{\mu}^{\boldsymbol{Y}}_{t|t-1}), \tag{3}$$

$$\boldsymbol{\Sigma}^{\boldsymbol{X}}_{t|t} = \boldsymbol{\Sigma}^{\boldsymbol{X}}_{t|t-1} - \boldsymbol{\Sigma}^{\boldsymbol{XY}}_{t|t-1}(\boldsymbol{\Sigma}^{\boldsymbol{Y}}_{t|t-1})^{-1}\boldsymbol{\Sigma}^{\boldsymbol{YX}}_{t|t-1}, \tag{4}$$

where the subscript $t|t$ denotes the updated estimation at time t using all the observations until time t, and where the covariance matrix $\boldsymbol{\Sigma}^{\boldsymbol{XY}}_{t|t-1} = \boldsymbol{\Sigma}^{\boldsymbol{X}}_{t|t-1}\mathbf{F}^\intercal$.

3.2 Residual components in BDLM

BDLM are built using various generic components which can be categorized into (a) baseline components that describe the local level (LL), trend (LT), and acceleration (LA) of time series; (b) periodic components that capture the harmonic periodic (PD) or non-harmonic kernel regressive (KR) pattern, and (c) a residual component which represents the model error that cannot fit in any other explainable components. Residual components are key to ensuring the estimation accuracy of other hidden states [28]. A common residual component is the first order autoregressive model, denoted simply as AR. This model will be reviewed next.

3.2.1 Formulation of the autoregressive (AR) component

The matrices defining the transition and observation models of the AR component are

$$\mathbf{A}^{\mathtt{AR}} = \phi^{\mathtt{AR}}, \mathbf{F}^{\mathtt{AR}} = 1, \mathbf{Q}^{\mathtt{AR}} = (\sigma^{\mathtt{AR}})^2.$$

The transition matrix $\mathbf{A}^{\mathtt{AR}}$ is defined by the autoregression coefficient $\phi^{\mathtt{AR}}$, which represents the linear dependency between successive time steps and allows the AR model to predict the current hidden state from the past step. $\mathbf{F}^{\mathtt{AR}}$ is set to 1, indicating that $\boldsymbol{x}^{\mathtt{AR}}$ is directly observable. $\sigma^{\mathtt{AR}}$ is the standard deviation of process error, and the stationary standard deviation, $\sigma^{\mathtt{AR},0}$, measures the dispersion of the stationary AR process and is defined following

$$(\sigma^{\mathtt{AR},0})^2 = \frac{(\sigma^{\mathtt{AR}})^2}{1-(\phi^{\mathtt{AR}})^2}, \tag{5}$$

where $\phi^{\mathtt{AR}}$ is constrained in $[0, 1]$ in order to model stationary process errors.

3.2.2 Motivation for constraining the AR component

In BDLM, when there is a small progressive anomaly, or when $\phi^{\texttt{AR}}$ is optimized to a value close to 1, the AR component has a tendency to capture patterns that should not belong to the residual term. A drifting AR could delay the detection of anomalies through regime switches and hinder the hidden states' interpretability. A value of $\phi^{\texttt{AR}}$ approaching 1 can be attributed to

(a) modelling errors: When the baseline or periodic components are not correctly defined, the patterns of these components may be partially attributed to the AR component and eventually lead to a $\phi^{\texttt{AR}}$ value close to 1 to enable capturing the strong dependencies between residual values;

(b) small time steps (δ_t): When the time-step length is short, e.g., 1 hour, consecutive model predictions are likely to be similar. The model errors describing the discrepancy between the model predictions and the structural states are thus highly correlated. In order to capture such a high correlation, $\phi^{\texttt{AR}}$ typically approaches 1. On the other hand, with long time-step lengths, e.g., 1 month, the model errors become mostly independent from each other and $\phi^{\texttt{AR}}$ tends to be close to zero.

We will see in the next section how the shortcomings of the AR component can be overcome using constraints without negatively impacting the estimation of other components and while maintaining the analytical tractability of the BDLM-SKF approach.

4 Bounded autoregressive (BAR) component

This section introduces the bounded autoregressive (BAR) component for modelling constrained error processes in the BDLM framework. The moments of the BAR hidden states as well as its covariance with respect to other hidden states are approximated using a new probabilistic activation function called the Mixture Rectified Linear Unit (mReLU). The mReLU function is developed in Section 4.2 based on the existing ReLU activation function and the theory of Gaussian mixture reduction in Section 4.1. The formulations for the BDLM-BAR are presented in Section 4.3.

4.1 Lower- and upper-bounded ReLU and Gaussian mixture reduction

The rectified Linear Unit (ReLU) is one of the most popular activation functions in deep learning models [35, 36]. A ReLU function bounds its input values to the positive range by assigning zero to all negative input values, i.e., $\phi_r(x) = \max(0, x)$. With the superposition of a minimum and a maximum function, one can obtain an activation function that bounds input values to a closed interval. Similar to a RuLU6 [37] that bounds the input in [0, 6], a general formulation of a lower- and upper-bounded ReLU that symmetrically constrains input values between $-b$ and b is given by

$$\phi_r^{\pm}(x) = \min\left(\max\left(-b, x\right), b\right).$$

A Gaussian mixture combines multiple Gaussian distributions, each represented by $\left\{\pi_i, \mu_i, \sigma_i^2\right\}$, where π_i is the probability of each distribution, and μ_i and σ_i^2 denote the mean and variance. This approach enables calculating the exact moments for the combined distribution following

$$\tilde{\mu} = \sum_{i=1}^{n} \pi_i \mu_i, \tag{6}$$

$$\tilde{\sigma}^2 = \sum_{i=1}^{n} \pi_i \left(\sigma_i^2 + \left(\mu_i - \tilde{\mu}\right)^2\right). \tag{7}$$

The Gaussian mixture reduction [27] approximates this combined distribution by a Gaussian distribution using the moments defined by eq. (6) and eq. (7). Figure 4 presents examples of Gaussian mixture reduction for two and three components.

4.2 Mixture Rectified Linear activation Unit (mReLU)

The BAR component relies on a Gaussian mixture reduction between a truncated Gaussian and two real values obtained from a lower- and upper-bounded ReLU function. Figure 5 illustrates an example

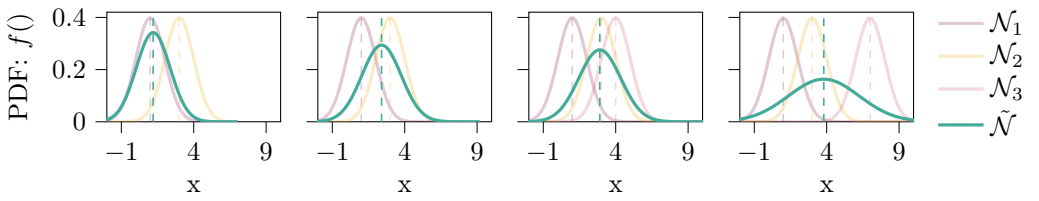


Figure 4: Four Gaussian mixture reduction examples. The first and the second plots exhibit the mixture $\tilde{\mathcal{N}}$ for two Gaussian components, which are $\mathcal{N}_1 = \{\pi_1 = 0.9, \mu_1 = 1, \sigma_1^2 = 1^2\}$, $\mathcal{N}_2 = \{\pi_2 = 0.1, \mu_2 = 3, \sigma_2^2 = 1^2\}$ for the first plot, and $\mathcal{N}_1 = \{\pi_1 = 0.3, \mu_1 = 1, \sigma_1^2 = 1^2\}$, $\mathcal{N}_2 = \{\pi_2 = 0.7, \mu_2 = 3, \sigma_2^2 = 1^2\}$ for the second plot. The third and the fourth plots show the mixtures $\tilde{\mathcal{N}}$ for three Gaussian components, which are $\mathcal{N}_1 = \{\pi_1 = 0.2, \mu_1 = 1, \sigma_1^2 = 1^2\}$, $\mathcal{N}_2 = \{\pi_2 = 0.5, \mu_2 = 3, \sigma_2^2 = 1^2\}$, and $\mathcal{N}_3 = \{\pi_3 = 0.3, \mu_3 = 4, \sigma_3^2 = 1^2\}$ for the third plot, and $\mathcal{N}_1 = \{\pi_1 = 0.2, \mu_1 = 1, \sigma_1^2 = 1^2\}$, $\mathcal{N}_2 = \{\pi_2 = 0.5, \mu_2 = 3, \sigma_2^2 = 1^2\}$, and $\mathcal{N}_3 = \{\pi_3 = 0.3, \mu_3 = 7, \sigma_3^2 = 1^2\}$ for the fourth plot.

of the mReLU transformation, where the input is a Gaussian distribution $f(x^{\texttt{AR}})$ and the output is approximated by another Gaussian distribution $f(x^{\texttt{BAR}})$ given $X^{\texttt{BAR}} = \phi_r^{\pm}(X^{\texttt{AR}})$. Starting from the plot on the second row, $f(x^{\texttt{AR}})$ is truncated into three segments by two boundary values $-b$ and b. By passing the truncated segments through a ReLU function, the intervals $[-\infty, -b]$ and $[b, +\infty]$ are collapsed to $-b$ and b, as shown in the second plot on the first row. These can be regraded as two Gaussian components with variances equal to zero, $\{\pi_1 = F_{X^{\texttt{AR}}}(-b), \mu_1 = -b, \sigma_1^2 = 0\}$ and $\{\pi_2 = F_{X^{\texttt{AR}}}(b), \mu_2 = b, \sigma_2^2 = 0\}$, following the same notation for the Gaussian mixture reduction mentioned in Section 4.1. The probability density function (PDF) for the middle interval $f(\tilde{x}^{\texttt{AR}})$ remains unchanged by the ReLU and is approximated by matching its moments to a Gaussian PDF illustrated by a solid line in the third plot on the first row, which acts as the third component $\{\pi_3 = F_{X^{\texttt{AR}}}(b) - F_{X^{\texttt{AR}}}(-b), \mu_3, \sigma_3^2\}$. The Gaussian mixture reduction represents the distribution of the BAR hidden state $f(x^{\texttt{BAR}})$ as shown in the rightmost plot.

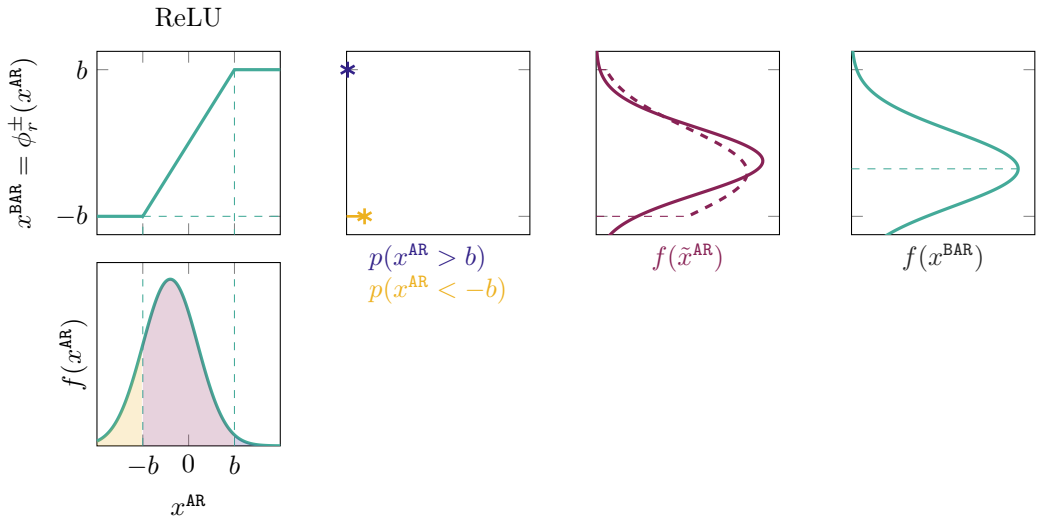


Figure 5: A mReLU transformation example in the context of BAR. The input distribution $f(x^{AR})$ has a mean value that is close to -b within the boundary and a variance that extends beyond the $\pm b$.

The calculations of the moments μ_3, σ_3, $\mu^{\texttt{BAR}}$, and $\sigma^{\texttt{BAR}}$, as well as the covariance between $X^{\texttt{BAR}}$ and other hidden states are carried in the mReLU function which can be summarized as

$$\left\{\mu^{\texttt{BAR}}, \sigma^{\texttt{BAR}}, \text{cov}\left(X^{\texttt{BAR}}, X^{\cdot}\right)\right\} = \text{mReLU}\left(\mu^{\texttt{AR}}, \sigma^{\texttt{AR}}, \text{cov}(X^{\texttt{AR}}, X^{\cdot}), \gamma\right), \tag{8}$$

where $\text{cov}\left(X^{\texttt{BAR}}, X^{\cdot}\right)$ is the covariance between BAR and other hidden states, $\text{cov}(X^{\texttt{AR}}, X^{\cdot})$ is the covariance between AR and other hidden states, and γ is a hyper-parameter that defines the boundaries

according to the autoregressive process's stationary standard deviation $\sigma^{\texttt{AR},0}$ as defined by eq. (5), so that the boundary value is

$$b = \gamma\sigma^{\texttt{AR},0} = \gamma\sqrt{\frac{\sigma_W^2}{1-(\phi^{\texttt{AR}})^2}}. \tag{9}$$

When the constant $\gamma = 2$, the bounded area corresponds to a 95% coverage region for the autoregressive process's stationary distribution.

