RESEARCH ARTICLE

Damage Detection for Structural Health Monitoring Using Reinforcement and Imitation Learning

Shervin Khazaeli[a] and James-A. Goulet[a]

[a]École Polytechnique de Montréal, Quebec, Canada

ARTICLE HISTORY
Compiled January 3, 2024

Abstract
Structural damages are responsible for expenses associated with maintaining the safety and serviceability of infrastructures. Detecting damages is difficult because they often develop over years affecting structural responses in orders of magnitudes smaller than external effects such as temperature. When a damage occurs, structural responses depart from a normal condition to an abnormal one, which is referred to as an anomaly. Existing anomaly detection methodologies lack a mechanism to quantify the probability of rightfully detecting anomalies as a function of the anomaly's characteristics, e.g., duration and magnitude, and associate them with the severity of structural damages. This paper proposes a framework addressing these challenges by relying on Bayesian dynamic linear models as well as reinforcement and imitation learning approaches. The former allows separating the changes in the structural responses from the ones caused by external effects, while the latter two enable incorporating information obtained from the changes in the structural responses for detecting anomalies. The proposed methodologies are validated using measurements collected on three instrumented bridge spans in Canada. The results show a good performance of the methods proposed in detecting structural damages with different severity levels and lay the foundation for further applications for other civil infrastructures.

KEYWORDS
Anomaly detection; damage detection; reinforcement learning; imitation learning; Bayesian dynamic linear models; structural health monitoring; bridge

1 Introduction

The deterioration and degradation of civil infrastructures are common issues where they are responsible for societal expenses affecting economical growth (ASCE, 2021; Black, 2022; BluePlan, 2019). However, limited financial resources prohibit the maintenance of all infrastructures at the same time. *Structural health monitoring* (SHM) aims at providing information about structural conditions, e.g., normal and abnormal, for decision makers to support the maintenance of structures (Balsamo, Betti, & Beigi, 2014; C. R. Farrar & Worden, 2010). SHM uses sensors in order to collect measurements about the structural responses and their surrounding environment. The measurements are then interpreted to infer the condition of a structure over time. In this regard, an *anomaly*, is defined as a change in the structural responses from a *reference* condition that is referred to as the *normal condition*. Changes from the normal condition of a

CONTACT Shervin Khazaeli. Email: shrvin.khazaeli@gmail.com

structure could be associated with various *damages* that can develop over a period spanning several years. Therefore, a SHM system should focus on the long-term monitoring of structures in order to detect anomalies at an early stage before they become critical. Note that anomalies may not always indicate a structural damage. They can result from various factors, including but not limited to, intervention actions like repair work or sensor-related issues such as drift. Therefore, while anomalies signify a broader range of changes within a structural responses, the central focus of this work is on the ones caused by structural damages. In this regard, true detection is referred to correctly identifying an anomaly when it happens. On the other hand, a false detection, i.e., a *false alarm*, is referred to as wrongfully identifying the presence of an anomaly. Therefore, detecting anomalies while avoiding false alarms is the primary objective when developing anomaly detection methods in the context of structural health monitoring.

A SHM system requires building a model to analyze structural responses; *Data-driven* models (C. Farrar & Worden, 2013; Nassif, Talib, Nasir, & Dakalbab, 2021; Yu & Sun, 2020) solely rely on structural responses without the necessity of building a physics-based model. The vast majority of existing data-driven anomaly detection approaches are based on a threshold-based alarm-triggering paradigm. In general, these methods create a model from a training dataset representing the normal signature of a structure, and the newly collected data are analyzed and tested against the model to identify whether the predictions fit the test data within an acceptable threshold. The reader is referred to (Bull, Worden, Manson, & Dervilis, 2018; C. Farrar & Worden, 2013; Li, Morgantini, & Betti, 2023; Malekzadeh, Atia, & Catbas, 2015; Mohammadi Ghazi & Büyüköztürk, 2016; Ying et al., 2013) for examples of threshold-based anomaly detection methods. The key factor that limits the scalability of a SHM system for detecting anomalies is the changes in the conditions surrounding a structure causing frequent false alarms that need to be mitigated by raising the detection threshold, which in return reduces the detectability of anomalies. Another limitation is that most existing methods have not been developed or validated with real operational conditions, where a full-scale structure is studied in its operational environment, and where it is subject to external loading and effects such as traffic, temperature, and weather changes(C. Farrar & Worden, 2013; Li et al., 2023; Mohammadi Ghazi & Büyüköztürk, 2016). Moreover, existing methods typically study punctual and sudden damage scenarios whereas in practice we are interested in detecting anomalies that have a small magnitude, and that develop over years.

One method that has tackled several of the limitations concerning the threshold-based approaches for anomaly detection relies on the Bayesian dynamic linear models (BDLM) (Goulet, 2017); It is a particular type of state-space model that enables decomposing observations into the superposition of *reversible* and *irreversible* responses. Reversible responses are typically due to external effects such as temperature and loading that are modeled using components such as a kernel regression (Nguyen, Gaudot, Khazaeli, & Goulet, 2019) component. Irreversible responses are associated with the underlying structural responses; They represent the *degradation's kinematics* consisting of the structure's degradation, degradation's speed, and degradation's acceleration. Nguyen and Goulet (2018) extended the BDLM approach in order to detect anomalies using switching Kalman filtering (SKF). This method uses two kinematics identifying stationary and non-stationary regimes that represent the normal and abnormal state of a structure, respectively. The probability of each regime is estimated and used as a basis for detecting anomalies when a normal regime switches to the abnormal one. However, as pointed out by Khazaeli, Nguyen, and Goulet (2021) making decisions

only based on the probability values of non-stationary regimes likely results in false or missed alarms, especially for cases with low probability values. In addition, the SKF-based anomaly detection method does not account for anomaly characteristics such as the anomaly's *magnitude*, its *duration* and *time to detection*; Anomalies affect the measurements collected on a structure by creating a deviation in the behavior of structures associated with the normal condition. This deviation can be quantified in terms of its magnitude, which is referred to as the *anomaly magnitude*. In addition, such a deviation occurs over a period of time known as the *anomaly duration*. The *time to detection* refers to the time that an SHM method takes to identify the presence of an anomaly after its starting time.

One of the methods that addressed the above-mentioned limitations is based on employing the reinforcement learning approach (Khazaeli et al., 2021), where a decision maker, called the *agent*, makes decisions by taking *actions* of triggering or not triggering an alarm in an *environment* issued from the SKF-BDLM analysis that includes the degradation's speed and the probability of the non-stationary regime. The main limitation of this method is that it does not account for the anomaly duration and it does not associate an anomaly's magnitude with the severity level of structural damages. Prior to this work, reinforcement learning has been used in the field of active control where a controller takes some actions based on feedback from an instrumented structure (Korkmaz, 2011; Soong & Cimellaro, 2009). Few researches are dedicated to precisely formulating structural control problems in the form of an RL framework (Adam & Smith, 2008; Khalatbarisoltani, Soleymani, & Khodadadi, 2019) or employing it, in particular, for detecting anomalies (Huang, Wu, Zuo, Pei, & Min, 2018; Yu & Sun, 2020). Although these works have shown the advantages of using reinforcement learning, the time series employed do not account for the specificities of SHM data collected on civil infrastructures.

One of the limitations in the context of reinforcement learning is the requirement to design a reward function representing the task in hand, e.g., trigger an alarm or not, that an agent requires to accomplish. To overcome this limitation, Pomerleau (1989) proposed an approach that consists in learning from demonstrations, also known as *imitation learning* (IL). In IL, an expert, usually a human, provides the optimal policy within a set of demonstrations associated with the desired behavior. Then, the agent learns the optimal policy by imitating the expert's policy. There are two approaches towards imitation learning in the literature. In the first approach, an agent learns to directly imitate an expert's policy. The simplest form of this approach is called *behavior cloning*. In this setup, a classifier is trained from a demonstration set in order to predict the expert's behavior. The demonstration set consists of state-action pairs that is derived from an expert's policy. Although behavior cloning employs supervised learning techniques, it is distinct in its application. Unlike generic supervised learning that broadly applies to predicting outputs from inputs across various contexts, imitation learning is specifically tailored to mimic an expert's behavior in sequential decision-making tasks. This distinction is crucial because behavior cloning is shown to yield poor results when the demonstrations issued by the expert are not uniformly sampled from the environment. To mitigate this limitation, Ross, Gordon, and Bagnell (2011) introduced an algorithm known as *data aggregation* (DAgger). As such, demonstrations are used incrementally whenever needed; An agent learns to imitate the expert's policy with given demonstrations, and whenever there are states which are not in the current demonstrations set, the agent makes a query about optimal actions from the expert for these new states. The reader is referred to Hussein, Gaber, Elyan, and Jayne (2017) for a comprehensive review of existing methods relying on directly imitating the expert's

policy. In this work, while we employ behavior cloning techniques, our focus remains on imitation learning in the context of task-specific expert policy mimicry, hence our preference for using the term imitation learning over supervised learning. In the second approach, which is not considered in the current work, an agent learns to indirectly imitate the expert's policy by first learning the reward function and then learning the optimal policy. This approach is known as inverse reinforcement learning and the reader is referred to Arora and Doshi (2021) for a detailed review of existing methods.

This paper develops a method for characterizing probabilistically the anomaly detectability of a SHM system based on both reinforcement and imitation learning approaches. The characterization needs to be presented probabilistically not only as a function of an anomaly's magnitude, but also as its duration as well as the time required to detect an anomaly. Furthermore, it develops a methodology to associate the magnitudes of anomalies to the severity levels of structural damages in order to discriminate the damages that are detectable from the ones that are not. In this regard, Section 2 presents a glimpse at the steps taken in this paper in order to address the above-mentioned objectives. Section 3 presents the formalism of BDLM and its extension for detecting anomalies using regime switching, i.e., SKF-BDLM. Section 4 presents the core idea behind anomaly detection using reinforcement and imitation learning approaches; First, we present a procedure for simulating an environment in which an agent interacts. Afterwards, we present the formalism of the reinforcement and imitation learning approaches. Section 5 introduces a procedure for quantifying anomaly detectability with respect to false alarms and the probability of rightfully detecting anomalies. In addition, it presents a method to associate the detected anomalies with structural damages. Section 6 examines the methods proposed on three instrumented spans of a bridge in Canada and compares with each other the results obtained from both anomaly detection methods proposed, and with the results obtained from decision-making solely based on the probability of the non-stationary regime.

2 Damage Detection Framework

To better contextualize the methodological approach and corresponding result used in this research, this section presents the general framework for quantifying anomaly detection and associating detected anomalies to structural damages. The general flow of the framework is illustrated in Figure 1. The arrows indicate the input-output of eight distinct elements. These elements are categorized in four general stages, namely measurement, state inference, decision making, and damage detection. The following subsections present the detailed descriptions of these stages and their elements.

[Figure 1 about here.]

2.1. Measurement

The goal of this stage is to present the bridge, sensor technology, and the measurements collected from instrumented spans. The first three elements, from left to right, shown in Figure 1, corresponding to this stage are: (i) The sensor technology used on the bridge, including the sensors types and their locations, (ii) The instrumented spans involving their general layout and sensors configuration, and (iii) The data collected. We provide these information in this section as the authors think that they help the reader to easily follow the next steps.

Starting in 2019, three spans of a bridge in Canada have been instrumented. We refer to these spans as "$4-5$", "$5-6$", and "$6-7$". Excepted for repairs due to the structural steel reinforcement program, the three spans are identical with respect to their superstructures, geometry, and boundary conditions. Figure 2 shows a typical span of the bridge divided in seven sections that are number from 1 to 7. The span measures 24 meters in width and 45 meters in length. The superstructure of the span comprises a concrete deck and a steel structure. Furthermore, each span is supported by fixed supports at section 7, and by expansion joints and bearing devices at section 1.

[Figure 2 about here.]

The instrumentation configuration for each span is identical and consists in three inclinometers [d°], an extensometer [mm], and a thermometer [C°], which are installed on the downstream side of the span. The two longitudinal inclinometers are located on the sections 1 and 3 and measure rotations around the y axis. The lateral inclinometer is located on the section 3 and measures rotations around the x axis. Inclinometers are uniaxial employing a force-balance measurement method. They are capable of detecting a change in angle as small as 0.00005° within a temperature range of -40°C to $+80$°C. The 2 meters long extensometer is located on the lower cord of the span between the sections 3 and 5 (centered at section 4). It measures elongations along the x axis with a resolution capable of detecting changes as small as 0.00313 mm, operating effectively in a temperature range from -20°C to $+80$°C. Thermometers offers a resolution of 0.1°C and an accuracy of ± 0.5°C.

We define a generic name for each sensor in order to readily identify a sensor's type and location; Each name consists of the span's name followed by an abbreviation, that is $\mathtt{I}_1$ for a longitudinal inclinometer on section 1, $\mathtt{I}_3$ for a longitudinal inclinometer on section 3, $\mathtt{I}_\mathrm{L}$ for a lateral inclinometer on section 3, $\mathtt{E}$ for an extensometer on section 4, and $\mathtt{T}$ for a thermometer on section 4. For instance, the longitudinal inclinometer on the section 1 on the span $5-6$ is referred to "$5-6:\mathtt{I}_1$".

The data acquisition system consists of central acquisition units, which form the core of the panel, as well as interfaces for sensor connectivity and channel multiplication, and other components for power and communication. The data acquisition started in May 2019 with an average period of acquisition of 10 min. This research analyzes the data from August 2019 until November 2021; We discarded the first three months of the data due to multiple sensor calibration issues that resulted in missing or improperly collected measurements. We use weekly averaged data for all measurements in the current research as it facilitates and speeds up the analysis without limiting the interpretation capacity. Figures 3a and 3b show the weekly averaged data corresponding to the extensometer and temperature measurements on the span $5-6$ with the respective labels $5-6:\mathtt{E}$ and $5-6:\mathtt{T}$.

[Figure 3 about here.]

2.2. State Inference

The goal of this stage is to interpret the preprocessed data in order to remove the external effects from the measurements collected on a structure. The proposed interpretation technique uses BDLM to create empirical models. An empirical model learns to reproduce the behavior of a system, e.g., a bridge span, from the measurements

recorded on it. In contrast to physics-based models, e.g., finite element model (FEM), the empirical models do not contain information about the geometry and a structure's properties. Therefore, empirical models allow the development of a single generic model that can then be used to interpret newly generated data from any bridge. These generated data are evaluated by using the empirical models developed in conjunction with anomalies that are characterized by wide range of magnitude and duration. As shown in Figure 1, the elements corresponding to this stage are as follows: (iv) Bayesian dynamic linear model in order to build generic empirical models, and (v) Measurements decomposition consisting in the underlying structural responses and the external effects.

2.3. Decision Making

As it's discussed in details in Section 5, the main objective of this stage is to quantify the anomaly detectability of each sensor using the underlying structural responses obtained from the previous step. To this end, the first step is to develop an anomaly detection method capable of operating for a wide range of anomalies. The proposed anomaly detection method uses the reinforcement learning (RL) and imitation learning (IL) approaches. In the reinforcement learning approach, instead of defining rule-based alarm-triggering thresholds, we let agents figure out how to optimally trigger alarms while satisfying our preference constraints encoded in a reward function for true anomaly detections, false alarms, and missed alarms. On the other hand, in the imitation learning approach, an agent learns the optimal policy by imitating a demonstration set for which the corresponding optimal policy is issued beforehand. With the anomaly detection method developed, the detectability of anomalies is quantified through an annual false alarm ratio and a probability of true detections that is a function of the anomaly's magnitude, its duration and the time required to detect an anomaly. The former expresses the number of false alarms that the anomaly detection method yields, and the latter is its capacity at rightfully detecting anomalies. Upon addressing the objectives of this stage, the anomaly detectability quantification will serve for the next stage. From Figure 1, the associated elements are: (vi) Reinforcement learning or imitation learning approaches as the basis for detecting anomalies, and (vii) Alarm triggering policies which dictate the optimal action, in this case triggering an alarm in the presence of an anomaly or not triggering an alarm otherwise.

2.4. Damage Detection

The ensemble of methods developed in the previous stages enables associating the magnitudes of detected anomalies to the severity levels of structural damage scenarios, and evaluating the capacity of different sensor in discriminating between these damages. In this regard, different damage scenarios are proposed and modeled using finite element method. The severity of these damages are defined as the changes in structural responses at the location of sensors. The crux of our damage detection method lies in comparing the minimum detectable anomaly magnitude obtained from previous step with the FEM-derived changes in structural responses. This comparison is critical for accurately identifying and quantifying damage scenarios; It enables discriminating between the damage scenarios that will be detectable from those that will not, thus setting realistic expectations with respect to the capacity of a structural health monitoring system. In Figure 1, this stage is associated with the last element, i.e., (viii) damage detection.

3 Empirical Model Estimation

BDLMs use generic sub-components for representing the underlying behaviors for the observation vector $\mathbf{y}_t = [y_1\ y_2\ \cdots\ y_{\mathtt{Y}}]_t^\intercal$ that consists in $\mathtt{Y}$ observations at a time t. Sub-components are assembled in order to form the *hidden state* vector $\mathbf{x}_t = [x_1\ x_2\ \cdots\ x_{\mathtt{X}}]_t^\intercal$ consisting in $\mathtt{X}$ hidden states; At a time t, hidden state variables are estimated based on the information obtained from two linear models: A *transition model* that predicts the evolution of the hidden states variables across time, and an *observation model* that updates the knowledge of the predicted values of the hidden state variables according to the observations. The transition and observation models are defined following

$$\begin{aligned} \mathbf{x}_t &= \mathbf{A}_t\mathbf{x}_{t-1} + \mathbf{w}_t, \qquad \mathbf{w} : \boldsymbol{W} \sim \mathcal{N}(\mathbf{w}; \mathbf{0}, \mathbf{Q}_t), \\ \mathbf{y}_t &= \mathbf{C}\mathbf{x}_t + \mathbf{v}_t, \qquad \mathbf{v} : \boldsymbol{V} \sim \mathcal{N}(\mathbf{v}; \mathbf{0}, \mathbf{R}), \end{aligned} \tag{1}$$

where, $\mathbf{A}_t$ and $\mathbf{C}$ are the transition and observation matrices, respectively. $\mathbf{w}_t$ and $\mathbf{v}_t$ are the respective transition and observation model error vectors following Gaussian distributions with zero means and the covariance matrices $\mathbf{Q}_t$ and $\mathbf{R}$. The formulation of the model matrices $\mathcal{M} = \{\mathbf{A}_t, \mathbf{C}, \mathbf{Q}_t, \mathbf{R}\}$ is based on the assembly of generic sub-components which can model behaviors such as a structure's degradation, a degradation's speed, a degradation's acceleration, periodic phenomena, and so on. The reader is referred to (Goulet, 2017; Goulet & Koo, 2018; Murphy, 2013; West & Harrison, 1999) for a complete review of the different sub-components available. Note that the subscript t in Equation 1 indicates that the corresponding matrix is time variant. The formulations of matrices used in this work are presented in Appendix A.

3.1. Regime Switching

This section provides the concept of anomaly detection using regime switching (Nguyen & Goulet, 2018). Consider Figure 4a where the observed structural responses y_t are shown by gray circles. Figures 4b-4d illustrate the baseline ($\mathtt{B}$) generic sub-components including the structure's degradation $x_t^{\mathtt{B}}$, degradation's speed $\dot{x}_t^{\mathtt{B}}$, and degradation's acceleration $\ddot{x}_t^{\mathtt{B}}$. In order to simplify the illustration, the uncertainty around the estimations is not shown in Figures 4b-4d. The structure's degradation, its speed, and its acceleration are hidden state variables that jointly define the *degradation's kinematics* of the system. They are irreversible structural responses describing the evolution of the hidden state variables vector $\mathbf{x}_t^{\mathtt{B}} = \left[x_t^{\mathtt{B}}\ \dot{x}_t^{\mathtt{B}}\ \ddot{x}_t^{\mathtt{B}}\right]^\intercal$.

[Figure 4 about here.]