The truncated Gaussian between $-b$ and b is approximated by a Gaussian PDF whose moments match those of a truncated Gaussian, for which the expected value is

$$\begin{aligned}
\alpha_{\texttt{L}} &= -\tfrac{b+\mu^{\texttt{AR}}}{\sigma^{\texttt{AR}}},\\
\alpha_{\texttt{U}} &= \tfrac{b-\mu^{\texttt{AR}}}{\sigma^{\texttt{AR}}},\\
\omega &= \Phi(\alpha_{\texttt{U}}) - \Phi(\alpha_{\texttt{L}}) = \Pr(-b < X^{\texttt{AR}} < b),\\
\beta &= \tfrac{\phi(\alpha_{\texttt{U}})-\phi(\alpha_{\texttt{L}})}{\omega},\\
\tilde{\mu}^{\texttt{AR}} &= \mu^{\texttt{AR}} - \beta\sigma^{\texttt{AR}},
\end{aligned}$$

the variance is

$$\begin{aligned}
\kappa &= 1 - \tfrac{\alpha_{\texttt{U}}\phi(\alpha_{\texttt{U}})-\alpha_{\texttt{L}}\phi(\alpha_{\texttt{L}})}{\omega} - \beta^2,\\
(\tilde{\sigma}^{\texttt{AR}})^2 &= \kappa(\sigma^{\texttt{AR}})^2,
\end{aligned}$$

and the covariance between $\tilde{X}^{\texttt{AR}}$ and another variable $X^{\cdot}$ is

$$\text{cov}(\tilde{X}^{\texttt{AR}}, X^{\cdot}) = \kappa^{1/2}\text{cov}(X^{\texttt{AR}}, X^{\cdot}).$$

A Gaussian mixture reduction is performed on the three Gaussian components, $\{\pi_1 = F_{X^{\texttt{AR}}}(-b), \mu_1 = -b, \sigma_1^2 = 0\}$, $\{\pi_2 = F_{X^{\texttt{AR}}}(b), \mu_2 = b, \sigma_2^2 = 0\}$ and $\{\pi_3 = F_{X^{\texttt{AR}}}(b) - F_{X^{\texttt{AR}}}(-b), \mu_3 = \tilde{\mu}^{\texttt{AR}}, \sigma_3^2 = (\tilde{\sigma}^{\texttt{AR}})^2\}$, to obtain the output of BAR hidden state

$$X^{\texttt{BAR}} = -b\Phi(\alpha_{\texttt{L}}) + \omega\tilde{X}^{\texttt{AR}} + b(1-\Phi(\alpha_{\texttt{U}})).$$

Following eq. (6), the expected value of $X^{\texttt{BAR}}$ is a mixture of the expected values of the three Gaussian components according to their probabilities so that

$$\mu^{\texttt{BAR}} = -b\Phi(\alpha_{\texttt{L}}) + \omega\tilde{\mu}^{\texttt{AR}} + b(1-\Phi(\alpha_{\texttt{U}})).$$

By setting $\sigma_1 = 0$ and $\sigma_2 = 0$ in eq. (7), the variance of $X^{\texttt{BAR}}$ is given by

$$(\sigma^{\texttt{BAR}})^2 = \omega(\tilde{\sigma}^{\texttt{AR}})^2 + \omega(\tilde{\mu}^{\texttt{AR}} - \mu^{\texttt{BAR}})^2 + \Phi(\alpha_{\texttt{L}})(b+\mu^{\texttt{BAR}})^2 + (1-\Phi(\alpha_{\texttt{U}}))(b-\mu^{\texttt{BAR}})^2,$$

and the covariance between $X^{\texttt{BAR}}$ and other hidden states $X^{\cdot}$ are

$$\begin{aligned}
\lambda &= (\omega\kappa)^{1/2},\\
\text{cov}(X^{\texttt{BAR}}, X^{\cdot}) &= \lambda\text{cov}(\tilde{X}^{\texttt{AR}}, X^{\cdot}).
\end{aligned}$$

4.3 BDLM-BAR formulation

This section outlines the integration of the BAR component within the BDLM framework. The model matrices for the BAR component are

$$\boldsymbol{x_t} = \begin{bmatrix} x^{\texttt{AR}} \\ x^{\texttt{BAR}} \end{bmatrix}, \mathbf{A} = \begin{bmatrix} \phi^{\texttt{AR}} & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{F} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\intercal}, \mathbf{Q} = \begin{bmatrix} (\sigma_W^{\texttt{AR}})^2 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{R} = \sigma_V^2.$$

The BAR approach only changes the transition model in the prediction step presented in Section 3.1 by bounding its autoregressive output using mReLU and by incorporating the bounded values to the

predicted hidden states. Following eq. (1), the moments for the output of the transition model can be calculated as

$$\tilde{\boldsymbol{\mu}}^{\boldsymbol{X}}_{t|t-1} = \mathbf{A}\boldsymbol{\mu}^{\boldsymbol{X}}_{t-1} = \begin{bmatrix} \phi^{\texttt{AR}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mu^{\texttt{AR}}_{t-1} \\ \mu^{\texttt{BAR}}_{t-1} \end{bmatrix} = \begin{bmatrix} \phi^{\texttt{AR}}\mu^{\texttt{AR}}_{t-1} \\ 0 \end{bmatrix}, \tag{10}$$

$$\begin{aligned} \tilde{\boldsymbol{\Sigma}}^{\boldsymbol{X}}_{t|t-1} &= \mathbf{A}\boldsymbol{\Sigma}^{\boldsymbol{X}}_{t-1}\mathbf{A}^{\intercal} + \mathbf{Q} \\ &= \begin{bmatrix} \phi^{\texttt{AR}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} (\sigma^{\texttt{AR}}_{t-1})^2 & \operatorname{cov}(X^{\texttt{AR}}_{t-1}, X^{\texttt{BAR}}_{t-1}) \\ \operatorname{cov}(X^{\texttt{AR}}_{t-1}, X^{\texttt{BAR}}_{t-1}) & (\sigma^{\texttt{BAR}}_{t-1})^2 \end{bmatrix} \begin{bmatrix} \phi^{\texttt{AR}} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} (\sigma^{\texttt{AR}}_{W})^2 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (\phi^{\texttt{AR}})^2(\sigma^{\texttt{AR}}_{t-1})^2 + (\sigma^{\texttt{AR}}_{W})^2 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{11}$$

Note that $X^{\texttt{BAR}}$ is not involved in the transition model as indicated by the components of the $\mathbf{A}$ matrix that are all 0 except in the AR position $[\mathbf{A}]_{1,1}$. On the other hand, $X^{\texttt{BAR}}$ is the only observable hidden state. This is because the BAR method constrains a variable that is analogous to the autoregressive component without modifying the original variable, whose entire distribution remains and is transferred through time steps. Without this procedure, the approximation errors caused by the truncation would accumulate at each time step, leading to a monotonically increasing bias and monotonically decreasing variance [38].

The moments of $X^{\texttt{BAR}}_t$ as well as its covariance with respect to $X^{\texttt{AR}}_t$ are computed by bounding the AR hidden state using the mReLU function defined in eq. (8). They are then added to eq. (10) and eq. (11) to obtain the predicted hidden states with constrained residuals, whose moments conditional on all the observations up to $t-1$ are defined by

$$\begin{aligned} \boldsymbol{\mu}^{\boldsymbol{X}}_{t|t-1} &= \tilde{\boldsymbol{\mu}}^{\boldsymbol{X}}_{t|t-1} + \begin{bmatrix} 0 \\ \mu^{\texttt{BAR}}_{t|t-1} \end{bmatrix} = \begin{bmatrix} \phi^{\texttt{AR}}\mu^{\texttt{AR}}_{t-1} \\ \mu^{\texttt{BAR}}_{t|t-1} \end{bmatrix}, \\ \boldsymbol{\Sigma}^{\boldsymbol{X}}_{t|t-1} &= \tilde{\boldsymbol{\Sigma}}^{\boldsymbol{X}}_{t|t-1} + \begin{bmatrix} 0 & \operatorname{cov}(X^{\texttt{BAR}}_t, X^{\texttt{AR}}_t) \\ \operatorname{cov}(X^{\texttt{BAR}}_t, X^{\texttt{AR}}_t) & (\sigma^{\texttt{BAR}}_{t|t-1})^2 \end{bmatrix} \\ &= \begin{bmatrix} (\phi^{\texttt{AR}})^2(\sigma^{\texttt{AR}}_{t-1})^2 + (\sigma^{\texttt{AR}}_{W})^2 & \operatorname{cov}(X^{\texttt{BAR}}_t, X^{\texttt{AR}}_t) \\ \operatorname{cov}(X^{\texttt{BAR}}_t, X^{\texttt{AR}}_t) & (\sigma^{\texttt{BAR}}_{t|t-1})^2 \end{bmatrix}, \end{aligned}$$

and which are then passed to the observation model described by eq. (2), and finally to the update step following eq. (3) and eq. (4) as presented in Section 3.1. The process of integrating the BAR component in the BDLM framework is illustrated by the flowchart in Figure 6 with the modified steps highlighted by dashed lines.

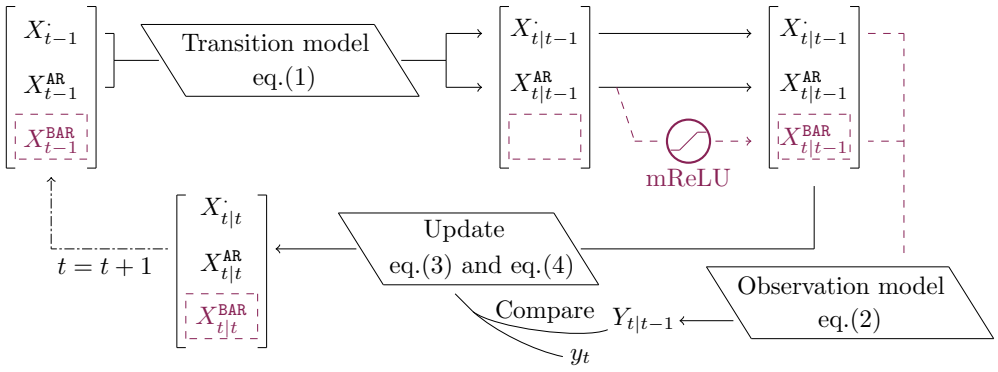


Figure 6: Integration process of the BAR component in the BDLM framework. The modifications made by the BAR method are highlighted by dashed lines.

5 Experiments

The BDLM-BAR coupled with a Switching Kalman Filter (SKF) [39] is tested in comparison with the existing BDLM-AR framework on three distinct case studies: (1) a synthetic time series with a

fixed anomaly in Section 5.1, (2) a real time series measured on a bridge with synthetic anomalies having three different magnitudes in Section 5.2, and (3) in Section 5.3 eight independent time series recorded on a dam with synthetic anomalies having varying magnitudes, shapes and locations.

5.1 Synthetic data

A synthetic time series is generated using a BDLM comprising a local level (LL) and an autoregressive (AR) component, covering the period from $2020/01/01$ to $2021/01/01$ with a daily time-step length. The model matrices for this experiment can be found in Appendix A.2.1. A constant-speed anomaly is introduced starting at the midpoint ($2020/07/01$) of the time series as depicted in Figure 7.

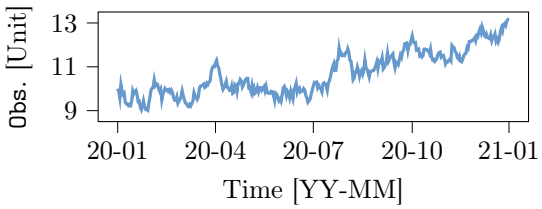


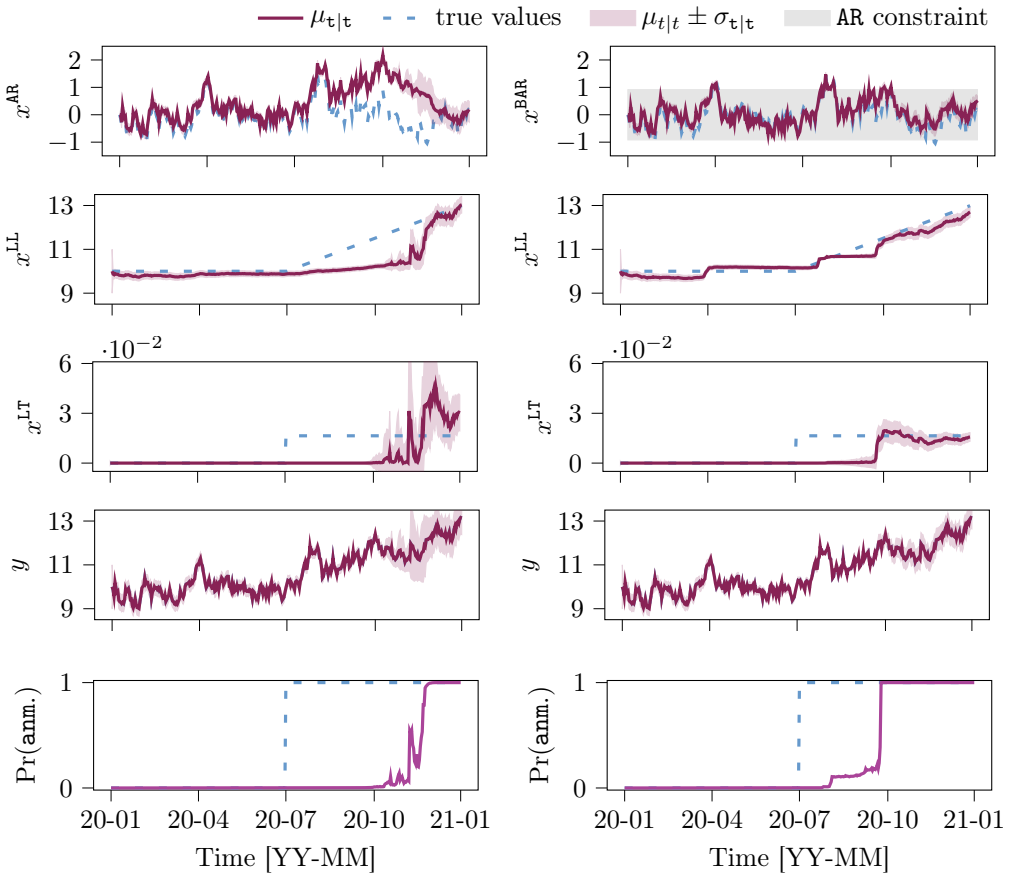
Figure 7: A one-year synthetic time series with daily time-step length generated with LL and AR components. A constant-speed anomaly starts at the midpoint $2020/07/01$.

A SKF involving a *stationary* and a *trend-stationary regime* is used to detect the simulated anomaly. The stationary regime is modelled using $\texttt{LL} + \texttt{AR}/\texttt{BAR}$ components, while the trend-stationary regime is modelled by a local trend (LT) instead of a local level component. The parameter γ in the BAR component is set to 2. The model matrices for these regimes are summarized in Appendix A.2.2 and A.2.3. The probability of detecting the synthetic anomaly, denoted as $\Pr(\texttt{anm.})$, is represented by the probability of the trend-stationary regime.
Figure 8 compares the estimations for the BDLM-AR and the BDLM-BAR with the true values. In the first plot from Figure 8a, the estimations of the AR hidden state gradually drift away from the true values after the introduction of the anomaly, implying that the changes in the trend caused by the anomaly are wrongfully captured by the residual component, which was initially intended to only model a stationary error-process. This drift results in delays in the anomaly detection quantified by $\Pr(\texttt{anm.})$, as well as inaccuracies in the hidden state estimation for the LL and LT components.
On the other hand, the BAR approach in Figure 8b prevents the drift and is capable to quickly catch up with the changes in the LL and LT hidden states. Additionally, its predicted responses y after the anomaly is introduced exhibit a smaller variance compared to those from the BDLM-AR model. As a result, the BDLM-BAR model detects the anomaly with $\Pr(\texttt{anm.}) \geq 0.5$ almost two months earlier than the BDLM-AR. This delay in the detection time can be further reduced for the BDLM-BAR model by imposing a smaller γ parameter to further constrain the AR hidden state, which will be discussed in Section 5.3.