From Figure 4c, two regimes are identified: a *trend-stationary* regime, where the degradation's speed remains constant, and a *non-stationary regime*, where the degradation's speed varies over time. The non-stationary regime is shown by the red dashed lines in Figures 4b–4d. Note that in this work, the terms trend-stationary and stationary are used interchangeably. In Figure 4c, the dynamic model starts with a constant degradation speed during the stationary regime and then switches to a non-stationary regime with a locally constant acceleration before it returns to the stationary regime with a different constant speed. In BDLMs, two dynamic models are built corresponding to the two regimes for which the *probabilities* are evaluated using the switching Kalman filter (SKF) (Nguyen & Goulet, 2018). In this configuration, the task of *anomaly detection* consists in identifying the presence of a switch from a stationary regime to a

non-stationary one. The BDLM models non-stationary system responses by estimating at each time t, the probability of each regime along with their hidden state values. Each regime is defined by its own model matrices, that is $\mathcal{M}^{\mathtt{s}}$ for a stationary regime and $\mathcal{M}^{\mathtt{ns}}$ for a non-stationary one. Consider Figure 4b where a system that has two possible regimes are identified by the regime state variable $s_t \in \{1, 2\}$ and the corresponding probability $\pi_t^{s_t}$, where 1: stationary ($\mathtt{s}$) or 2: non-stationary ($\mathtt{ns}$). Therefore, at a time t, the objective of the SKF is to estimate: (i) the hidden state variables vector $\mathbf{x}_t$ and (ii) the regime state probability $\pi_t^{s_t}$ associated with each regime state s_t. The BDLM procedure for hidden states and probabilities estimation can be summarized by

$$(\boldsymbol{\mu}_{t|t}^{i(j)}, \boldsymbol{\Sigma}_{t|t}^{i(j)}, \mathcal{L}_t^{i(j)}, \pi_t^{i(j)}) = \text{BDLM}(\boldsymbol{\mu}_{t-1|t-1}^{j}, \boldsymbol{\Sigma}_{t-1|t-1}^{i}, \pi_{t-1|t-1}^{j}, \mathbf{y}_t, \mathbf{A}_t^j, \mathbf{C}_t^j, \mathbf{Q}^{i(j)}, \mathbf{R}^j), \tag{2}$$

where, at a time t, $\boldsymbol{\mu}_{t|t} \equiv \mathbb{E}[\boldsymbol{X}_t|y_{1:t}]$ and $\boldsymbol{\Sigma}_{t|t} \equiv \text{cov}[\boldsymbol{X}_t|y_{1:t}]$ respectively denote the expected value vector and covariance matrix of the hidden states vector. Note that, both quantities are conditional on all the observations $y_{1:t}$ ranging from time 1 up to time t. In addition, $\mathcal{L}_t^{i(j)}$ is the marginal likelihood, and the superscript inside the parentheses $^{i(j)}$ indicates the current regime state $s_t = j$ at time t given the regime state $s_{t-1} = i$ at time $t-1$. In addition, a BDLM involves a set of parameters $\mathcal{P}$ that depends on the sub-components used and that is learned from the observed data (Goulet, 2020; Nguyen & Goulet, 2018). Section 6.1 presents the sub-components and corresponding parameters used in this work.

4 Anomaly Detection

In this work, the information used for identifying the presence of an anomaly are the expected value of the degradation's speed $\mu_t^{\mathtt{T}}$ known as the *trend* ($\mathtt{T}$), and the probability of a non-stationary ($\mathtt{ns}$) regime $\pi_t^{\mathtt{ns}}$ (Khazaeli et al., 2021). As mentioned earlier, anomalies are rare events, so it is not possible to draw probabilistic conclusions by only relying on real anomalies that occurred on a structure. To address this limitation, we present a procedure for generating synthetic time series representing the structural responses subject to a wide range of anomalies. Furthermore, we present a simulated environment for decision-making based on the decomposition of the synthetic time series using Bayesian dynamic linear models (BDLM). Finally, we provide the formalism of detecting anomalies in the context of the reinforcement and imitation learning approaches.

4.1. Anomaly Function

We define an anomaly as a switch from a stationary to a non-stationary regime. Such a transition is simulated by an anomaly function that represents a change in the structure's degradation. It is defined in the form of a continuous logistic function as

$$f_{\text{a}}(t; \mathcal{J}) = \frac{h_{\text{a}}}{|h_a|} \cdot \frac{|h_a| + 2\epsilon_a}{1 + e^{-k(t-t_{\text{c}})}}, \tag{3}$$

where $k = \frac{2}{w_a} \ln\left(\frac{|h_{\text{a}}|}{\epsilon_{\text{a}}} + 1\right)$ is its steepness. The anomaly function parameters are generated from $\mathcal{J} = \{h_a : H_a,\ w_{\text{a}} : W_a,\ t_{\text{c}} : T_c\}$, a set of random variables consisting in the anomaly's magnitude h_{a}, its duration w_a, and the center time t_{c} of the anomaly

function. The *anomaly tolerance* $f_{\mathrm{a}}(t_{\mathrm{s}}) = \epsilon_a$ is the magnitude associated with the beginning time t_{s} of a non-stationary regime and $w_0 > 0$ is the number of initial time steps without any anomaly. In addition, the anomaly window w_{a} represents the duration of an anomaly and is defined based on the center time t_{c}, so the anomaly's end and start times are $t_{\mathrm{e,s}} = t_{\mathrm{c}} \pm w_a/2$. Note that the formalism of the logistic function presented in Equation 3 is similar to the one presented by Khazaeli et al. (2021). However, in this work, the logistic function is used to represent a change in the degradation instead of a change in the degradation's speed. This allows for a direct comparison between the magnitude of anomalies, h_{a}, and changes in structural responses due to damage scenarios as will be discussed in Section 5.4. Figure 5a illustrates an example of an anomaly function simulated to generate the synthetic time series shown in Figure 5c. The anomaly's magnitude is $h_a = 0.011$ [unit], its duration is $w_a = 1$ year, and the two vertical dashed lines indicate the beginning and the end of the anomaly. For a better illustration, the anomaly duration is indicated by a shaded area. Note that the unit of the anomaly function values is the same as the measurements that the anomaly function is used for.

[Figure 5 about here.]

Figures 5b and 5d show the trend expected value $\mu^{\mathtt{T}}$ and the probability of the non-stationary regime $\pi^{\mathtt{ns}}$ obtained from the BDLM analysis of the simulated time series. The blue and red lines in Figure 5 indicate the normal and abnormal time steps, respectively.

4.2. Simulated Environment for Decision Making

The process of simulating an environment's state-space involves generating synthetic abnormal structural responses and then estimating hidden-state variables and regime probabilities; Generating abnormal structural responses is carried out by simulating synthetic time series using the BDLM matrices associated with a normal regime, i.e., $\mathcal{M}^{\mathtt{s},\mathrm{obs}}$, and afterwards, adding the values corresponding to a realization of the anomaly function $f_{\mathrm{a}}(t;\mathcal{J})$ to the simulated synthetic time-series. This process is repeated to build synthetic times series containing a wide range of anomalies. Choosing a logistic function for an anomaly is an effective approach is due to its infinite differentiability and integrability are key for accurately modeling the kinematics of anomalies, capturing the changes in local level, local trend, and local acceleration, i.e., degradation's kinematics. Furthermore, the logistic function's capability to represent non-linear behaviors and smooth transitions is essential for modeling the underlying changes in time series indicative of long-term structural damages. In particular, the parameters presented in Equation 3 allows for the early detection of subtle changes that are developed over several years. The simulation of synthetic structural responses $\mathbf{y}_t^{\mathrm{sim}}$ is summarized by

$$\mathbf{y}_t^{\mathrm{sim}} = \mathrm{SIM}(\boldsymbol{\mu}_{t|t}^{\mathrm{obs}}, \boldsymbol{\Sigma}_{t|t}^{\mathrm{obs}}, \mathcal{M}^{\mathtt{s},\mathrm{obs}}, \mathcal{P}^{*,\mathrm{obs}}, f_{\mathrm{a}}(t;\mathcal{J})), \tag{4}$$

where, the superscripts $^{\mathrm{obs}}$ and $^{\mathrm{sim}}$ refer to the quantities that are estimated from the observed measurements and simulated structural responses, respectively. Figure 6 shows an example of simulated time series for an extensometer over the span of ten years. For these time series, the anomaly's height, duration, and the time corresponding to the center of the non-stationary regime are sampled from the respective uniform distributions $H_{\mathrm{a}} \sim \mathcal{U}(h_{\mathrm{a}}; -0.03, 0.03)$ mm, $W_{\mathrm{a}} \sim \mathcal{U}(w_{\mathrm{a}}; 0, 5)$ year, $T_{\mathrm{c}} \sim \mathcal{U}(t_{\mathrm{c}}; \text{2019-08}, \text{2029-08})$,

and the number of time steps without synthetic anomaly is $w_0 = 90$ days.

[Figure 6 about here.]

Once abnormal synthetic structural responses are obtained, we use the procedure summarized in Equation 2 to estimate its hidden-state variables and regime probabilities. Therefore, at a time t, the *environment's state vector* for which with an agent interacts, is defined as $\mathbf{s}_t = [\mu_t^{\mathtt{T}}\ \pi_t^{\mathtt{ns}}]^{\intercal} \in \mathbb{R}^2$. Figure 7a illustrates an example of the established environment's state space. Each blue dot indicates a pair of $(\mu_t^{\mathtt{T}}, \pi_t^{\mathtt{ns}})$ at time t obtained from the BDLM analysis of the synthetic time series.

[Figure 7 about here.]

4.3. Anomaly Detection in the Context of Reinforcement Learning

In the context of reinforcement learning, we follow the same procedure introduced by Khazaeli et al. (2021) to detect anomalies using the tabular Q-learning approach for a *discretized* environment state space. An example of such a discretized environment is shown in Figure 7b, where the shaded areas represent the visited states $\hat{\mathbf{s}}$, while the white ones correspond to non-visited states $\neg\hat{\mathbf{s}}$. In this setup, an agent takes an action of triggering an alarm a^0 or not triggering an alarm a^1 for a discretized environment's state $\hat{\mathbf{s}}$; The agent's optimal action is to trigger an alarm after an anomaly begins such that its corresponding time to detection is the closest to the anomaly's start time; The agent should not trigger an alarm for other time steps. Training an agent results in issuing a policy $\pi(\hat{\mathbf{s}})$ for each discretized state $\hat{\mathbf{s}}$ identifying the optimal actions for the visited states, i.e., *policy map*. Note that here, we consider the action of triggering an alarm for non-visited and out-of-domain, i.e., $\neg\hat{\mathbf{s}}$ states because these states correspond to either high probabilities of the non-stationary regime or high degradation speed values.

4.4. Anomaly Detection in the Context of Imitation Learning

The main difference between a reinforcement learning agent and an imitation learning one is that the former uses a reward function as a feedback signal to identify the optimal policy, while the latter uses demonstrations set established by an expert to imitate the expert's policy. In the simplest form, we can see the imitation learning approach as a classification problem in the context of the supervised learning approach where the labels (state-action pairs) are provided by an expert. A demonstration is represented by an episode $\hat{\mathcal{Z}}$ that is defined as a set of state-action pairs for consecutive time steps following

$$\hat{\mathcal{Z}} = \{(\mathbf{s}_1, a_1^{\text{expert}}), (\mathbf{s}_2, a_2^{\text{expert}}), \cdots\}, \tag{5}$$

where, at a time t, $\mathbf{s}_t$ is the environment's state vector, and $a_t^{\text{expert}} = \pi^{\text{expert}}(\mathbf{s}_t)$ is the action issued by an expert. An agent uses a set of expert demonstrations $\{\hat{\mathcal{Z}}_1, \hat{\mathcal{Z}}_2, \cdots, \hat{\mathcal{Z}}_{\mathtt{Z}}\}$ consisting in $\mathtt{Z}$ episodes to learn a policy that imitates the expert's policy. In this work, an agent directly learns to imitate the expert's policy using the classification approach. Therefore, the optimal policy $\pi^*(\mathbf{s})$ is obtained by building a classifier for a set of $\mathtt{D}$ joint observations $\mathcal{D} = \{(\mathbf{s}_i, a_i^{\text{expert}}), \forall i \in \{1 : \mathtt{D}\}\}$, where, $\mathbf{s}_i$ are the covariates vector and $a_i^{\text{expert}} \in \{-1 : a^0 \equiv \text{trigger an alarm}, +1 : a^1 \equiv \text{not trigger an alarm}\}$ are the system responses representing the actions taken by

the expert. Here, we employ the *tractable approximate Gaussian inference* (TAGI) method (Goulet, Nguyen, & Amiri, 2021) for Bayesian neural networks to build a classification model from demonstration sets. For the problem at hand, the probabilities of the two classes are: (i) the probability of triggering an alarm $\Pr(A = +1 : a^0|\mathbf{s})$, conditional on the covariates vector $\mathbf{s}$, and (ii) the probability of not triggering an alarm $\Pr(A = -1 : a^1|\mathbf{s}) = 1 - \Pr(A = +1 : a^0|\mathbf{s})$.

4.4.1. Demonstration Set Based on an Expert's Policy

To build a demonstration set, we choose $\mathtt{Z}$ episodes from the training dataset for each sensor. For each episode, we assume that (i) the expert policy $\pi^{\text{expert}}(\mathbf{s})$ for a visited state $\mathbf{s}$ is to trigger an alarm for the time steps during the anomaly window and not to trigger an alarm for remaining time steps, and (ii) the expert policy $\pi^{\text{expert}}(\neg\mathbf{s})$ for a non-visited or out-of-domain state $\neg\mathbf{s}$ is to trigger an alarm for all time steps. Therefore, the expert policy is

$$\pi^{\text{expert}}(\mathbf{s}) = \begin{cases} a^0, & \forall t \in (t_s, t_e) \\ a^1, & \text{Otherwise} \end{cases}, \quad \pi^{\text{expert}}(\neg\mathbf{s}) = a^0, \forall t, \tag{6}$$

where, t_s and t_e are the anomaly's start and end time, respectively. Recall Figure 5 as an example of an abnormal episode used by an expert to build a demonstration. According to the expert's decision presented in Equation 6, the action for the normal time steps indicated by blue lines is to not trigger an alarm, while the action of triggering an alarm is only assigned for the time steps associated with the anomaly window such as the one identified by the shaded area in Figure 5a. This consideration is because the anomaly function for $t \notin (t_s, t_e)$ is constant and it results in a zero derivative for these values. This implies that the structural degradation's speed is zero for these time steps and therefore, the optimal policy is to not trigger an alarm. We employ the aforementioned procedure on simulated time series with a wide range of anomalies to establish a training dataset and build a classifier. Before training an agent to imitate the expert's policy, we reduce the number of samples to accelerate the training procedure; This was already noted in the work of Syed and Schapire (Syed & Schapire, 2010), where they proved that imitation learning requires fewer demonstrations than reinforcement learning to imitate an expert. To this end, first, we use a two-dimensional grid $d_{\mathtt{M}\times\mathtt{N}}$, and apply it to the environment's state-space to create $\mathtt{M} \times \mathtt{N}$ unique regions. Afterwards, wherever it is possible, we randomly draw 10 samples from each of these regions. Moreover, we do not use the classifier built to extrapolate the actions predicted for the non-visited or out-of-domain states. Instead, we force the agent to imitate the expert's decision for these states which is to trigger an alarm. To this end, we add a third class for such states when building the classifier, and later directly assign the action of triggering an alarm for these states. To generate data for this new class, we draw one sample from each of the non-visited regions.

Figure 8 shows an example of a demonstration set, where we choose $\mathtt{Z} = 100000$ episodes from the training dataset. This results in a dataset with a total of approximately 33000 samples, which is further split into a training and test sets such that 70% of the samples are used to train a classifier.

[Figure 8 about here.]

The red dots are associated with the abnormal states which correspond to the time steps during an anomaly. For these states, the expert's decision is to trigger an alarm, i.e., a^0.

On the other hand, the blue dots are associated with the normal states corresponding to the remaining time steps. For these states, the expert's decision is to not trigger an alarm, i.e., a^1. The yellow dots are associated with the non-visited states for which the agent is forced to imitate the expert's decision for these states, that is to trigger an alarm, i.e., a^0.

5 Anomaly Detectability Quantification

This section presents the methods developed for quantifying anomaly detectability concerning two metrics, namely the annual false positive ratio and the probability of true positive detections. Furthermore, we propose two metrics identifying the capacity of a sensor for detecting anomalies with respect to the anomaly's magnitude as well as the time required for an anomaly to develop.

5.1. Categorization of an Agent's Decisions

To quantify an agent's performance at detecting anomalies, we need to categorize the outcome of its decisions: If an action is to trigger an alarm at any time t after the anomaly's start time t_s, the decision leads to a true positive detection, while if an agent triggers an alarm at any time step before t_s, it is a false positive one. In addition, if an action is to not trigger an alarm for any time step after t_s, it leads to false negative detection, while for any time steps before t_s it is a true negative one. To draw probabilistic conclusions about the detectability of anomalies, we generate independent synthetic datasets with labelled anomalies and test an agent against them. As a result, we can quantify the anomaly detectability capacity of an agent concerning two quantities: (i) the annual false positive ratio, and (ii) the probability of true positive detections.

5.2. Annual False Positive Detection Rate

The estimation of the annual false positive ratio ζ_FP involves testing the agent against an environment $\mathcal{S}^{\mathrm{test,\,FP}}$ for which all the episodes are free from the anomaly, and then calculating the fraction of false positive detections from the *total number of false positive detections*. To this end, two quantities are evaluated, namely the total number of time steps *#time steps* obtained from the total number of episodes in the test environment, and the average of the number of time steps per year *#time steps/year* for each episode. Therefore, the annual false positive ratio ζ_FP is expressed in events per year and is estimated by

$$\zeta_\mathrm{FP} = \frac{\#\text{false positive detections} \times \#\text{time steps/year}}{\#\text{time steps}}. \tag{7}$$

5.3. Probability of True Positive Detection

In order to evaluate the probability of a true positive detection, we only consider the actions for the time steps after the beginning of an anomaly, i.e, $t \geq t_\mathrm{s}$. This implies that only true positives and false negative detections are considered in order

to build a probabilistic model for true positive detections. The task of building such a probabilistic model corresponds to a classification problem with two classes $\{-1 \equiv$ false negative, $+1 \equiv$ true positive$\}$. The covariates for this classification problem are the absolute value of the anomaly's magnitude $|h_{\mathrm{a}}|$, its duration w_{a}, and the number of time steps $\mathtt{N}_{t>t_{\mathrm{s}}}$ that are required to detect an anomaly after its start. Note that the covariate $\mathtt{N}_{t>t_{\mathrm{s}}}$ determines the delay caused by an agent in detecting an anomaly, while the other two covariates are the anomaly's characteristics. For a test environment $\mathcal{S}^{\mathrm{test,\,TP}}$, a set of $\mathtt{D}$ joint observations $\mathcal{D} = \{(\mathbf{x}_i, y_i), \forall i \in \{1 : \mathtt{D}\}\}$, where, $\mathbf{x}_i = [|h_{\mathrm{a}}|\ w_{\mathrm{a}}\ \mathtt{N}_{t>t_{\mathrm{s}}}]^\intercal$ are the covariates vector and $y_i \in \{-1, +1\}$ are the system responses. Note that whenever the policy is to trigger an alarm after the beginning of an anomaly, we consider the policy for the remaining time steps of the episode to be triggering an alarm. This consideration is due to the termination criterion for which the agent-environment interaction is terminated whenever the agent triggers an alarm.

Figure 9 shows an example of an abnormal episode used in building a test dataset $\mathcal{D}$.

[Figure 9 about here.]

Figure 9a illustrates a realization of the anomaly function $f_{\mathrm{a}}(t; \mathcal{J})$. The anomaly's magnitude is $h_{\mathrm{a}} = 0.11$ and its duration is $w_{\mathrm{a}} = 1$ year. The unit of the anomaly magnitude is the same as the measurements. Figure 9c depicts the resulting simulated time series using an anomaly realization along with the simulation procedure expressed in Equation 4. Figures 9b and 9d show the trend expected value $\mu^{\mathtt{T}}$ and the probability of the non-stationary regime, respectively. Here, the agent triggers an alarm approximately 10 months = 0.83 year after the anomaly begins. The point where the alarm is triggered is identified by a circle marker and a vertical dashed line. The covariates vector instance corresponding to this episode is $\mathbf{x}_i = [0.011\ 1\ 0.83]^\intercal$.

Figure 10 shows an example of the results on a test environment containing 1200 episodes, where each consists of an anomaly having a different magnitude, duration, and time of occurrence.

[Figure 10 about here.]

The three axes correspond to the covariates vector $\mathbf{x}$. The red dots are associated with the true positive detections, where the action consists in rightfully triggering an alarm, i.e., $y = +1$. The blue dots indicate the false negative detections, $y = -1$, for which the agent wrongfully avoids triggering an alarm. Note that the episodes resulting in false positives are discarded because they are considered separately for evaluating the annual false positive ratio as explained in Section 5.2. We employ the *tractable approximate Gaussian inference* (TAGI) method (Goulet et al., 2021) for Bayesian neural networks in order to build a classification model that will characterize the probability of a true positive detection as a function of the anomaly's magnitude, its duration, and the time to detection. TAGI is parameterized by the neural network's weights and biases. Once these parameters are estimated from the data presented in Figure 10, the probability of each class for a given covariates vector $\mathbf{x}$ can be evaluated. For the problem at hand, the probabilities of the two classes are: (i) the probability of a true positive detection $\Pr(Y = +1 : \mathrm{TP}|\mathbf{x})$, conditional on the covariates vector $\mathbf{x}$, and (ii) the probability of a false negative detection $\Pr(Y = -1 : \mathrm{FN}|\mathbf{x}) = 1 - \Pr(Y = +1 : \mathrm{TP}|\mathbf{x})$.