5.2 Bridge data

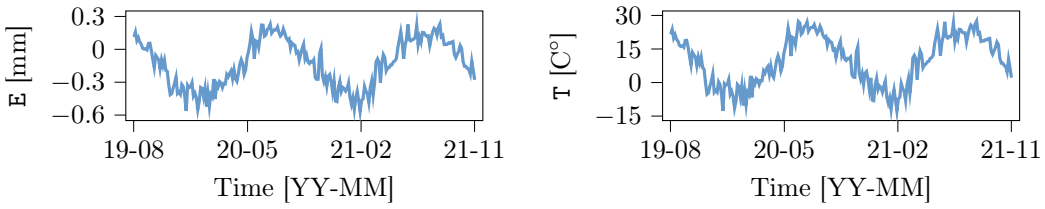
The time series studied in this section comprise measurements of elongation (E) and air temperature (T) obtained from a bridge. Both raw time series have a time-step length of 10 minutes, which is aggregated during pre-processing to 3.5 days as demonstrated in Figure 9. Elongation measurements exhibit a dependency on the air temperature, which can be accounted for in the observation model described in Section 3.1. The two time series, covering the period between $2019/08/16$ and $2021/11/16$, are considered to be stationary, given that no structural interventions or anomalies were recorded during this period, and the temperature remained stable. Three constant-speed anomalies with magnitude of -0.5, -0.1 and -0.025 mm/yr are introduced on the elongation measurements starting at the timestamp $2021/01/22$ as depicted by the true values for the $x^{\texttt{LT}}$ hidden state in Figure 10.

The BDLM components for the temperature measurement are $\{\texttt{LL}^{\texttt{T}} + \texttt{PD}^{\texttt{T}} + \texttt{AR}^{\texttt{T}}\}$ for both the stationary and trend-stationary regimes because there is no regime switch in the temperature. A local level component ($\texttt{LL}^{\texttt{T}}$) is used to describe the constant level of the temperature, a periodic component ($\texttt{PD}^{\texttt{T}}$) is used for modelling its yearly periodic pattern, and an autoregressive component



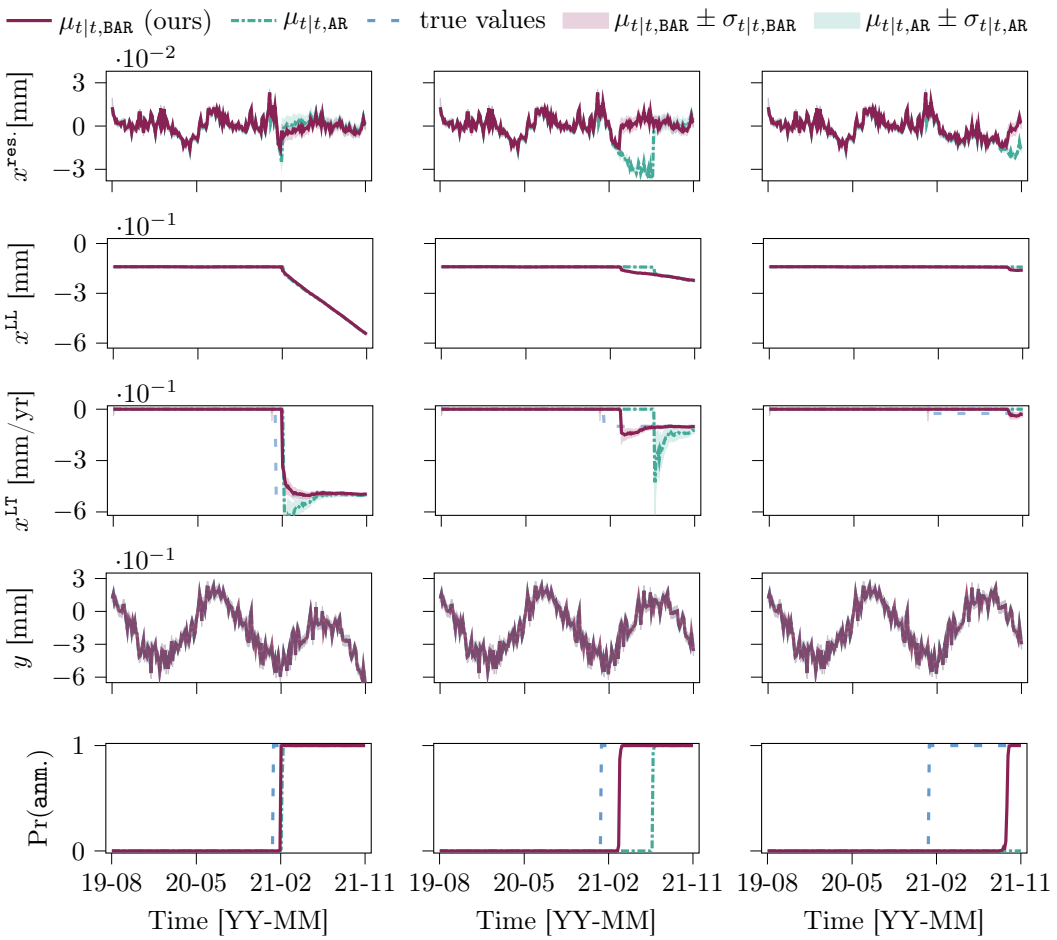
(a) BDLM-AR (b) BDLM-BAR (ours)

Figure 8: Hidden states estimation on the synthetic data with the standard AR and BAR in comparing with the true values.



(a) Elongation (b) Temperature

Figure 9: Elongation and temperature measurement on a bridge, where time steps are aggregated from 10 minutes to 3.5 days.



(a) Large anomaly (b) Medium anomaly (c) Small anomaly

Figure 10: Comparison of hidden states between BDLM-AR and BDLM-BAR under different anomaly magnitudes in a time series measured on a bridge. Large anomaly indicates an LT anomaly of -0.5 mm/year; the medium is -0.1 mm/year; and the small is -0.025 mm/year. The residual components (res.) include the AR and the BAR components.

(AR) is used to describe the model's errors. For the elongation measurement, $\{\texttt{LL}^{\texttt{E}} + \texttt{AR/BAR}^{\texttt{E}}\}$ and $\{\texttt{LT}^{\texttt{E}} + \texttt{AR/BAR}^{\texttt{E}}\}$ components are employed for the stationary and trend-stationary regimes without using a periodic component, which is accounted for through the dependency upon the temperature measurements. The constant level of the elongation for the stationary regime is described by a local level component ($\texttt{LL}^{\texttt{E}}$), while its constant speed for the trend-stationary regime is modelled by a local trend component ($\texttt{LT}^{\texttt{E}}$) with the model errors captured by either an AR or a BAR component. The parameter γ in the BAR component is set to 2. All model matrices are listed in Appendix A.3.

Figure 10 compares the results for the BDLM-BAR and the BDLM-AR models on the elongation time series with synthetic anomalies. When the anomaly is large enough to be seen visually as shown in Figure 10a, both models yield similar estimations with a slight delay in the anomaly detection. Yet, the BDLM-BAR outperforms the BDLM-AR in capturing the shift in trend $x^{\texttt{LT}}$ without overestimation at the beginning when there is a regime switch. As the anomaly magnitude decreases, it becomes almost impossible to identify it visually as shown in Figure 10b-c, and it takes a longer time for both models to detect it. BDLM-AR detects the second anomaly of -0.1 mm/yr after a delay of more than five months (on 2021/07/09), while BDLM-BAR detects it after only two (on 2021/03/23). For this anomaly, the BDLM-AR model displays a constant trend in the estimations of the residual component (res.) and fails to quickly capture the changes in $x^{\texttt{LL}}$ and $x^{\texttt{LT}}$ after the anomaly develops, which is consistent with the findings from Section 5.1. The smallest anomaly magnitude that BDLM-BAR can

detect is -0.025 mm/yr, as shown in Figure 10c, where the BDLM-BAR model correctly estimates the $x^{\texttt{LT}}$ hidden state after detecting the anomaly while maintaining a stationary residual process. In contrast, for the third case, the BDLM-AR is unable to detect or identify the changes in the hidden states.

5.3 Dam data

This section presents a probabilistic study of the BAR's performance under different γ values (BAR-γ) on eight time series measured on a dam. When γ is too small, false alarms may occur due to the strict constraint, when it is too large, BAR is similar to AR and is capable to capture trends. To estimate BAR's performance, not only we must consider the anomaly detectability but also the false alarm rate. However, the ground truth of whether anomalies are present is typically unknown for real time series, making it impossible to categorize an alarm as true or false. To overcome this issue, we generate synthetic time series using BDLM in Section 5.3.1, on which synthetic LA, LT and LL anomalies are overlaid. In Section 5.3.2, the performance of BAR-γ is verified on these synthetic time series with simulated anomalies and validated on the real time series in Section 5.3.3.

5.3.1 Overview of datasets and synthetic time series

The datasets used in this study consist of eight time series, each representing either automatic or manual measurements of crack openings or displacements on a dam. The first time series A01C spanning from 1998-01 to 2005-12 with a daily time interval measures the cracking openings on a dam. It demonstrates an increasing trend and yearly periodicity. The second time series namely A16D measures the displacements, and spans from 2011-11 to 2022-12 with a half-day time interval. It shows a decreasing trend and yearly periodicity. The M08C time series measuring the crack openings has a time interval of 91 days. It shows a decreasing trend between 2013-12 and 2022-12. The crack openings, measured in M09C time series with a 49-day time interval, increase between 1996-04 and 2007-03. The displacements, measured in M11D time series with a 28-day time interval, increase between 2007-02 and 2022-12. The M12D displacement time series has a 63-day time interval and a decreasing trend from 2011-11 to 2022-12. The M13D displacement time series decreases from 2009-01 to 2020-07, having a 35-day time interval. The M14D displacement time series is measured every 41 days, and increases between 1991-06 and 1999-08. Yearly periodicity and missing data are observed throughout all the time series as depicted in Appendix B. Table 2 provides a summary of the measurement type, start and end times, acquisition period, number of observations, and BDLM model components employed for each time series. Among them, results for M08C are showcased in the following sections, while those for the other time series are documented in Appendices C, D, and E.

Table 2: Summary of eight time series (TS) measured on a dam with respect to time-series name, measurement type, start & end time, acquisition period, number of observation and BDLM model components. Crack openings and displacement are notated as C.O. and Disp. Stationary/ non-stationary regimes are referred to as S./ N.S. The abbreviation of res. includes AR or BAR residual component. The *emphasized row* is the TS example showcased in the following sections.

Name	Type	Start time	End time	Period [day]	#obs.	Components (S./ N.S.)
A01C	C.O.	1998-01-29	2005-12-29	1	2789	LT/LA+ KR (10)+ res.
A16D	Disp.	2011-11-17	2022-12-07	0.5	8154	LT/LA+ KR (10)+ res.
M08C	*C.O.*	*2013-12-09*	*2022-12-06*	*91*	*39*	LT/LA+ PD+ *res.*
M09C	C.O.	1996-04-11	2007-03-06	49	56	LT/LA+ PD+ res.
M11D	Disp.	2007-02-14	2022-12-06	28	149	LT/LA+ PD+ res.
M12D	Disp.	2011-11-17	2022-12-07	63	95	LT/LA+ PD+ res.
M13D	Disp.	2009-01-14	2020-07-28	35	122	LT/LA+ PD+ res.
M14D	Disp.	1991-06-10	1999-08-17	41	74	LT/LA+ PD+ res.

Figure 11 illustrates the time series M08C in black, along with 100 similar synthetic time series in colours. The real observations range between -3.4 mm and -2.2 mm exhibiting an annual periodic pattern and a decreasing trend. The synthetic time series are generated using a BDLM with optimized parameters, whose matrices are listed in Appendix A.4. In Section 5.3.2 where we use the synthetic

time series to verify the performance of BAR-γ, we use a unified time span of 10 years instead of the original ones.

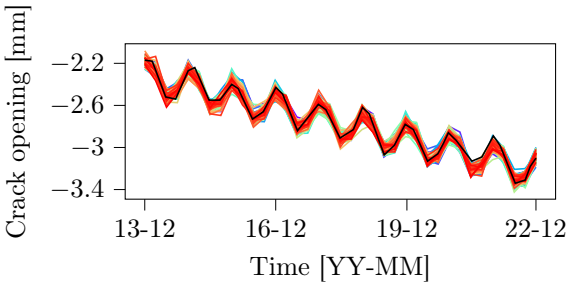
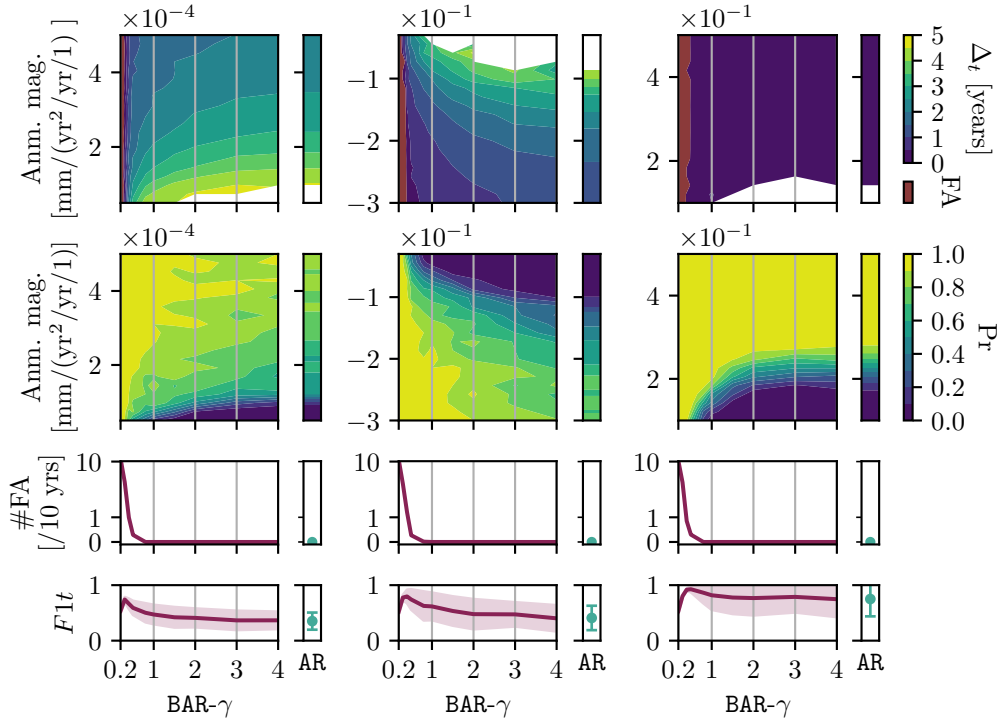


Figure 11: 100 synthetic time series (coloured) similar to a real time series (black) measured on a dam.

5.3.2 Verification of BAR-γ performance

The detection time of anomalies may vary according to their starting times, even when they have identical magnitudes. To account for this factor, for each anomaly magnitude and γ value, the BAR-γ model is run on 100 synthetic time series with anomalies introduced at random timestamps within the first five years. An anomaly is considered as detected when the probability of the non-stationary regime is higher than 0.5 within a detection window $l_{dw} = 5$ years, so that even if an anomaly is imposed at the end of the first five years, there is enough time to detect it in the second half. An alarm triggered on a time series where there is no simulated anomaly is considered as a false alarm. Whenever a false alarm is triggered, the analysis on this time series is terminated so that the maximum number of false alarms for each time series is one.
Figure 12 evaluates BAR-γ and AR using four metrics: (1) the average detection time Δ_t over the 100 synthetic time series; (2) the detection probability (Pr) which quantifies, out of the 100 synthetic time series, how many have correctly triggered alarms; (3) the number of false alarms per 10 years; (4) the $F1t$ metric which aggregates the anomaly detectability, detection time, as well as false alarm rate as described in Section 2.2. In Figure 12, the BAR and AR methods are verified on the LA, LT, and LL anomalies presented in Section 2.1. The optimal parameter γ^* for detecting the LA anomalies using BAR is found to be 0.3 according to its highest $F1t$ score of 0.742, where the expected number of false alarms is 3.5 in every ten years. Compared to AR, BAR-0.3 exhibits a better performance with a shorter detection time and higher detection probability, as shown in the first and the second rows of plots by dark blue and yellow colours, respectively. The white regions in the Δ_t subfigures indicate non-detectable anomalies and the dark red regions mark the negative detection time caused by false alarms (FA). By increasing γ, i.e., relaxing the constraints imposed on the AR hidden state, the BAR eventually converges to an AR component with both methods having a similar performance in the four tested metrics.
Similar convergences are observed in the detection of LT and LL anomalies, where BAR-0.4 and BAR-0.5 perform best with expected $F1t$ scores of 0.796 and 0.927, respectively. When a LL anomaly occurs, both BDLM-BAR and BDLM-AR models either detect it with little delay or fail to trigger an alarm altogether. This is because LL anomalies only shift the hidden state x^{LL} at one timestamp after which they switch back to the stationary regime, in contrast to the accumulated changes with the LT or LA anomalies. The additional verification on the other seven time series are provided in Appendix C.

The expected values and standard deviations of the $F1t$ scores for BAR-γ^* and AR verified on the synthetic datasets with LA anomalies are summarized in Figure 13 for the eight real time series. Similar plots for LT and LL anomalies are included in Appendix D. The $F1t$ scores for the three types of anomalies, as well as the optimal parameter γ^* and the false alarm rate under each γ^*, are listed in Table 3. Throughout a better tuning of the γ parameter, BAR significantly outperforms AR with higher expected $F1t$ scores and lower standard deviations. $F1t$ scores of BAR that are 80% higher than those of AR are highlighted in Table 3.



(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 12: Performance of the BAR method under different γ on synthetic time series with three types of anomalies in comparison with the AR. Evaluation metrics are detection time (Δ_t), detection probability (Pr), false alarm number (#FA) per ten years and $F1t$ score. The white regions on the top three subfigures indicate the undetectable anomalies and their dark red regions represent all the negative detection time caused by the false alarms (FA). False alarm rate is drawn in logarithmic scale. A higher value of $F1t$ indicates better performance. Repetition N = 100.

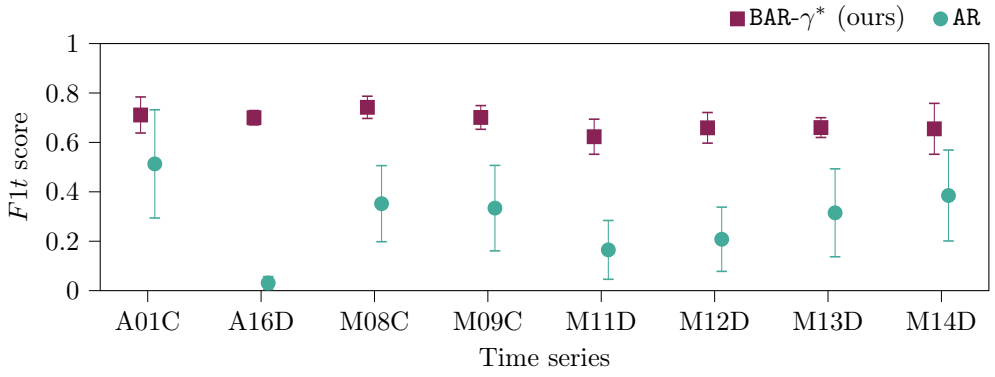


Figure 13: The $F1t$ scores for BAR-γ^* and AR verified on the synthetic datasets similar to the eight real time series measured on a dam. LA anomalies are overlaid on the synthetic time series. Additional $F1t$ scores with respect to the LT and LL anomalies are shown in Appendix D.

Table 3: Summary of AR v.s. BAR-γ^* performance. γ^* is the optimal parameter corresponding to the highest $F1t$ in the BAR method. **Bolded** numbers indicate the $F1t$ scores improved by more than 80% in BAR compared to AR.

		AR		BAR-γ^* (ours)			
Anm.	TS	$\mathbb{E}(F1t)$	$\sigma(F1t)$	$\mathbb{E}(F1t)$	$\sigma(F1t)$	γ^*	#FA/10 yrs
LA	A01C	0.513	0.219	0.711	0.073	2.0	1.533
	A16D	0.031	0.026	**0.700**	0.029	1.2	6.749
	M08C	0.352	0.154	**0.742**	0.045	0.3	3.466
	M09C	0.334	0.173	**0.701**	0.048	0.4	3.009
	M11D	0.165	0.119	**0.623**	0.071	0.5	1.853
	M12D	0.208	0.130	**0.659**	0.062	0.4	1.865
	M13D	0.315	0.178	**0.660**	0.040	0.4	5.635
	M14D	0.385	0.184	0.655	0.103	0.5	1.032
LT	A01C	0.302	0.211	**0.659**	0.140	2.2	0.661
	A16D	N/A	N/A	**0.686**	0.044	1.2	6.083
	M08C	0.409	0.220	**0.796**	0.122	0.4	1.207
	M09C	0.194	0.137	**0.643**	0.038	0.4	3.739
	M11D	0.380	0.202	**0.740**	0.102	0.5	1.782
	M12D	0.429	0.224	0.771	0.076	0.4	1.585
	M13D	0.443	0.255	0.772	0.108	0.5	1.382
	M14D	0.402	0.244	0.712	0.127	0.5	1.376
LL	A01C	0.466	0.291	0.536	0.226	2.4	0.250
	A16D	0.901	0.213	0.898	0.214	4.0	0.000
	M08C	0.750	0.312	0.927	0.055	0.5	0.298
	M09C	0.792	0.304	0.886	0.137	0.5	0.801
	M11D	0.692	0.358	0.932	0.103	0.8	0.038
	M12D	0.795	0.299	0.897	0.137	0.5	0.494
	M13D	0.791	0.318	0.862	0.075	0.5	1.403
	M14D	0.694	0.355	0.859	0.063	0.5	1.226

5.3.3 Validation on real data with synthetic anomalies

This section presents a validation experiment on the real time series M08C with LA synthetic anomalies to confirm the verification results obtained in Section 5.3.2. Additional validation results for other time series and anomaly types are included in Appendix E. The detection window is set to be half the number of data points. LA anomalies of magnitudes varying between 10^{-5} and 10^{-3} mm/yr^2 are randomly added on the first half of the time series, while the second half is reserved as the detection window. In order to validate the entire time series, the same process is also repeated on a copy of the time series for which the time direction is flipped. Here, γ is set to 0.4 instead of the optimal value of 0.3 as suggested by the highest $F1t$ score in order to reduce the false alarm rate to the target value of one per ten years.

Figure 14 shows the anomaly detection time and the probability of regime switch achieved by the BAR-0.4 model in comparison with the AR model. In the second plot from Figure 14, the faded circles provide an overview of the anomaly detection time for each real time series with anomalies introduced at a random location, while the solid lines and and the shaded regions represent their means and standard deviations for each anomaly magnitude. The BAR-0.4 consistently exhibits a higher detection probability than the AR until they both converge to one, at which point it has the potential to detect anomalies twice as fast as the AR. The smallest detectable anomaly magnitude of BAR-0.4 is around 2×10^{-5} mm/yr^2 compared to 7×10^{-5} for the AR. These results are consistent with those in Figure 12. Take the range of anomaly magnitude between 2×10^{-4} and 4×10^{-4} as an example, the results in Figure 12 indicate that the detection time of BAR-0.4 lies in the interval between 1 and 2 years with a probability close to one, while the AR has a detection time ranging from 2 to 3.5 years with a probability between 0.6 and 0.8. Similar ranges are observed in Figure 14 as well, indicating that the model evaluation on the synthetic time series in Section 5.3.2 is an effective representation of the models' performance on the real time series.

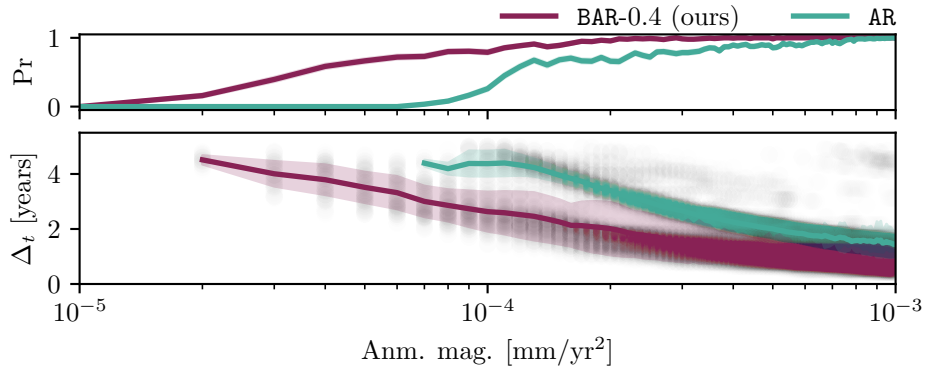


Figure 14: Validation results with regard to the anomaly detection probability (Pr) and time (Δ_t) on the real time series M08C with LA anomalies of different magnitudes by using BAR-0.4 and AR components. 0.4 is the selected value for the hyper-parameter γ in BAR. Additional validation results for other time series and anomaly types are reported in Appendix E.

6 Conclusions

This study presents a new approach within the existing BDLM framework for improving anomaly detection for SHM time series. The method proposes a novel residual component consisting of a constrained autoregressive hidden state. The performance of this new approach is evaluated using synthetic as well as bridge and dam datasets which consider a stochastic performance evaluation scheme involving three common types of anomalies encountered in infrastructure measurements. The experimental results demonstrate that the bounded autoregressive model surpasses the performance of the existing autoregressive model with (1) an improved accuracy at estimating hidden states, (2) an early detection of anomalies, (3) a capacity to detect smaller anomaly magnitudes, and (4) the ability to control the tradeoff between the anomaly detectability and the false alarm rate. The anomaly detectability, the detection time, and the false alarm rate are all integrated into an $F1t$ evaluation metric proposed in this paper. Overall, these findings highlight the effectiveness of the proposed method in enhancing the detection of anomalies in infrastructure.

7 Acknowledgement

The authors acknowledge the contribution of Hydro-Québec who provided the datasets and who have funded this research, as well as funding from the Hydro-Québec Research Institute (IREQ) and Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] ASCE. 2013 report card for America's infrastructure. Technical report, 2013.

[2] S. El-Zahab, E. Mohammed Abdelkader, and T. Zayed. An accelerometer-based leak detection system. *Mechanical Systems and Signal Processing*, 108:276–291, 2018.

[3] F. Cerda, S. Chen, J. Bielak, J.H. Garrett, P. Rizzo, and J. Kovacevic. Indirect structural health monitoring of a simplified laboratory-scale bridge model. *Smart Structures and Systems*, 13:849–868, May 2014.

[4] S. S. Saidin, S. A. Kudus, A. Jamadin, M. A. Anuar, N. M. Amin, A. B. Z. Ya, and K. Sugiura. Vibration-based approach for structural health monitoring of ultra-high-performance concrete bridge. *Case Studies in Construction Materials*, 18:e01752, 2023.

[5] R. Kromanis and P. Kripakaran. Data-driven approaches for measurement interpretation: analysing integrated thermal and vehicular response in bridge structural health monitoring. *Advanced Engineering Informatics*, 34:46–59, 2017.

[6] J.-A. Goulet. Bayesian dynamic linear models – a generic framework for structural health monitoring bayesian dynamic linear models – a generic framework for structural health monitoring. *Conference proceedings form ICOSSAR*, 2017.

[7] F. Cavadas, I. F.C. Smith, and J. Figueiras. Damage detection using data-driven methods applied to moving-load responses. *Mechanical Systems and Signal Processing*, 39(1):409–425, 2013.

[8] Z. Hamida and J.-A. Goulet. A stochastic model for estimating the network-scale deterioration and effect of interventions on bridges. *Structural Control and Health Monitoring*, 29(4):e2916, 2022.

[9] L. Ierimonti, I. Venanzi, N. Cavalagli, E. García-Macías, and F. Ubertini. A Bayesian-based data fusion methodology and its application for seismic structural health monitoring of the Consoli Palace in Gubbio, Italy. *Procedia Structural Integrity*, 2023.