Figure 11 presents an example of the results obtained from the TAGI classification.

[Figure 11 about here.]

Figure 11a presents an example of a model trained using TAGI with the observations obtained for 12000 episodes corresponding to the case presented in Figure 10. The

colours correspond to the probability of true positive detections for any instance of the covariates vector $\mathbf{x}$. Figure 11b shows the isosurface extracted from Figure 11a for a 90% probability of true positive detections. Such an isosurface enables quantifying the anomaly detectability of an agent with respect to two quantities identified by the two vertical planes: (i) the *minimum detectable anomaly magnitude* $|h_a|_{\min}$, which is shown by the blue colour plane, and (ii) the *minimum detectable anomaly growth rate per year* $\psi_{\min} = \psi_0 \times |h_a|_{\min}$, which is shown by the yellow colour plane. $\psi_{\min}$ is a function of the minimum detectable anomaly magnitude, where ψ_0 is a constant obtained from the slope of this plane with respect to the anomaly duration w_a. In Figure 11b, the minimum detectable anomaly magnitude $|h_a|_{\min} = 0.01$. In addition the minimum detectable anomaly growth rate $\psi_{\min} = 0.15|h_a|_{\min}$ unit per year. For this case, the anomaly with the magnitude of 0.01 is detectable with the probability of 90%, if it fully develops in less than $\mathtt{N}_\psi = 1/0.15 \approx 7$ years. For instance, the black circle in Figure 11b with the minimum anomaly's magnitude and the duration of 7 years. Such an anomaly is detectable within 7 years after it starts, with a probability of 90%. Another interpretation of using the yellow plane is related to the point shown by a black diamond in Figure 11b. This point shows an anomaly with a minimum magnitude of 0.01 and a duration of 5 years. As it is seen, this anomaly is detectable with a probability of 90% within 6 years after its starting time. However, the anomaly's magnitude for this point could also be associated with the anomaly shown by the black circle which is detected by the agent after 6 years. This discussion is also valid for other anomalies with the same magnitude but different durations.

5.4. Damage Detection

We have introduced methodologies in order to detect anomalies from structural responses and to quantify the anomaly detectability with respect to the probability of true positive detections and the annual false positive ratio. We use the finite element (FE) method to build a model of a structure in order to study the changes in structural responses caused by different structural damage scenarios. The FE analysis quantifies the elements' deformations subject to various loading types. The deformations obtained under the structure's self-weight without the application of damages are referred to as the *intact* state of the structure. Applying structural damages on a FE model results in new deformation values that are referred to as the *damaged* state of a structure. The corresponding nodal displacements and rotations associated with the deformations can be simulated at the locations where sensors are installed. The difference in nodal displacements and rotations between the intact and damaged states determines the change in structural responses due to each damage scenario. This change is associated with an anomaly's magnitude used in the process of defining the anomaly function's characteristics as explained in Section 4.1. Therefore, a structural damage is deemed to be detectable if the resulting changes in the structural responses at the location of a sensor are greater than the minimum detectable anomaly's magnitude.

To illustrate the relationship between changes in sensor measurements and a damage event, consider Figure 12a which shows the deformation of an element due to a damage scenario. $\Delta\ell$ represents the change in structural responses before and after the deformation of an element for an extensometer, and $\Delta\theta$ indicates this change for an inclinometer. The difference between inclinations of the intact θ_{intact} and damaged θ_{damaged} state of the structure is $\Delta\theta = \theta_{\text{damaged}} - \theta_{\text{intact}}$. The extensometer readings are based on the relative displacements between the supports of the extensometer. The

length of the sensor is evaluated for the intact ℓ_{intact} and damaged ℓ_{damaged} states of the structure, and the difference is calculated as $\Delta\ell = \ell_{\text{damaged}} - \ell_{\text{intact}}$. Figure 12b shows the displacement for the two nodes corresponding to an extensometer supports. Each node has a coordinate and the difference between the nodal coordinates are $\Delta x_{2,1} = x_2 - x_1$ and $\Delta y_{2,1} = y_2 - y_1$.

[Figure 12 about here.]

By defining the local horizontal u_i and vertical v_i displacements for each node $i \in \{1, 2\}$, the new length of the extensometer due to a deformation is obtained as $\ell = \sqrt{[\Delta x_{2,1} + (u_2 - u_1)]^2 + [\Delta y_{2,1} + (v_2 - v_1)]^2}$. The horizontal and vertical displacements are obtained from the FE analysis subject to different damage scenarios. An example of a damage scenario and corresponding changes in structural responses is shown in Figure 16. The induced changes $\Delta\ell$ and $\Delta\theta$ are compared with the minimum detectable anomaly magnitude $|h_a|_{\min}$ such as the one shown in Figure 11b, to identify whether the corresponding damage scenarios is detectable or not.

6 Case Study

This section provides the results obtained using the damage detection framework discussed in Section 2. In this regard, we present the empirical models built for each sensor on the three instrumented spans using the Bayesian dynamic linear models. Afterwards, we present the results associated with the structural damages detected using the methods proposed.

6.1. Empirical Model Estimation

Structural degradation can manifest in various forms, such as alterations in material properties or reduced stiffness. This degradation typically leads to observable changes in sensor data. For example, steady corrosion in a bridge can cause increased deflections, which in turn are reflected in sensor readings. We model these effects using the baseline component $\mathbf{x}_t^{\mathtt{B}}$ explained in Section 3.1 in order to capture the evolving nature of structural degradation.

In this regard, for non-stationary regimes, we employ a level hidden state variable $x^{\mathtt{L}}$, it's rate of change called trend hidden state variable $x^{\mathtt{T}}$, and the rate of change in the trend that is the local *acceleration* hidden state variable $x^{\mathtt{LA}}$. These three variables represent the degradation's kinematics describing the irreversible responses of the structure's degradation, the degradation's speed, and the degradation's acceleration; They jointly form the local acceleration vector $\mathbf{x}_t^{\mathtt{LA}} = \left[x_t^{\mathtt{L}}\ x_t^{\mathtt{T}}\ x_t^{\mathtt{LA}}\right]^{\intercal}$ similar to the one presented in Section 3.1. For stationary regimes, the acceleration is zero for all the time steps. Therefore, the irreversible responses consist of a level and a local trend hidden-state variables. They jointly form the local trend vector $\mathbf{x}_t^{\mathtt{LT}} = \left[x_t^{\mathtt{L}}\ x_t^{\mathtt{LT}}\right]^{\intercal}$. Note that the only irreversible response to model temperature data is the local level $x^{\mathtt{LL}}$ hidden state variable. In addition, we use the kernel regression component (Nguyen et al., 2019) in order to model the reversible *periodic* patterns; For the measurements associated with inclinometers and thermometers, we use 10 hidden control points for the kernel regression which result in the vector of hidden states $\mathbf{x}_t^{\mathtt{KR}} = \left[x_{t,0}^{\mathtt{KR}}\ x_{t,1}^{\mathtt{KR}}\ \cdots\ x_{t,10}^{\mathtt{KR}}\right]^{\intercal}$. Furthermore, we use the time-dependent *autoregressive* component $x_t^{\mathtt{AR}}$ in order to model the effect of other phenomena that are not accounted for by other components.

The parameter set $\mathcal{P}$ consists in $\sigma_w \in \mathbb{R}^+$, the standard deviation of the process

noise for each irreversible component; $\ell \in \mathbb{R}^+$ and $\sigma \in \mathbb{R}^+$ are the kernel length and the standard deviation of the hidden periodic pattern corresponding to the kernel regression component; and $\phi^{\mathtt{AR}} \in (0, 1]$, the autoregressive coefficient. Here, we assume a fixed kernel period equal to 365.2422 days and a zero standard deviation for the hidden control points. Apart from the parameters corresponding to the components, there are additional ones that need to be learned; They are $\sigma_v \in \mathbb{R}^+$, the observation standard deviation for each measurement, and regression coefficients β expressing the dependencies between time series data. We use an identical empirical model structure for each sensor type. The corresponding hidden-state variables vector $\mathbf{x}_t$ follows

$$\begin{aligned}
\mathbf{x}_t^{\mathtt{T}} &= \left[x_t^{\mathtt{LL,T}}; \mathbf{x}_t^{\mathtt{KR,T}}; x_t^{\mathtt{AR,T}}\right], \\
\mathbf{x}_t^{\mathtt{E}} &= \left[\mathbf{x}_t^{\mathtt{LA,E}}; x_t^{\mathtt{AR,E}}; \mathbf{x}_t^{\mathtt{T}}\right], \\
\mathbf{x}_t^{\mathtt{I}} &= \left[\mathbf{x}_t^{\mathtt{LA,I}}; \mathbf{x}_t^{\mathtt{KR,I}}; x_t^{\mathtt{AR,I}}; \mathbf{x}_t^{\mathtt{T}}\right],
\end{aligned}$$

where $[\,;]$ indicates the column-wise concatenation of the vectors and the additional superscripts $^{\mathtt{T}}$, $^{\mathtt{I}}$, and $^{\mathtt{E}}$ indicate that the observations are respectively related to the temperature, inclination, and elongation data. Note that the generic BDLM model for the stationary regime $\mathcal{M}^{\mathtt{s}}$ remains the same as the non-stationary one except that it does not include the acceleration. Therefore, the irreversible responses consist in a level $x^{\mathtt{L}}$ and a local trend $x^{\mathtt{LT}}$ hidden state variables. Moreover, building an empirical model for temperature data only requires the stationary model $\mathcal{M}^{\mathtt{s}}$ because we do not expect to have a regime switch. Appendix A provides the detailed model matrices for both regimes, i.e., $\mathcal{M}^{\mathtt{s}}$ and $\mathcal{M}^{\mathtt{ns}}$. The model parameters that learned from data are summarized as

$$\mathcal{P}^{\mathtt{T}} = \left\{\ell^{\mathtt{KR,T}}, \phi^{\mathtt{AR,T}}, \sigma^{\mathtt{AR,T}}, \sigma_v^{\mathtt{T}}\right\}, \tag{8}$$