[10] L.H. Nguyen and J.-A. Goulet. Anomaly detection with the switching Kalman filter for structural health monitoring. *Structural Control and Health Monitoring*, 25(4):e2136, 2018.

[11] Z. Tang, Z. Chen, Y. Bao, and H. Li. Convolutional neural network-based data anomaly detection method using multiple information for structural health monitoring. *Structural Control and Health Monitoring*, 26(1):e2296, 2019.

[12] L. H. Nguyen, I. Gaudot, S. Khazaeli, and J.-A. Goulet. A kernel-based method for modeling non-harmonic periodic phenomena in Bayesian dynamic linear models. *Frontiers in Built Environment*, 5:8, 2019.

[13] X. Xia, X. Pan, N. Li, X. He, L. Ma, X. Zhang, and N. Ding. Gan-based anomaly detection: A review. *Neurocomputing*, 493:497–535, 2022.

[14] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. J. Inman. A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mechanical Systems and Signal Processing*, 147:107077, 2021.

[15] A. Moallemi, A. Burrello, D. Brunelli, and L. Benini. Model-based vs. data-driven approaches for anomaly detection in structural health monitoring: a case study. In *2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pages 1–6, 2021.

[16] Zhi Ma, Chung-Bang Yun, Hua-Ping Wan, Yanbin Shen, Feng Yu, and Yaozhi Luo. Probabilistic principal component analysis-based anomaly detection for structures with missing data. *Structural Control and Health Monitoring*, 28(5):e2698, 2021.

[17] D. Jana, J. Patil, S. Herkal, S. Nagarajaiah, and L. Duenas-Osorio. CNN and convolutional autoencoder (CAE) based real-time sensor fault detection, localization, and correction. *Mechanical Systems and Signal Processing*, 169:108723, 2022.

[18] J. Mao, H. Wang, and B. F. Spencer. Toward data anomaly detection for automated structural health monitoring: Exploiting generative adversarial nets and autoencoders. *Structural Health Monitoring*, 20(4):1609–1626, 2021.

[19] J. Zhang, J. Zhang, and Z. Wu. Long-short term memory network-based monitoring data anomaly detection of a long-span suspension bridge. *Sensors*, 22(16), 2022.

[20] W.-J. Jiang, C.-W. Kim, Y. Goi, and F.-L. Zhang. Data normalization and anomaly detection in a steel plate-girder bridge using lstm. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, 8(1):04021082, 2022.

[21] X. Lei, Y. Xia, A. Wang, X. Jian, H. Zhong, and L. Sun. Mutual information based anomaly detection of monitoring data with attention mechanism and residual learning. *Mechanical Systems and Signal Processing*, 182:109607, 2023.

[22] S. Khazaeli, L. H. Nguyen, and J.-A. Goulet. Anomaly detection using state-space models and reinforcement learning. *Structural Control and Health Monitoring*, 28(6):e2720, 2021.

[23] C. Zhong, M. C. Gursoy, and S. Velipasalar. Controlled sensing and anomaly detection via soft actor-critic reinforcement learning. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4198–4202, 2022.

[24] Y.-M. Zhang, H. Wang, Y. Bai, J.-X. Mao, X.-Y. Chang, and L.-B. Wang. Switching bayesian dynamic linear model for condition assessment of bridge expansion joints using structural health monitoring data. *Mechanical Systems and Signal Processing*, 160:107879, 2021.

[25] N. Aswal, S. Sen, and L. Mevel. Switching Kalman filter for damage estimation in the presence of sensor faults. *Mechanical Systems and Signal Processing*, 175:109116, 2022.

[26] L. C. K. Reuben and D. Mba. Diagnostics and prognostics using switching kalman filters. *Structural Health Monitoring*, 13(3):296–306, 2014.

[27] A.R. Runnalls. Kullback-Leibler approach to Gaussian mixture reduction. *IEEE Transactions on Aerospace and Electronic Systems*, 43(3):989–999, 2007.

[28] J.-A. Goulet. *Probabilistic Machine Learning for Civil Engineers*. MIT Press, 2020.

[29] J. Paparrizos, Y. Kang, P. Boniol, R. S. Tsay, T. Palpanas, and M. J. Franklin. TSB-UAD: An end-to-end benchmark suite for univariate time-series anomaly detection. In *Proceedings of the VLDB Endowment*, 2022.

[30] M. Goswami, C. Challu, L. Callot, L. Minorics, and A. Kan. Unsupervised model selection for time-series anomaly detection. In *Proceedings of ICLR*, 2023.

[31] C. J. Van Rijsbergen. *Information Retrieval*. Butterworth-Heinemann, 2nd edition, 1979.

[32] A. Bhatnagar, P. Kassianik, C. Liu, T. Lan, W. Yang, R. Cassius, D. Sahoo, D. Arpit, S. Subramanian, G. Woo, A. Saha, A. K. Jagota, G. Gopalakrishnan, M. Singh, K C Krithika, S. Maddineni, D. Cho, B. Zong, Y. Zhou, C. Xiong, S. Savarese, S. Hoi, and H. Wang. Merlion: A machine learning library for time series, 2021.

[33] C. D. Petersen, R. Fraanje, B. S. Cazzolato, A. C. Zander, and C. H. Hansen. A Kalman filter approach to virtual sensing for active noise control. *Mechanical Systems and Signal Processing*, 22(2):490–508, 2008.

[34] J.A. Goulet. Bayesian dynamic linear models for structural health monitoring. *Structural Control and Health Monitoring*, 24(12):1545–2263, 2017.

[35] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. Attention is all you need. In *Conference on Neural Information Processing Systems*, 2017.

[36] M. Mirza and S. Osindero. Conditional generative adversarial nets. In *Conference on Neural Information Processing Systems*, 2014.

[37] A. Krizhevsky and G. Hinton. Convolutional deep belief networks on CIFAR-10. CiteSeer preprint, 2010.

[38] D. Simon and D. L. Simon. Constrained Kalman filtering via density function truncation for turbofan engine health estimation. *International Journal of Systems Science*, 41(2):159–171, 2010.

[39] K.P. Murphy. Switching Kalman filters. Technical report, 1998.

A Appendix: BDLM matrices and parameters

This appendix provides all the model matrices and their parameters used in the three experiments shown in Section 5.

A.1 Component matrices and parameters overview

The baseline components in BDLM are LL, LT and LA. The matrices for LL component are

$$\mathbf{A}^{\mathtt{LL}} = 1, \mathbf{F}^{\mathtt{LL}} = 1, \mathbf{Q}^{\mathtt{LL}} = \left(\sigma_w^{\mathtt{LL}}\right)^2.$$

The matrices for LT are

$$\mathbf{A}^{\mathtt{LT}} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \mathbf{F}^{\mathtt{LT}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\intercal}, \mathbf{Q}^{\mathtt{LT}} = \left(\sigma_w^{\mathtt{LT}}\right)^2 \begin{bmatrix} \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^2}{2} & \Delta t \end{bmatrix}.$$

The matrices for LA are

$$\mathbf{A}^{\mathtt{LA}} = \begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{F}^{\mathtt{LA}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^{\intercal}, \mathbf{Q}^{\mathtt{LA}} = \left(\sigma_w^{\mathtt{LA}}\right)^2 \begin{bmatrix} \frac{\Delta t^2}{20} & \frac{\Delta t^4}{8} & \frac{\Delta t^3}{6} \\ \frac{\Delta t^4}{8} & \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^3}{6} & \frac{\Delta t^2}{2} & \Delta t \end{bmatrix}.$$

The periodic components in BDLM are harmonic periodic component (PD) and kernel regressive component (KR). The matrices for the PD component are

$$\mathbf{A}^{\mathtt{PD}} = \begin{bmatrix} \cos\omega & \sin\omega \\ \sin\omega & \cos\omega \end{bmatrix}, \mathbf{F}^{\mathtt{PD}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\intercal}, \mathbf{Q}^{\mathtt{PD}} = \left(\sigma_w^{\mathtt{PD}}\right)^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

where $\omega = \frac{2\pi \cdot \Delta t}{p}$ and p is the period.
The matrices for the KR component with 10 control points are

$$\mathbf{A}^{\mathtt{KR}} = \begin{bmatrix} 0 & \tilde{\boldsymbol{k}}_t^{\mathtt{KR}} \\ \mathbf{0} & \mathbf{I}_{10\times 10} \end{bmatrix}, \mathbf{F}^{\mathtt{KR}} = \begin{bmatrix} 1 \\ \mathbf{0}_{10\times 1} \end{bmatrix}^{\intercal}, \mathbf{Q}^{\mathtt{KR}} = \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix},$$

where $\tilde{\boldsymbol{k}}_t^{\mathtt{KR}} = \left[\tilde{k}_{t,1}^{\mathtt{KR}}, \tilde{k}_{t,2}^{\mathtt{KR}}, ..., \tilde{k}_{t,10}^{\mathtt{KR}}\right]$ is the normalized kernel depending on the time t.

The matrices for the residual components, AR and BAR, are presented in Section 3.2.1 and 4.3, respectively.

The matrices of baseline components are adapted in SKF. The $\mathbf{A}$ and $\mathbf{F}$ matrices for the the stationary model need to be compatible with the non-stationary model matrices by having the same size. The $\mathbf{Q}$ matrix in SKF incorporates the uncertainty of regimes switching. From the stationary regime to the trend-stationary regime, the compatible matrices, $\mathbf{A}$ and $\mathbf{F}$, for stationary regime are

$$\mathbf{A}^{\mathtt{LcT}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{F}^{\mathtt{LcT}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\intercal}.$$

The $\mathbf{Q}$ matrices differ in each regime switch so that

$$\mathbf{Q}^{\mathtt{LL(LL)}} = \begin{bmatrix} \left(\sigma_w^{\mathtt{LL}}\right)^2 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{Q}^{\mathtt{LT(LT)}} = \mathbf{Q}^{\mathtt{LT}},$$

$$\mathbf{Q}^{\mathtt{LL(LT)}} = \begin{bmatrix} \left(\sigma_w^{\mathtt{LT}}\right)^2 \frac{\Delta t^3}{3} & 0 \\ 0 & \left(\sigma_w^{\mathtt{LTT}}\right)^2 \Delta t \end{bmatrix}, \quad \mathbf{Q}^{\mathtt{LT(LL)}} = \begin{bmatrix} \left(\sigma_w^{\mathtt{LL}}\right)^2 & 0 \\ 0 & 0 \end{bmatrix}.$$

From the LT stationary regime to the LA non-stationary regime, the compatible matrices for LT stationary regime are

$$\mathbf{A}^{\mathtt{TcA}} = \begin{bmatrix} 1 & \Delta t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{F}^{\mathtt{TcA}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^{\intercal}.$$

The $\mathbf{Q}$ matrices for regime switching are

$$\mathbf{Q}^{\mathtt{LT}(\mathtt{LT})} = \begin{bmatrix} \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} & 0 \\ \frac{\Delta t^2}{2} & \Delta t & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{Q}^{\mathtt{LA}(\mathtt{LA})} = \mathbf{Q}^{\mathtt{LA}},$$

$$\mathbf{Q}^{\mathtt{LT}(\mathtt{LA})} = \begin{bmatrix} \left(\sigma_w^{\mathtt{LA}}\right)^2 \frac{\Delta t^2}{20} & 0 & 0 \\ 0 & \left(\sigma_w^{\mathtt{LA}}\right)^2 \frac{\Delta t^3}{3} & 0 \\ 0 & 0 & \left(\sigma_w^{\mathtt{LTT}}\right)^2 \Delta t \end{bmatrix}, \quad \mathbf{Q}^{\mathtt{LA}(\mathtt{LT})} = \begin{bmatrix} \left(\sigma_w^{\mathtt{LT}}\right)^2 \frac{\Delta t^3}{3} & 0 & 0 \\ 0 & \left(\sigma_w^{\mathtt{LA}}\right)^2 \Delta t & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

For all the experiments in this paper, $\sigma_w^{\mathtt{LL}}$, $\sigma_w^{\mathtt{LT}}$ and $\sigma_w^{\mathtt{LA}}$ are set to be 0. The process error's standard deviations for the periodic components, $\sigma_w^{\mathtt{PD}}$ and $\sigma_w^{\mathtt{KR}}$, are set to 0 and their periods p are 365.24. The uncertainties in regime switching, $\left(\sigma_w^{\mathtt{LTT}}\right)^2 \Delta t$, is fixed to 10^{-6}. The remaining model parameters are kernel width $\ell^{\mathtt{KR}}$ for kernel regressive component, $\phi^{\mathtt{AR}}$ and $\sigma^{\mathtt{AR}}$ for residual components, observation error σ_v, z^{11} and z^{22} for SKF transition matrix. Among them, $\ell^{\mathtt{KR}}$, $\phi^{\mathtt{AR}}$ and $\sigma^{\mathtt{AR}}$ are optimized by Newton-Raphson gradient descend. Other parameters are decided based on engineering heuristics.

A.2 Synthetic data

This appendix provides a summary of the model matrices and parameters used in the synthetic time series generation and the SKF analysis.

A.2.1 Synthetic data generation

In the toy problem, the synthetic data generation employs LL and AR components. After the synthetic time series are generated, a LT anomaly is applied on top of it. Model matrices to generate synthetic data (GS) are

$$\begin{aligned} \mathbf{A}^{\mathrm{GS}} &= \text{blockdiag}\left(\mathbf{A}^{\mathtt{LL}}, \mathbf{A}^{\mathtt{AR}}\right) \\ \mathbf{F}^{\mathrm{GS}} &= \left[\mathbf{F}^{\mathtt{LL}}, \mathbf{F}^{\mathtt{AR}}\right] \\ \mathbf{Q}^{\mathrm{GS}} &= \text{blockdiag}\left(\mathbf{Q}^{\mathtt{LL}}, \mathbf{Q}^{\mathtt{AR}}\right). \end{aligned}$$

Parameters values for synthetic data generation are

$$\mathcal{P}^{\mathrm{GS}} = \{\phi^{\mathtt{AR}}, \sigma^{\mathtt{AR}}, \sigma_v\} = \{0.9, 0.2, 0.001\}.$$

A.2.2 Stationary model matrices in SKF

The stationary model matrices for the synthetic time series, referred to as SS, are

$$\begin{aligned} \mathbf{A}^{\mathrm{SS}} &= \text{blockdiag}\left(\mathbf{A}^{\mathtt{LcT}}, \mathbf{A}^{\mathtt{AR/BAR}}\right) \\ \mathbf{F}^{\mathrm{SS}} &= \left[\mathbf{F}^{\mathtt{LcT}}, \mathbf{F}^{\mathtt{AR/BAR}}\right] \\ \mathbf{Q}^{\mathrm{SS},\mathtt{LL}(\mathtt{LL})} &= \text{blockdiag}\left(\mathbf{Q}^{\mathtt{LL}(\mathtt{LL})}, \mathbf{Q}^{\mathtt{AR/BAR}}\right) \\ \mathbf{Q}^{\mathrm{SS},\mathtt{LT}(\mathtt{LL})} &= \text{blockdiag}\left(\mathbf{Q}^{\mathtt{LT}(\mathtt{LL})}, \mathbf{Q}^{\mathtt{AR/BAR}}\right). \end{aligned}$$

A.2.3 Non-stationary model matrices in SKF

The non-stationary model matrices for the synthetic time series, referred to as NS, are

$$\begin{aligned} \mathbf{A}^{\mathrm{NS}} &= \text{blockdiag}\left(\mathbf{A}^{\mathtt{LT}}, \mathbf{A}^{\mathtt{AR/BAR}}\right) \\ \mathbf{F}^{\mathrm{NS}} &= \left[\mathbf{F}^{\mathtt{LT}}, \mathbf{F}^{\mathtt{AR/BAR}}\right] \\ \mathbf{Q}^{\mathrm{NS},\mathtt{LT}(\mathtt{LT})} &= \text{blockdiag}\left(\mathbf{Q}^{\mathtt{LT}(\mathtt{LT})}, \mathbf{Q}^{\mathtt{AR/BAR}}\right) \\ \mathbf{Q}^{\mathrm{NS},\mathtt{LL}(\mathtt{LT})} &= \text{blockdiag}\left(\mathbf{Q}^{\mathtt{LL}(\mathtt{LT})}, \mathbf{Q}^{\mathtt{AR/BAR}}\right). \end{aligned}$$

Parameters values for both regimes in SKF on synthetic time series are

$$\mathcal{P}^{\mathrm{S}} = \{\phi^{\mathtt{AR}}, \sigma^{\mathtt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.9, 0.2, 0.001, 0.999999, 0.999999\}.$$

A.3 Bridge data

This appendix presents the model matrices and their parameters used in the SKF analysis for the bridge dataset including measurements of air temperature (T) and elongation (E).

A.3.1 Stationary model matrices in SKF

The stationary model matrices for the bridge time series, referred to as SB, are

$$\begin{aligned}
\mathbf{A}^{\mathrm{SB}} &= \text{blockdiag}\left(\overbrace{\mathbf{A}^{\mathtt{LL,T}}, \mathbf{A}^{\mathtt{PD,T}}\mathbf{A}^{\mathtt{AR,T}}}^{\text{Temperature}}, \overbrace{\mathbf{A}^{\mathtt{LcT,E}}, \mathbf{A}^{\mathtt{AR/BAR,E}}}^{\text{Elongation}}\right)\\
\mathbf{Q}^{\mathrm{SB},\mathtt{LL}(\mathtt{LL})} &= \text{blockdiag}\left(\overbrace{\mathbf{Q}^{\mathtt{LL,T}}, \mathbf{Q}^{\mathtt{PD,T}}\mathbf{Q}^{\mathtt{AR,T}}}^{\text{Temperature}}, \overbrace{\mathbf{Q}^{\mathtt{LL}(\mathtt{LL}),\mathtt{E}}, \mathbf{Q}^{\mathtt{AR/BAR,E}}}^{\text{Elongation}}\right)\\
\mathbf{Q}^{\mathrm{SB},\mathtt{LT}(\mathtt{LL})} &= \text{blockdiag}\left(\overbrace{\mathbf{Q}^{\mathtt{LL,T}}, \mathbf{Q}^{\mathtt{PD,T}}\mathbf{Q}^{\mathtt{AR,T}}}^{\text{Temperature}}, \overbrace{\mathbf{Q}^{\mathtt{LT}(\mathtt{LL}),\mathtt{E}}, \mathbf{Q}^{\mathtt{AR/BAR,E}}}^{\text{Elongation}}\right).
\end{aligned}$$

The observation matrix $\mathbf{F}$ models the dependency between temperature and elongation. For the standard `AR` model, the observation matrix is

$$\mathbf{F}^{\mathrm{SB},\mathtt{AR}} = \begin{bmatrix} \overbrace{1}^{\mathtt{LL,T}} & \overbrace{1 \quad 0}^{\mathtt{PD,T}} & \overbrace{1}^{\mathtt{AR,T}} & \overbrace{0 \quad 0}^{\mathtt{LcT,E}} & \overbrace{0}^{\mathtt{AR,E}} \\ 0 & \beta^{\mathtt{PD,E|T}} \quad 0 & \beta^{\mathtt{AR,E|T}} & 1 \quad 0 & 1 \end{bmatrix}.$$

For `BAR`, the observation matrix is

$$\mathbf{F}^{\mathrm{SB},\mathtt{BAR}} = \begin{bmatrix} \overbrace{1}^{\mathtt{LL,T}} & \overbrace{1 \quad 0}^{\mathtt{PD,T}} & \overbrace{1}^{\mathtt{AR,T}} & \overbrace{0 \quad 0}^{\mathtt{LcT,E}} & \overbrace{0 \quad 0}^{\mathtt{BAR,E}} \\ 0 & \beta^{\mathtt{PD,E|T}} \quad 0 & \beta^{\mathtt{AR,E|T}} & 1 \quad 0 & 0 \quad 1 \end{bmatrix},$$

where the first row corresponds to the temperature observation and the second corresponds to the elongation. β is the dependency ratio between the elongation hidden states and the temperature hidden states.

A.3.2 Non-stationary model matrices in SKF

The non-stationary model matrices for the bridge time series, referred to as NB, are

$$\begin{aligned}
\mathbf{A}^{\mathrm{NB}} &= \text{blockdiag}\left(\overbrace{\mathbf{A}^{\mathtt{LL,T}}, \mathbf{A}^{\mathtt{PD,T}}\mathbf{A}^{\mathtt{AR,T}}}^{\text{Temperature}}, \overbrace{\mathbf{A}^{\mathtt{LT,E}}, \mathbf{A}^{\mathtt{AR/BAR,E}}}^{\text{Elongation}}\right)\\
\mathbf{Q}^{\mathrm{NB},\mathtt{LT}(\mathtt{LT})} &= \text{blockdiag}\left(\overbrace{\mathbf{Q}^{\mathtt{LL,T}}, \mathbf{Q}^{\mathtt{PD,T}}\mathbf{Q}^{\mathtt{AR,T}}}^{\text{Temperature}}, \overbrace{\mathbf{Q}^{\mathtt{LT}(\mathtt{LT}),\mathtt{E}}, \mathbf{Q}^{\mathtt{AR/BAR,E}}}^{\text{Elongation}}\right)\\
\mathbf{Q}^{\mathrm{NB},\mathtt{LL}(\mathtt{LT})} &= \text{blockdiag}\left(\overbrace{\mathbf{Q}^{\mathtt{LL,T}}, \mathbf{Q}^{\mathtt{PD,T}}\mathbf{Q}^{\mathtt{AR,T}}}^{\text{Temperature}}, \overbrace{\mathbf{Q}^{\mathtt{LL}(\mathtt{LT}),\mathtt{E}}, \mathbf{Q}^{\mathtt{AR/BAR,E}}}^{\text{Elongation}}\right).
\end{aligned}$$

The observation matrix $\mathbf{F}^{\mathrm{NB}}$ is identical as $\mathbf{F}^{\mathrm{SB}}$ for both methods.
Parameters values for both regimes in SKF on bridge time series are

$$\begin{aligned}
\mathcal{P}^{\mathrm{B}} &= \left\{\overbrace{\phi^{\mathtt{AR,T}}, \sigma^{\mathtt{AR,T}}, \sigma_v^{\mathtt{T}}}^{\text{Temperature}}, \overbrace{\phi^{\mathtt{AR,E}}, \sigma^{\mathtt{AR,E}}, \sigma_v^{\mathtt{E}}, \beta^{\mathtt{PD,E|T}}, \beta^{\mathtt{AR,E|T}}, z^{11}, z^{22}}^{\text{Elongation}}\right\}\\
&= \{0.36981, 3.3065, 0.1, 0.74033, 0.0040317, 0.001, 0.019987,\\
&\quad 0.019689, 0.999999999, 0.999999999\}.
\end{aligned}$$

A.4 Dam data

This appendix presents the model matrices and their parameters used in the SKF analysis for the dam dataset consisting of eight time series.

A.4.1 Stationary model matrices in SKF

For automatically collected dam time series, the stationary model matrices, referred to as SDA, are

$$\begin{aligned}
\mathbf{A}^{\text{SDA}} &= \text{blockdiag}\left(\mathbf{A}^{\texttt{TcA}}, \mathbf{A}^{\texttt{KR}}, \mathbf{A}^{\texttt{AR/BAR}}\right) \\
\mathbf{F}^{\text{SDA}} &= \left[\mathbf{F}^{\texttt{TcA}}, \mathbf{F}^{\texttt{KR}}, \mathbf{F}^{\texttt{AR/BAR}}\right] \\
\mathbf{Q}^{\text{SDA,LT(LT)}} &= \text{blockdiag}\left(\mathbf{Q}^{\texttt{LT(LT)}}, \mathbf{Q}^{\texttt{KR}}, \mathbf{Q}^{\texttt{AR/BAR}}\right) \\
\mathbf{Q}^{\text{SDA,LA(LT)}} &= \text{blockdiag}\left(\mathbf{Q}^{\texttt{LA(LT)}}, \mathbf{Q}^{\texttt{KR}}, \mathbf{Q}^{\texttt{AR/BAR}}\right).
\end{aligned}$$

For manually collected dam time series, the stationary model matrices, referred to as SDM, are

$$\begin{aligned}
\mathbf{A}^{\text{SDM}} &= \text{blockdiag}\left(\mathbf{A}^{\texttt{TcA}}, \mathbf{A}^{\texttt{PD}}, \mathbf{A}^{\texttt{AR/BAR}}\right) \\
\mathbf{F}^{\text{SDM}} &= \left[\mathbf{F}^{\texttt{TcA}}, \mathbf{F}^{\texttt{PD}}, \mathbf{F}^{\texttt{AR/BAR}}\right] \\
\mathbf{Q}^{\text{SDM,LT(LT)}} &= \text{blockdiag}\left(\mathbf{Q}^{\texttt{LT(LT)}}, \mathbf{Q}^{\texttt{PD}}, \mathbf{Q}^{\texttt{AR/BAR}}\right) \\
\mathbf{Q}^{\text{SDM,LA(LT)}} &= \text{blockdiag}\left(\mathbf{Q}^{\texttt{LA(LT)}}, \mathbf{Q}^{\texttt{PD}}, \mathbf{Q}^{\texttt{AR/BAR}}\right).
\end{aligned}$$

A.4.2 Non-stationary model matrices in SKF

For automatically collected dam time series, the non-stationary model matrices, referred to as NDA, are

$$\begin{aligned}
\mathbf{A}^{\text{NDA}} &= \text{blockdiag}\left(\mathbf{A}^{\texttt{LA}}, \mathbf{A}^{\texttt{KR}}, \mathbf{A}^{\texttt{AR/BAR}}\right) \\
\mathbf{F}^{\text{NDA}} &= \left[\mathbf{F}^{\texttt{LA}}, \mathbf{F}^{\texttt{KR}}, \mathbf{F}^{\texttt{AR/BAR}}\right] \\
\mathbf{Q}^{\text{NDA,LA(LA)}} &= \text{blockdiag}\left(\mathbf{Q}^{\texttt{LA(LA)}}, \mathbf{Q}^{\texttt{KR}}, \mathbf{Q}^{\texttt{AR/BAR}}\right) \\
\mathbf{Q}^{\text{NDA,LT(LA)}} &= \text{blockdiag}\left(\mathbf{Q}^{\texttt{LT(LA)}}, \mathbf{Q}^{\texttt{KR}}, \mathbf{Q}^{\texttt{AR/BAR}}\right).
\end{aligned}$$

For manually collected dam time series, the non-stationary model matrices (NDM) are

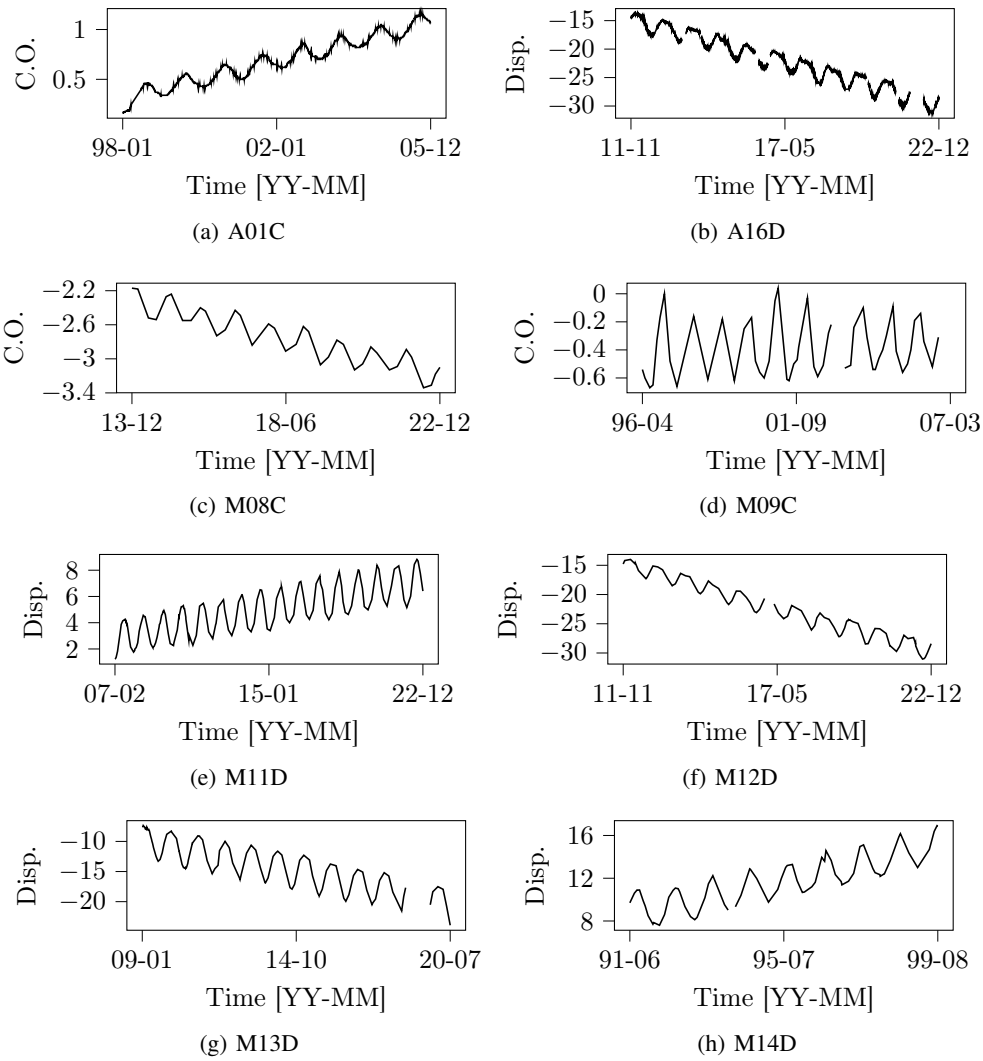
$$\begin{aligned}
\mathbf{A}^{\text{NDM}} &= \text{blockdiag}\left(\mathbf{A}^{\texttt{LA}}, \mathbf{A}^{\texttt{PD}}, \mathbf{A}^{\texttt{AR/BAR}}\right) \\
\mathbf{F}^{\text{NDM}} &= \left[\mathbf{F}^{\texttt{LA}}, \mathbf{F}^{\texttt{PD}}, \mathbf{F}^{\texttt{AR/BAR}}\right] \\
\mathbf{Q}^{\text{NDM,LA(LA)}} &= \text{blockdiag}\left(\mathbf{Q}^{\texttt{LA(LA)}}, \mathbf{Q}^{\texttt{PD}}, \mathbf{Q}^{\texttt{AR/BAR}}\right) \\
\mathbf{Q}^{\text{NDM,LT(LA)}} &= \text{blockdiag}\left(\mathbf{Q}^{\texttt{LT(LA)}}, \mathbf{Q}^{\texttt{PD}}, \mathbf{Q}^{\texttt{AR/BAR}}\right).
\end{aligned}$$