$$\mathcal{P}^{\mathtt{E}} = \left\{ \underbrace{z^{\mathrm{s,s}}, z^{\mathrm{ns,ns}}}_{\text{regime transition}}, \underbrace{\sigma_{\mathrm{Tr}}^{\mathtt{LA,E}}, \phi^{\mathtt{AR,E}}, \sigma^{\mathtt{AR,E}}, \sigma_v^{\mathtt{E}}}_{\text{elongation, E}}, \underbrace{\beta^{\mathtt{AR,I|T}}, \beta^{\mathtt{KR,E|T}}}_{\text{dependency, E|T}}, \mathcal{P}^{\mathtt{T}} \right\}, \tag{9}$$

$$\mathcal{P}^{\mathtt{I}} = \left\{ \underbrace{z^{\mathrm{s,s}}, z^{\mathrm{ns,ns}}}_{\text{regime transition}}, \underbrace{\sigma_{\mathrm{Tr}}^{\mathtt{LA,I}}, \ell^{\mathtt{KR,I}}, \phi^{\mathtt{AR,I}}, \sigma^{\mathtt{AR,I}}, \sigma_v^{\mathtt{I}}}_{\text{inclination, I}}, \underbrace{\beta^{\mathtt{AR,I|T}}, \beta^{\mathtt{KR,I|T}}}_{\text{dependency, I|T}}, \mathcal{P}^{\mathtt{T}} \right\}, \tag{10}$$

where $z^{i,i}$ is the transition probability between the regimes, and $\sigma_{\mathrm{Tr}}^{\mathtt{LA}} \in \mathbb{R}^+$ is the standard deviation of the transition model that corresponds to the transition from a stationary to a non-stationary regime. We use the same model parameters associated with temperature data presented in Equation 8 for other measurements. In Equations 9 and 10, two additional parameters are introduced; They correspond to the dependency between the temperature and inclination and elongation data with respect to the autoregressive component, i.e., $\beta^{\mathtt{AR}}$, and the kernel regression component $\beta^{\mathtt{KR}}$. These two coefficients determine how temperature data affect the structural measurements. The complete hidden state estimations for all the sensors are presented in Appendix B.

6.2. Anomaly Detection

This section presents the results regarding the detection of anomalies using the reinforcement and imitation learning approaches for two target annual false positive ratios of $\zeta_{\text{FP}} = 0.05$ and 0.1 per year. The configurations used for training and testing the reinforcement and imitation learning agents are provided in Appendix C.

Figures 13 and 14 illustrate the resulting policies for the two annual false positive ratios.

[Figure 13 about here.]

[Figure 14 about here.]

Due to the symmetrical property of the trend expected values, only the positive half of the environment is shown. The red colour indicates that the action is a^0: Trigger an alarm, and the blue one indicates that the action is a^1: Not trigger an alarm. The white areas are associated with the non-visited states. In addition, we consider that the agents' decision is to trigger an alarm for the out-of-domain states with respect to the degradation speed, that is $|\mu_t^{\mathtt{T}}| > 5 \times 10^{-5}$ for the inclinometers and $|\mu_t^{\mathtt{T}}| > 1 \times 10^{-4}$ for the extensometers. This consideration is due to the large degradation speed of the out-of-domain states for which the probability of the non-stationary regime is estimated by the BDLM to be close to one. Figures 13a-13d correspond to the case for which the annual false positive ratio is $\zeta_{\text{FP}} = 0.05$/year, while Figures 14a-14d are associated with the $\zeta_{\text{FP}} = 0.10$/year. For a $(|\mu_t^{\mathtt{T}}|, \pi_t^{\mathtt{ns}})$ pair in a continuous space, each figure shows the policy of the agent in the discretized environment. The agents' action is to do nothing for the region associated with low degradation speed and probability of the non-stationary regime, i.e., the blue region at the bottom left corners in Figures 13a-14d. Note that this region is larger for the cases with the false positive ratio of $\zeta_{\text{FP}} = 0.05$/year compared with the ones with $\zeta_{\text{FP}} = 0.1$/year. This difference is due to changes in the reward function values such that the higher is the penalty for wrongfully detecting an anomaly, the larger the blue region; Other differences in the agents' behavior are sensor dependent; For instance, in the cases of the extensometer shown in Figures 13d and 14d, the blue region is narrower for the case associated with a higher false positive ratio shown in Figure 14d, which indicates that the agent has the tendency to trigger more alarms for the same region compared with the agent shown in Figure 13d. Choosing a lower penalty for wrongfully detecting anomalies increases the false positive ratio, and subsequently brings about more false alarms. The current work is limited to the three instrumented spans. Changing the false positive ratio from 0.05 to 0.1 for each span increases the number of false alarms for all three spans from approximately one every $20/3 \approx 7$ years to one every $10/3 \approx 3$ years.

For the imitation learning approach, we use TAGI (Goulet et al., 2021) in order to build a classifier for each sensor's demonstration set as presented in Section 4.4. Note that the number of episodes we used to train an agent using the imitation learning approach is approximately 8% of the number of episodes that were used to train an agent using the reinforcement learning approach. In order to obtain the target annual false positive rations, the minimum probability of triggering an alarm $\Pr_{\text{min}}(A = -1 : a^0|\mathbf{s})$ needs to be determined; We applied a grid search on an independent test dataset over the alarm-triggering probabilities in order to obtain the target annual false positive ratios. Figure 15 illustrates the contours for the minimum probability of triggering an alarm $\Pr_{\text{min}}(Y = -1 : a^0|\mathbf{s})$ for the inclinometers and extensometers in order to obtain one false positive detection every 5, 10, 20, and 40 years. These values are equivalent

to annual false positive ratios ζ_{FP} of 0.2, 0.1, 0.05, and 0.025 false alarms per year. The black contour line indicates a 50% probability of triggering an alarm, and the corresponding annual false positive ratio is indicated by the superscript *. In some cases, the spacing between contour lines is so small that they are not distinguishable, e.g., Figure 15a and Figure 15b that correspond to the lateral inclinometer and longitudinal inclinometer 1, respectively. This shows the sensitivity of the imitation learning approach for smaller annual false positive ratios.

[Figure 15 about here.]

7 Anomaly Detectability Quantification

We follow the procedure described in Section 5.3 in order to evaluate the probability of true positive detections for each agent. Table 1 summarizes the minimum detectable anomaly's magnitude and its growth rate for all the sensors and for the target annual false positive ratios. The first column shows the name of the sensors. The second column is associated with the agents with the annual false positive ratios of $\zeta_{\text{FP}} = 0.05$ and 0.1 per year. The following three columns correspond to the metrics obtained from the imitation learning approach; They show the minimum detectable anomaly magnitude $|h_a|_{\min}$, the the constant ψ_0, and the minimum detectable anomaly growth rates $\psi_{\min}$ in years as a function of $|h_a|_{\min}$, i.e., $\psi_{\min} = |h_a|_{\min}$ per $\mathtt{N}_\psi$ years. For instance, for the agent corresponding to the lateral inclinometer with the annual false alarm ratio of $\zeta_{\text{FP}} = 0.05$/year, an anomaly with a minimum magnitude of 0.003°d is detectable, if it develops in less than $\mathtt{N}_\psi = 13$ years.

[Table 1 about here.]

In Table 1, the last two columns are associated with the minimum detectable anomaly magnitude and minimum detectable anomaly growth rates in years obtained from the reinforcement learning (RL) approach. Comparing these two sets of values with the ones obtained from the imitation learning (IL) approach reveals that both approaches result in a similar anomaly detectability for all the sensors except for the lateral inclinometers. In this case, the reinforcement learning approach shows a better anomaly detectability with respect to the minimum detectable anomaly's magnitude $|h_a|_{\min}$. In addition, comparing the maximum number of years for developing the minimum anomaly $\mathtt{N}_\psi$ between the RL and IL approaches reveals that IL-based anomaly detection allows for detecting anomalies that take a longer time to develop. In particular, this is the case for the anomaly detectability corresponding to the lateral inclinometer; The values for $\mathtt{N}_\psi$ are higher than the ones for the reinforcement learning approach.

7.1. Damage Detection

This section presents the structural damage scenarios issued for the instrumented spans; We consider three categories of structural damage scenarios that are affecting each span globally. They are *corrosion* (`COR`), *partial settlement* (`PST`), and *bearing device lockage* (`BDL`). The corrosion of a span is modelled as a reduction in the cross-section area of steel elements. Partial settlements are modelled as relative vertical displacements of the supports. Partial settlements are considered in two setups namely, *uniform* (`U`) and *non-uniform* (`NU`). The uniform settlement corresponds to the settlement of one support relative to the others. The side for which the vertical settlements are applied is

the one where the bearing devices are located as shown in Figure 2. The non-uniform settlements are associated with the relative settlement of only one support that is located on the upstream side of the bridge at the location of the bearing device. For simulating a bearing device lockage, we use a spring at the location of the supports in order to connect the fixed horizontal boundary condition and the bridge members. The stiffness of the spring is bounded between $k_{\min} = 0$ and $k_{\max}$. They represent the situation where they can behave like a roller or fixed support, respectively. When the loading is applied to the span, the corresponding horizontal displacements of the bearing devices change depending on the value of the spring's stiffness. In order to model various scenarios, different values of stiffness are selected and the reduction in the average of the bearing device horizontal displacement is evaluated in percentage. Bearing device lockage is modelled in two setups namely, *uniform* and *non-uniform*. The uniform configuration corresponds to applying a scenario on both bearing devices. The non-uniform configuration is associated with applying a damage scenario only on the bearing device located on the upstream side of the bridge.

Figure 16a shows the intact state of the span for which the deformations are due to the elements' self-weight. The numbers shown beside each sensor name are the value of either $\Delta\theta$ or $\Delta\ell$. Note that, $\Delta\ell$ is the change in structural responses before and after the deformation of an element for an extensometer in mm, and $\Delta\theta$ is this change for an inclinometer in °d. Figure 16b shows the damaged state of the span for which a 36.3% lockage is applied on the bearing device located on the upstream side of the span. Accordingly, the corresponding relative rotations and elongation, i.e., $\Delta\theta$ and $\Delta\ell$, at the location of the sensors are changed.

[Figure 16 about here.]

In order to identify damages that are detectable, we compare the minimum detectable anomaly magnitude $|h_a|_{\min}$ obtained from the Table 1 with the relative rotations and elongation obtained from the FE analysis. Table 2 shows the results of their comparison for all sensors, where the red colour indicates the damage scenarios that are detectable by the agents with an annual false alarm ratio ζ_{FP} of 0.1/year.

[Table 2 about here.]

The second column introduces five *damage types* used in order to analyze the changes in structural responses. The third column presents the abbreviation for damage scenarios names. In case of corrosion, we introduced two damage scenarios as 2% and 5% reduction in the cross-section areas. In the table, respective abbreviations are $\texttt{COR} - 2\%$ and $\texttt{COR} - 5\%$. In case of bearing device lockage, the percentages for the lockage varies from 5.4% to 74.1% depending whether the damage scenario is uniform or not. For instance, $\texttt{BDL} - \texttt{NU} - 11\%$ is the damage scenario associated with 11% bearing device lockage in the non-uniform setup. For partial settlement, the values varies from 1 mm to 5 mm depending on whether the settlement is uniform or not. For example, $\texttt{PST} - \texttt{NU} - 1\text{mm}$ indicates 1 mm partial settlement in non-uniform fashion. The fourth to sixth columns show the changes in structural responses at the location of the sensors due to each damage scenario. In Table 2, the severity of each damage type has a lower and an upper bound. For example, these severities are 2% and 5% for corrosion. The lower bound severity is associated with the damage scenario that is detectable by at least one sensor; Severities smaller than the lower bound are not detectable by any sensor. On the other hand, the upper bound severity corresponds to the damage scenario for which the number and types of sensors that are able to detect severities larger than the upper bound remain the same. For example, the minimum severity corresponding

to the corrosion is detectable by the extensometer and the longitudinal inclinometer 1; None of the sensors are able to detect the damages associated with corrosion of less than 2%. A corrosion of 5% is the upper bound severity. It is detectable by the longitudinal inclinometers and the extensometer meaning that corrosion larger than 5% is detectable with the same sensors. Note that the lateral inclinometer cannot detect this type of damage.

From Table 2, one can compare the damage detectability of each sensor individually. For instance, the best sensor is the extensometer as it detects more damage scenarios than any other. If we want to be able to detect all damage scenarios, we would need to use the extensometer, the lateral inclinometer, and either the longitudinal inclinometer 1 or 3. In order to illustrate this point, consider the two scenarios $\texttt{COR} - 2\%$ and $\texttt{PST} - \texttt{NU} - 5\,\text{mm}$ for the extensometer shown in the first and last row of Table 2. If the corresponding agent triggers an alarm for which the anomaly magnitude is $|h_a| = 0.02$, it is not possible to identify the cause of this anomaly only from the extensometer. Nevertheless, we can rely on multiple sensors in order to distinguish between these two scenarios. For example, by using the lateral inclinometer, one can realize that the $\texttt{PST} - \texttt{NU} - 5\,\text{mm}$ scenario is the cause of the anomaly because the lateral sensor cannot detect the $\texttt{COR} - 2\%$ scenario.

8 Conclusion

This research work presented a damage detection procedure based on Bayesian dynamic linear models (BDLMs) along with the reinforcement learning (RL) and imitation learning (IL) approaches. The methodologies proposed showed the capacity of reinforcement and imitation learning approaches to successfully quantify: (i) the anomaly detectability of different sensor types with respect to the target annual false positive ratio as well as the probability of true positive detections, which resulted in two metrics that are the minimum anomaly's magnitude $|h_a|_{\min}$ and the minimum anomaly's growth rate ψ, and (ii) the structural damage detection in order to associate an anomaly's magnitude to the severity levels of different structural damages. These methods showed a good performance on different measurements collected on a case-study bridge in Canada; They determined the minimum anomaly detectability and quantified the damage detectability of each sensor. The current sensor configuration is expected to detect the structural damages with a probability of 90% caused by corrosion of $\geq 2\%$, a uniform bearing device lockage of $\geq 5.4\%$, a non-uniform bearing device lockage of $\geq 10.2\%$, and uniform and non-uniform partial settlements of $\geq 1\,\text{mm}$.

Acknowledgments

The authors acknowledge Les Ponts Jacques Cartier et Champlain Incorporée (PJCCI), Canada, for supporting this research project. Also, the authors acknowledge Dahlia Tchoukuegno at Polytechnique Montréal, for building finite element models for the instrumented spans of the case-study bridge.

References

Adam, B., & Smith, I. (2008). Reinforcement learning for structural control. *Journal of Computing in Civil Engineering*, *22*, 133–139.

Arora, S., & Doshi, P. (2021). A survey of inverse reinforcement learning: Challenges, methods and progress. *Artificial Intelligence*, *297*, 103500.

ASCE. (2021). *2021 report card for America's infrastructure* (Tech. Rep.). American Society of Civil Engineers.

Balsamo, L., Betti, R., & Beigi, H. (2014). A structural health monitoring strategy using cepstral features. *Journal of Sound and Vibration*, *333*(19), 4526 - 4542.

Black, A. P. (2022). *2022 bridge report* (Tech. Rep.). American Road and Transportation Builder Association.

BluePlan. (2019). *Monitoring the state of Canada's core public infrastructure: The Canadian infrastructure report card 2019* (Tech. Rep.). BluePlan Engineering.

Bull, L., Worden, K., Manson, G., & Dervilis, N. (2018). Active learning for semi-supervised structural health monitoring. *Journal of Sound and Vibration*, *437*, 373–388.

Farrar, C., & Worden, K. (2013). *Structural health monitoring: A machine learning perspective.* John Wiley and Sons.

Farrar, C. R., & Worden, K. (2010). An introduction to structural health monitoring. Springer.

Goulet, J.-A. (2017). Bayesian dynamic linear models for structural health monitoring. *Structural Control and Health Monitoring*, *24*(12), e2035.

Goulet, J.-A. (2020). *Probabilistic machine learning for civil engineers.* The MIT Press.

Goulet, J.-A., & Koo, K. (2018). Empirical validation of Bayesian dynamic linear models in the context of structural health monitoring. *Bridge Engineering*, *23*(2), 05017017.

Goulet, J.-A., Nguyen, L. H., & Amiri, S. (2021). Tractable approximate Gaussian inference for Bayesian neural networks. *Journal of Machine Learning Research*, *22*(251), 1-23.

Huang, C., Wu, Y., Zuo, Y., Pei, K., & Min, G. (2018). Towards experienced anomaly detector through reinforcement learning. In *Proceedings of the 32nd conference on artificial intelligence.*

Hussein, A., Gaber, M. M., Elyan, E., & Jayne, C. (2017). Imitation learning: A survey of learning methods. *Computing Surveys*, *50*.

Khalatbarisoltani, A., Soleymani, M., & Khodadadi, M. (2019). Online control of an active seismic system via reinforcement learning. *Structural Control and Health Monitoring*, *26*(3), e2298.

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

Korkmaz, S. (2011). A review of active structural control: Challenges for engineering informatics. *Computers and Structures*, *89*(23), 2113–2132.

Li, L., Morgantini, M., & Betti, R. (2023). Structural damage assessment through a new generalized autoencoder with features in the quefrency domain. *Mechanical Systems and Signal Processing*, *184*, 109713.

Malekzadeh, M., Atia, G., & Catbas, F. N. (2015). Performance-based structural health monitoring through an innovative hybrid data interpretation framework. *Journal of Civil Structural Health Monitoring*, *5*(3), 287–305.

Mohammadi Ghazi, R., & Büyüköztürk, O. (2016). Damage detection with small data set using energy-based nonlinear features. *Structural Control and Health Monitoring*, *23*(2), 333–348.

Murphy, K. P. (2013). *Machine learning : A probabilistic perspective.* The MIT Press.

Nassif, A. B., Talib, M. A., Nasir, Q., & Dakalbab, F. M. (2021). Machine learning for anomaly detection: A systematic review. *IEEE Access*, *9*, 78658-78700.

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

Nguyen, L. H., & Goulet, J.-A. (2018). Anomaly detection with the switching Kalman filter

for structural health monitoring. *Structural Control and Health Monitoring*, *25*(4), e2136.
Pomerleau, D. (1989). ALVINN: An autonomous land vehicle in a neural network. In *Proceedings of neural information processing systems.*
Ross, S., Gordon, G., & Bagnell, D. (2011). A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (Vol. 15).
Soong, T. T., & Cimellaro, G. P. (2009). Future directions in structural control. *Structural Control and Health Monitoring*, *16*(1), 7–16.
Syed, U., & Schapire, R. E. (2010). A reduction from apprenticeship learning to classification. In *Advances in neural information processing systems* (Vol. 23).
West, M., & Harrison, J. (1999). *Bayesian forecasting and dynamic models*. Springer.
Ying, Y., Garrett, J., Oppenheim, I., Soibelman, L., Harley, J., Shi, J., & Jin, Y. (2013). Toward data-driven structural health monitoring : Application of machine learning and signal processing to damage detection. *Journal of Computing in Civil Engineering*, *27*, 667–680.
Yu, M., & Sun, S. (2020). Policy-based reinforcement learning for time series anomaly detection. *Engineering Applications of Artificial Intelligence*, *95*, 103919.

Appendix A. BDLM Models

This section presents the model matrices used in this work in order to build empirical models for temperature, inclination, and elongation data.

A.1. Model Matrices for the Inclinometers

The model matrices $\{\mathbf{A}_t, \mathbf{C}, \mathbf{Q}_t, \mathbf{R}\}$ for the stationary and non-stationary model matrices for the inclinometers $\mathtt{I}$ are defined following

Stationary Model Matrices $\mathcal{M}^{\mathtt{s},obs}$

$$\mathbf{A}_t^{\mathtt{s}} = \text{block diag}\left(\begin{bmatrix} 1 & \Delta t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \tilde{\boldsymbol{k}}_t^{\mathtt{KR,I}} \\ \mathbf{0} & \mathbf{I}_{10\times 10} \end{bmatrix}, \phi^{\mathtt{AR,I}}, 1, \begin{bmatrix} 0 & \tilde{\boldsymbol{k}}_t^{\mathtt{KR,T}} \\ \mathbf{0} & \mathbf{I}_{10\times 10} \end{bmatrix}, \phi^{\mathtt{AR,T}}\right)$$

$$\mathbf{C}^{\mathtt{s}} = \begin{bmatrix} 1 & 0 & 0 & 1 & \mathbf{0}_{1\times 10} & 1 & 0 & \beta^{\mathtt{KR,I|T}} & \mathbf{0}_{1\times 10} & \beta^{\mathtt{AR,I|T}} \\ 0 & 0 & 0 & 0 & \mathbf{0}_{1\times 10} & 0 & 1 & 1 & \mathbf{0}_{1\times 10} & 1 \end{bmatrix}$$

$$\mathbf{R}^{\mathtt{s}} = \text{block diag}\left(\left(\sigma_v^{\mathtt{I}}\right)^2, \left(\sigma_v^{\mathtt{T}}\right)^2\right)$$

$$\mathbf{Q}_t^{\mathtt{s(s)}} = \text{block diag}\left(\left(\sigma_w^{\mathtt{LT}}\right)^2 \begin{bmatrix} \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} & 0 \\ \frac{\Delta t^2}{2} & \Delta t & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,I}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,I}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,I}}\right)^2, \left(\sigma_w^{\mathtt{LL}}\right)^2, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,T}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,T}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,T}}\right)^2\right)$$

$$\mathbf{Q}_t^{\mathtt{ns(s)}} = \text{block diag}\left(\begin{bmatrix} \left(\sigma_w^{\mathtt{LT}}\right)^2 \cdot \frac{\Delta t^3}{3} & 0 & 0 \\ 0 & \left(\sigma_w^{\mathtt{LTT}}\right)^2 \cdot \Delta t & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,I}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,I}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,I}}\right)^2, \left(\sigma_w^{\mathtt{LL}}\right)^2, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,T}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,T}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,T}}\right)^2\right)$$