Parameters values for both regimes in SKF on dam time series series are

$$\begin{aligned}
\mathcal{P}^{\text{A01C}} &= \{\ell^{\texttt{KR}}, \phi^{\texttt{AR}}, \sigma^{\texttt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.98461, 0.91225, 0.0060635, 0.001, 0.99999, 0.99999\} \\
\mathcal{P}^{\text{A16D}} &= \{\ell^{\texttt{KR}}, \phi^{\texttt{AR}}, \sigma^{\texttt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.85519, 0.99412, 0.033654, 0.001, 0.999999, 0.999999\} \\
\mathcal{P}^{\text{M08C}} &= \{\phi^{\texttt{AR}}, \sigma^{\texttt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.34319, 0.036042, 0.001, 0.99999, 0.99999\} \\
\mathcal{P}^{\text{M09C}} &= \{\phi^{\texttt{AR}}, \sigma^{\texttt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.088333, 0.091778, 0.001, 0.99999, 0.99999\} \\
\mathcal{P}^{\text{M11D}} &= \{\phi^{\texttt{AR}}, \sigma^{\texttt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.39252, 0.23484, 0.001, 0.999999, 0.999999\} \\
\mathcal{P}^{\text{M12D}} &= \{\phi^{\texttt{AR}}, \sigma^{\texttt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.084549, 0.4741, 0.001, 0.999999, 0.999999\} \\
\mathcal{P}^{\text{M13D}} &= \{\phi^{\texttt{AR}}, \sigma^{\texttt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.32141, 0.39865, 0.001, 0.99999, 0.99999\} \\
\mathcal{P}^{\text{M14D}} &= \{\phi^{\texttt{AR}}, \sigma^{\texttt{AR}}, \sigma_v, z^{11}, z^{22}\} = \{0.16912, 0.34072, 0.001, 0.99999, 0.99999\}.
\end{aligned}$$

B Appendix: BDLM matrices and parameters

This appendix provides the plots of the eight dam datasets used in the experiments.



(a) A01C (b) A16D

(c) M08C (d) M09C

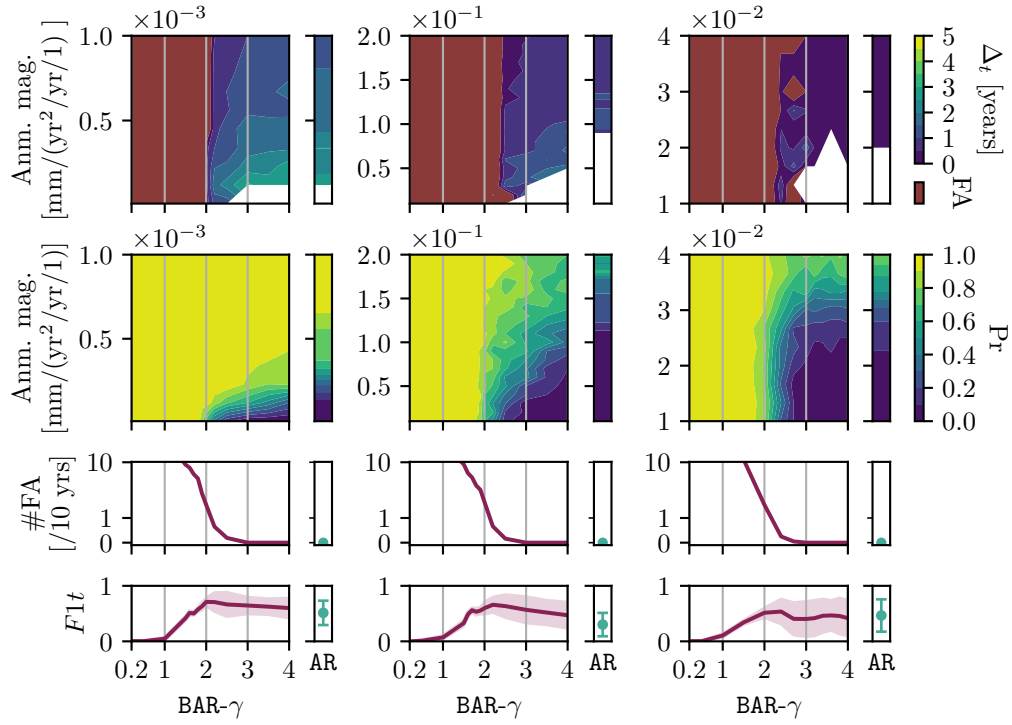
(e) M11D (f) M12D

(g) M13D (h) M14D

Figure 15: Eight time series measuring the cracking opening (C.O.) and displacement (Disp.) on a dam. They present missing data, different data acquisition frequencies, trends and periodic patterns. All the time series are assumed to be anomaly-free given that no structural deterioration was observed during these periods.

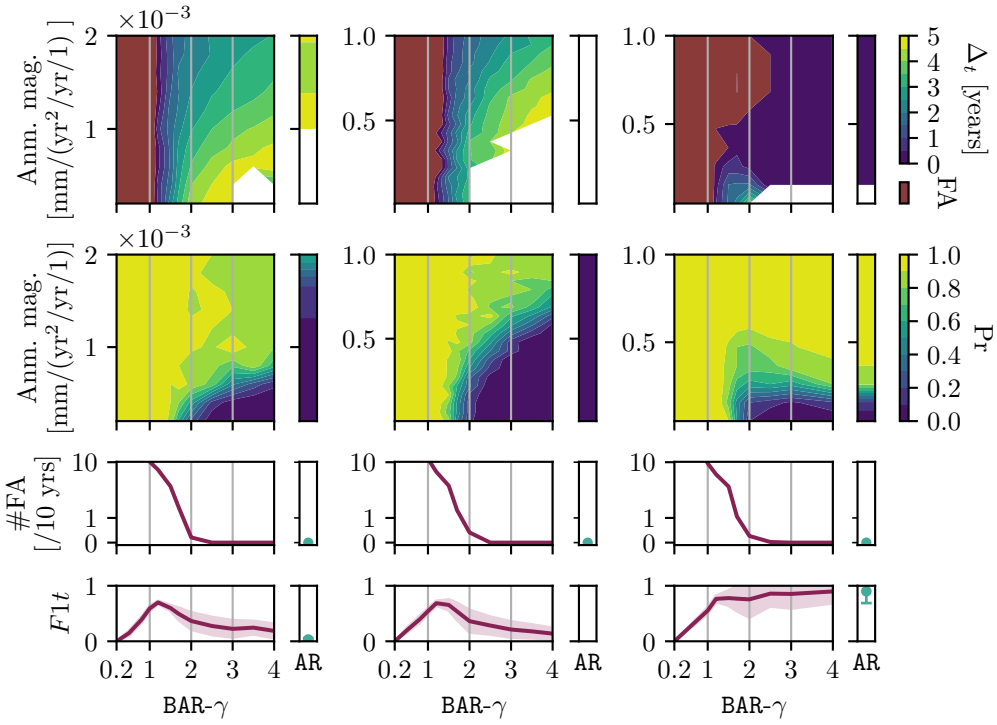
C Appendix: BAR-γ performance on synthetic time series in dam dataset

This appendix presents the BAR and the AR performance under different γ values evaluated on the synthetic datasets that are similar to seven time series measured on a dam. Three types of anomalies are introduced in the synthetic time series.



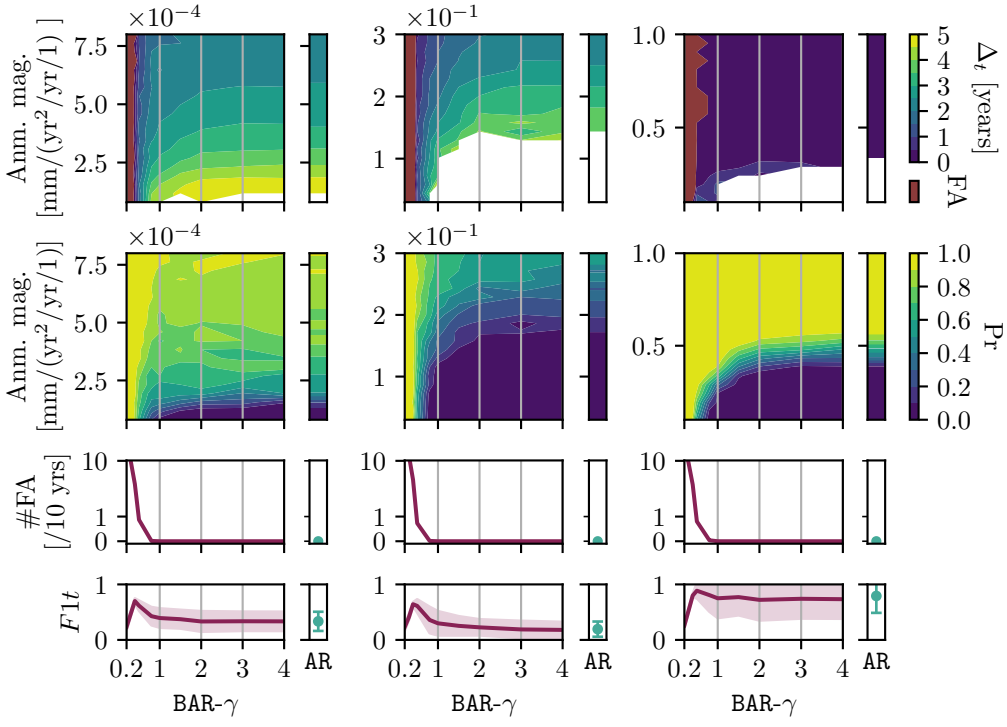
(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 16: BAR-γ and AR performance on synthetic time series A01C. Repetition $N = 50$



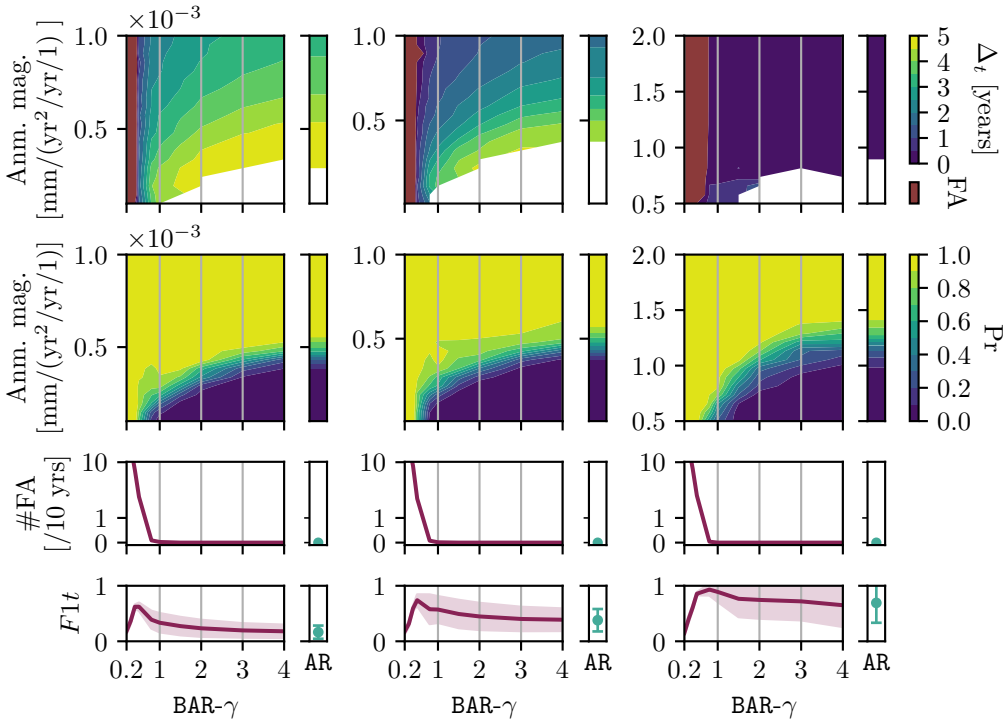
(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 17: BAR-γ and AR performance on synthetic time series A16D. Repetition N = 50



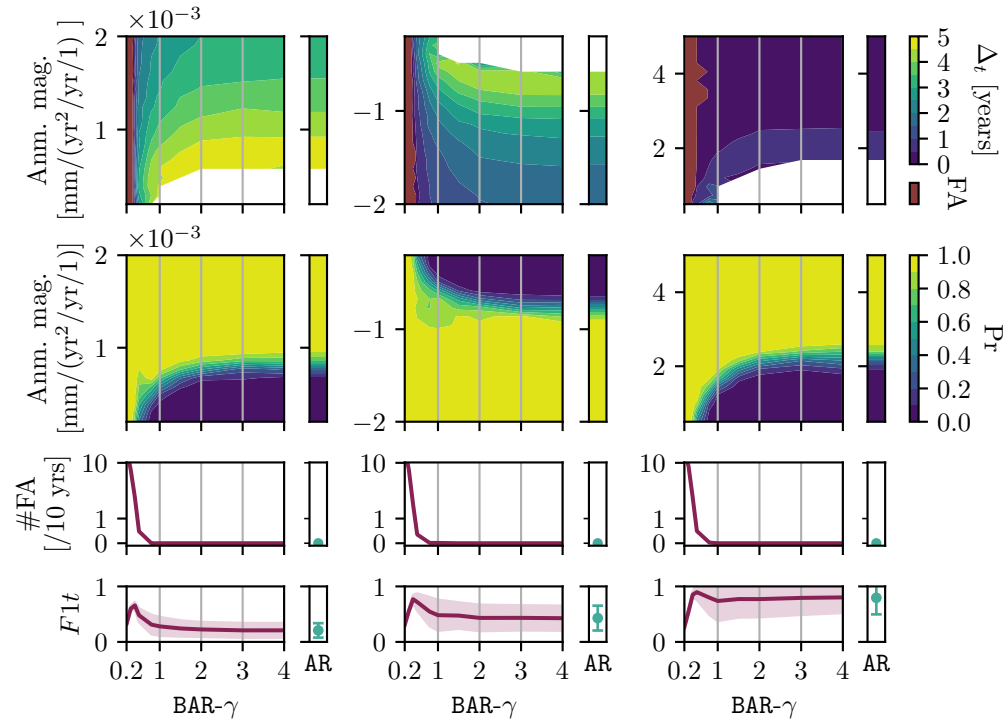
(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 18: BAR-γ and AR performance on synthetic time series M09C. Repetition N = 100



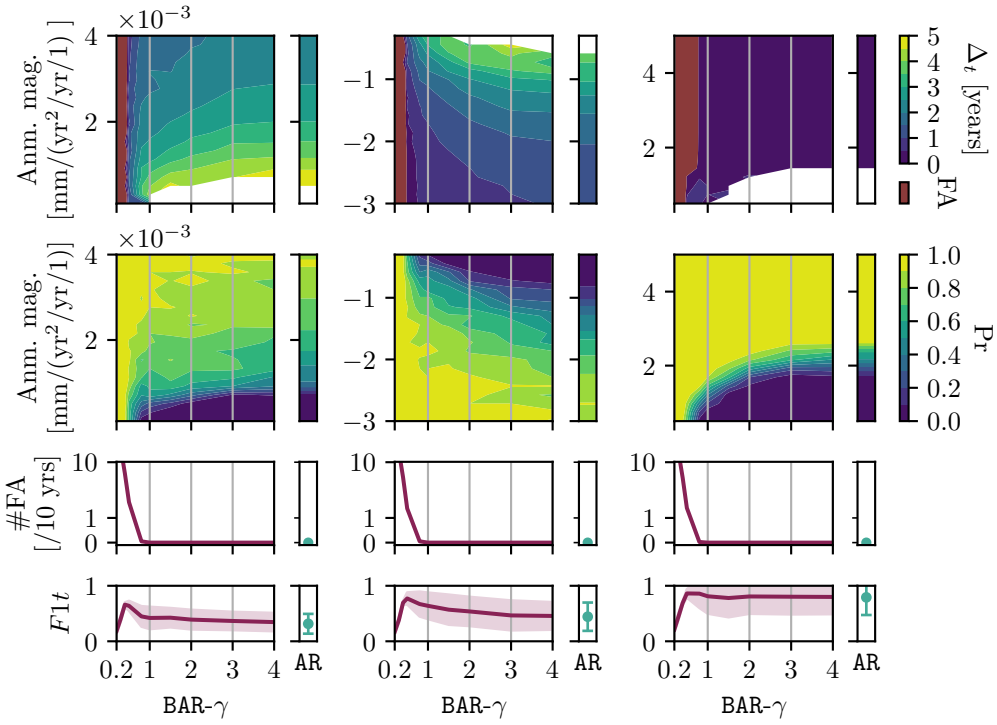
(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 19: BAR-γ and AR performance on synthetic time series M11D. Repetition N $= 100$



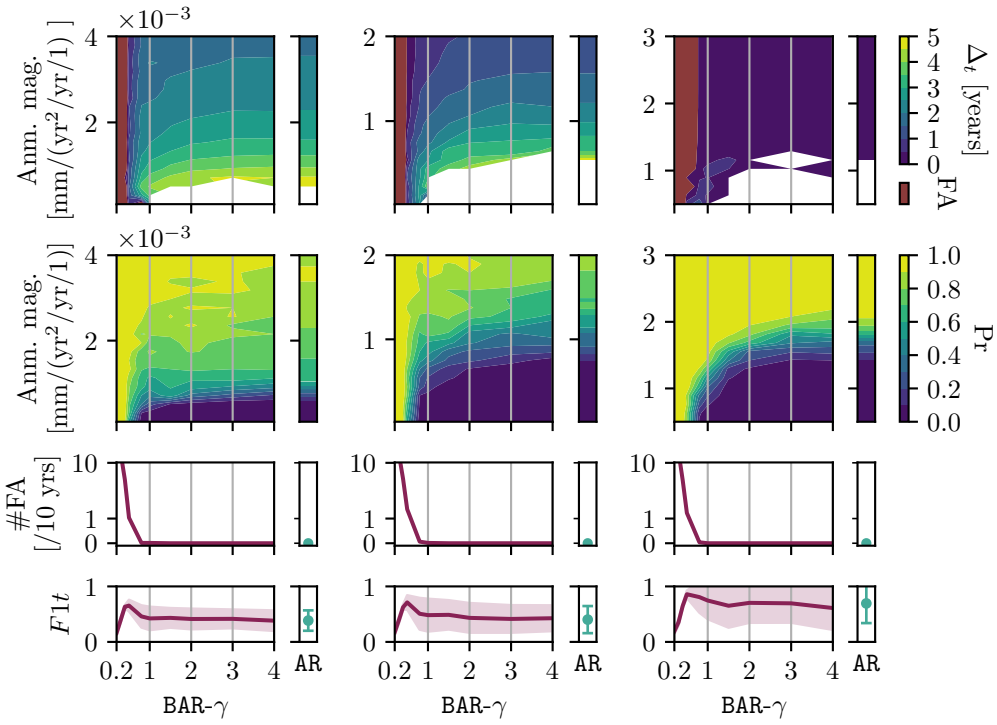
(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 20: BAR-γ and AR performance on synthetic time series M12D. Repetition N $= 100$



(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 21: BAR-γ and AR performance on synthetic time series M13D. Repetition N $= 100$



(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 22: BAR-γ and AR performance onsynthetic time series M14D. Repetition N $= 100$.

D Appendix: $F1t$ score summary on 8 dam time series

This appendix presents the $F1t$ scores for BAR and AR verified on the synthetic time series involving LT and LL anomalies.

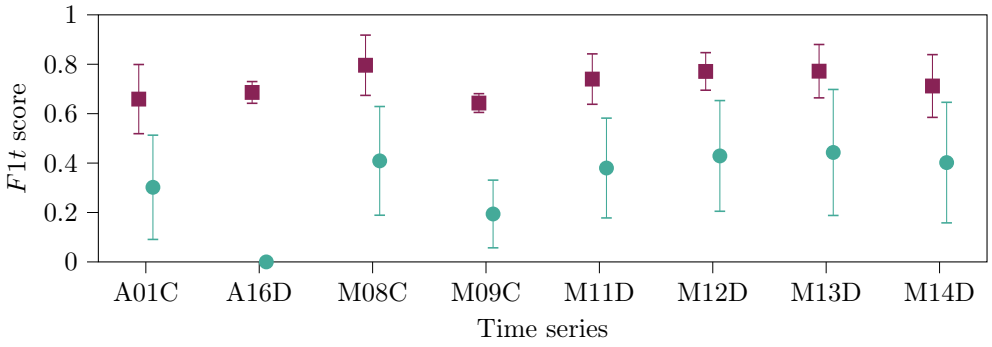


Figure 23: $F1t$ scores for BAR and AR verified on the synthetic datasets similar to eight real time series. LT anomalies are overlaid on the synthetic time series.

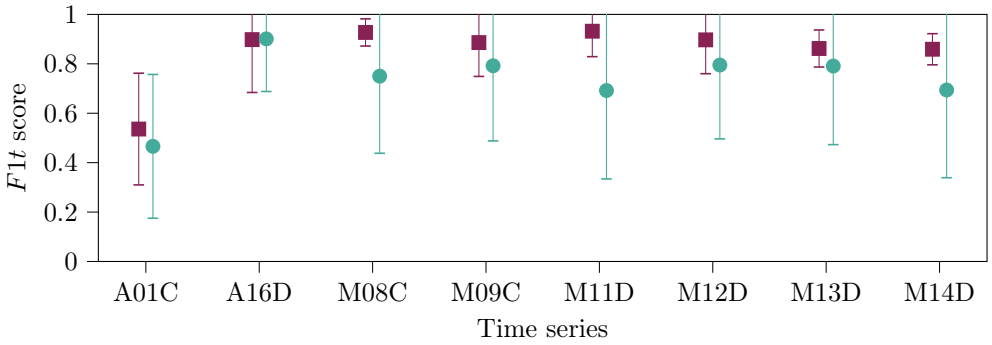
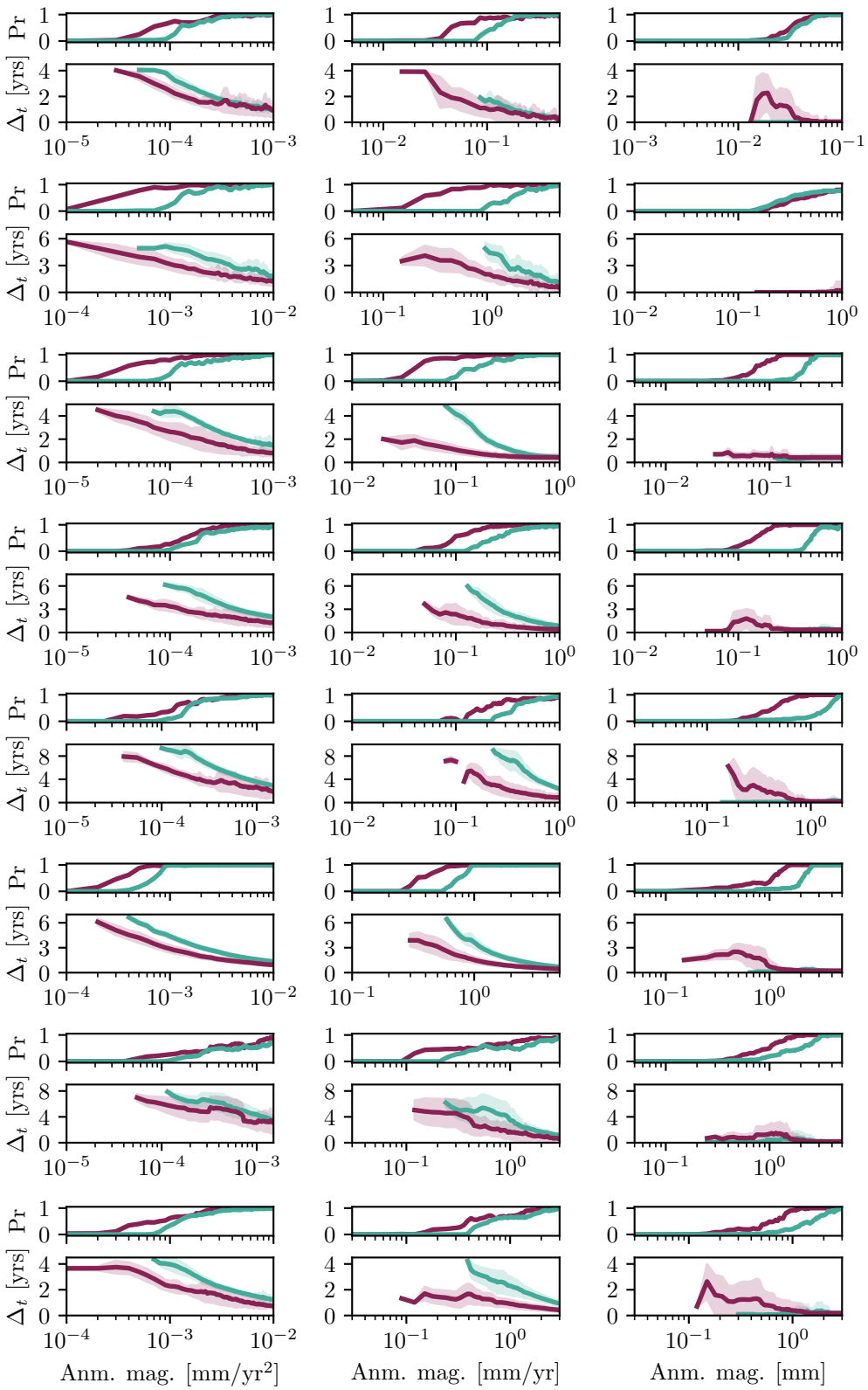


Figure 24: $F1t$ scores for BAR and AR verified on the synthetic datasets similar to eight real time series. LL anomalies are overlaid on the synthetic time series.

E Appendix: BAR-γ performance on real dam dataset

This appendix presents the BAR performance with selected γ on the eight real time series measured on a dam involving three types of anomalies. For the plots of $\Pr$, the higher the curve, the better, while for the plots of Δ_t, the lower and the more to the left, the better. Repetition is 30 for the automatically collected time series and 100 for the manually collected ones.



(a) LA anomaly (b) LT anomaly (c) LL anomaly

Figure 25: From top to bottom, each row shows the performance of BAR-γ and AR on eight real time series of A01C ($\gamma = 2.8$), A16D ($\gamma = 2.0$), M08C ($\gamma = 0.4$), M09C ($\gamma = 0.4$), M11D ($\gamma = 0.6$), M12D ($\gamma = 0.7$), M13D ($\gamma = 0.9$) and M14D ($\gamma = 0.5$) involving three types of anomalies.