Non-Stationary Model Matrices $\mathcal{M}^{\mathtt{ns},obs}$

$$\mathbf{A}_t^{\mathtt{ns}} = \text{block diag}\left(\begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & \tilde{\boldsymbol{k}}_t^{\mathtt{KR,I}} \\ \mathbf{0} & \mathbf{I}_{10\times 10} \end{bmatrix}, \phi^{\mathtt{AR,I}}, 1, \begin{bmatrix} 0 & \tilde{\boldsymbol{k}}_t^{\mathtt{KR,T}} \\ \mathbf{0} & \mathbf{I}_{10\times 10} \end{bmatrix}, \phi^{\mathtt{AR,T}}\right)$$

$$\mathbf{C}^{\mathtt{ns}} = \begin{bmatrix} 1 & 0 & 0 & 1 & \mathbf{0}_{1\times 10} & 1 & 0 & \beta^{\mathtt{KR,I|T}} & \mathbf{0}_{1\times 10} & \beta^{\mathtt{AR,I|T}} \\ 0 & 0 & 0 & 0 & \mathbf{0}_{1\times 10} & 0 & 1 & 1 & \mathbf{0}_{1\times 10} & 1 \end{bmatrix}$$

$$\mathbf{R}^{\mathtt{ns}} = \text{block diag}\left(\left(\sigma_v^{\mathtt{I}}\right)^2, \left(\sigma_v^{\mathtt{T}}\right)^2\right)$$

$$\mathbf{Q}_t^{\mathtt{s(ns)}} = \text{block diag}\left(\begin{bmatrix} \left(\sigma_w^{\mathtt{LA}}\right)^2 \cdot \frac{\Delta t^5}{20} & 0 & 0 \\ 0 & \left(\sigma_w^{\mathtt{LTT}}\right)^2 \cdot \frac{\Delta t^3}{3} & 0 \\ 0 & 0 & \left(\sigma_w^{\mathtt{LA}}\right)^2 \cdot \Delta t \end{bmatrix}, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,I}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,I}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,I}}\right)^2, \left(\sigma_w^{\mathtt{LL}}\right)^2, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,T}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,T}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,T}}\right)^2\right)$$

$$\mathbf{Q}_t^{\mathtt{ns(ns)}} = \text{block diag}\left(\left(\sigma_w^{\mathtt{LA}}\right)^2 \begin{bmatrix} \frac{\Delta t^5}{20} & \frac{\Delta t^4}{8} & \frac{\Delta t^3}{6} \\ \frac{\Delta t^4}{8} & \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^3}{6} & \frac{\Delta t^2}{2} & \Delta t \end{bmatrix}, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,I}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,I}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,I}}\right)^2, \left(\sigma_w^{\mathtt{LL}}\right)^2, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,T}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,T}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,T}}\right)^2\right)$$

where the superscript `I` indicates that the matrices belong to the inclinometers, the superscript `T` indicates that the matrices belong to the thermometers, and $\tilde{\boldsymbol{k}}_t^{\mathtt{KR}} = [\tilde{k}_{t,1}^{\mathtt{KR}}, \tilde{k}_{t,2}^{\mathtt{KR}}, \cdots, \tilde{k}_{t,10}^{\mathtt{KR}}]$ is the normalized kernel values and Δt is the time step at the time t.

A.2. Model Matrices for the Extensometers

The model matrices $\{\mathbf{A}_t, \mathbf{C}, \mathbf{Q}_t, \mathbf{R}\}$ for the stationary and non-stationary model matrices for the extensometers `E` are defined following

Stationary Model Matrices $\mathcal{M}^{\mathtt{s},obs}$

$$\mathbf{A}_t^{\mathtt{s}} = \text{block diag}\left(\left[\begin{array}{ccc} 1 & \Delta t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right], \phi^{\mathtt{AR,E}}, 1, \left[\begin{array}{cc} 0 & \tilde{\boldsymbol{k}}_t^{\mathtt{KR,T}} \\ \mathbf{0} & \mathbf{I}_{10\times 10} \end{array}\right], \phi^{\mathtt{AR,T}}\right)$$

$$\mathbf{C}^{\mathtt{s}} = \left[\begin{array}{cccccccccc} 1 & 0 & 0 & 1 & 1 & 0 & \beta^{\mathtt{KR,E|T}} & \mathbf{0}_{1\times 10} & \beta^{\mathtt{AR,E|T}} \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & \mathbf{0}_{1\times 10} & 1 \end{array}\right]$$

$$\mathbf{R}^{\mathtt{s}} = \text{block diag}\left(\left(\sigma_v^{\mathtt{E}}\right)^2, \left(\sigma_v^{\mathtt{T}}\right)^2\right)$$

$$\mathbf{Q}_t^{\mathtt{s(s)}} = \text{block diag}\left(\left(\sigma_w^{\mathtt{LT}}\right)^2 \left[\begin{array}{ccc} \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} & 0 \\ \frac{\Delta t^2}{2} & \Delta t & 0 \\ 0 & 0 & 0 \end{array}\right], \left(\sigma^{\mathtt{AR,E}}\right)^2, \left(\sigma_w^{\mathtt{LL}}\right)^2, \left[\begin{array}{cc} \left(\sigma_{w,0}^{\mathtt{KR,T}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,T}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{array}\right], \left(\sigma^{\mathtt{AR,T}}\right)^2\right)$$

$$\mathbf{Q}_t^{\mathtt{ns(s)}} = \text{block diag}\left(\left[\begin{array}{ccc} \left(\sigma_w^{\mathtt{LT}}\right)^2 \cdot \frac{\Delta t^3}{3} & 0 & 0 \\ 0 & \left(\sigma_w^{\mathtt{LTT}}\right)^2 \cdot \Delta t & 0 \\ 0 & 0 & 0 \end{array}\right], \left(\sigma^{\mathtt{AR,E}}\right)^2, \left(\sigma_w^{\mathtt{LL}}\right)^2, \left[\begin{array}{cc} \left(\sigma_{w,0}^{\mathtt{KR,T}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,T}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{array}\right], \left(\sigma^{\mathtt{AR,T}}\right)^2\right)$$

Non-Stationary Model Matrices $\mathcal{M}^{\mathtt{ns},obs}$

$$\mathbf{A}_t^{\mathtt{ns}} = \text{block diag}\left(\begin{bmatrix} 1 & \Delta t & \frac{\Delta t^2}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}, \phi^{\mathtt{AR,E}}, 1, \begin{bmatrix} 0 & \tilde{\boldsymbol{k}}_t^{\mathtt{KR,T}} \\ \mathbf{0} & \mathbf{I}_{10\times 10} \end{bmatrix}, \phi^{\mathtt{AR,T}}\right)$$

$$\mathbf{C}^{\mathtt{ns}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & \beta^{\mathtt{KR,E|T}} & \mathbf{0}_{1\times 10} & \beta^{\mathtt{AR,E|T}} \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & \mathbf{0}_{1\times 10} & 1 \end{bmatrix}$$

$$\mathbf{R}^{\mathtt{ns}} = \text{block diag}\left(\left(\sigma_v^{\mathtt{E}}\right)^2, \left(\sigma_v^{\mathtt{T}}\right)^2\right)$$

$$\mathbf{Q}_t^{\mathtt{s(ns)}} = \text{block diag}\left(\begin{bmatrix} \left(\sigma_w^{\mathtt{LA}}\right)^2 \cdot \frac{\Delta t^5}{20} & 0 & 0 \\ 0 & \left(\sigma_w^{\mathtt{LTT}}\right)^2 \cdot \frac{\Delta t^3}{3} & 0 \\ 0 & 0 & \left(\sigma_w^{\mathtt{LA}}\right)^2 \cdot \Delta t \end{bmatrix}, \left(\sigma^{\mathtt{AR,E}}\right)^2, \left(\sigma_w^{\mathtt{LL}}\right)^2, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,T}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,T}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,T}}\right)^2\right)$$

$$\mathbf{Q}_t^{\mathtt{ns(ns)}} = \text{block diag}\left(\left(\sigma_w^{\mathtt{LA}}\right)^2 \begin{bmatrix} \frac{\Delta t^5}{20} & \frac{\Delta t^4}{8} & \frac{\Delta t^3}{6} \\ \frac{\Delta t^4}{8} & \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^3}{6} & \frac{\Delta t^2}{2} & \Delta t \end{bmatrix}, \left(\sigma^{\mathtt{AR,E}}\right)^2, \left(\sigma_w^{\mathtt{LL}}\right)^2, \begin{bmatrix} \left(\sigma_{w,0}^{\mathtt{KR,T}}\right)^2 & \mathbf{0} \\ \mathbf{0} & \left(\sigma_{w,1}^{\mathtt{KR,T}}\right)^2 \cdot \mathbf{I}_{10\times 10} \end{bmatrix}, \left(\sigma^{\mathtt{AR,T}}\right)^2\right)$$

where the superscript $\mathtt{E}$ indicates that the matrices belong to the extensometers, the superscript $\mathtt{T}$ indicates that the matrices belong to the thermometers, and $\tilde{\boldsymbol{k}}_t^{\mathtt{KR}} = [\tilde{k}_{t,1}^{\mathtt{KR}}, \tilde{k}_{t,2}^{\mathtt{KR}}, \cdots, \tilde{k}_{t,10}^{\mathtt{KR}}]$ is the normalized kernel values and Δt is the time step at the time t.

Appendix B. Hidden State Estimations

This section presents the hidden states estimations corresponding to the thermometers, extensometers, and longitudinal inclinometer measurements. Table 3 shows the optimized parameters set $\mathcal{P}$ for all three spans. The first column shows the name of the

spans, and the remaining columns are the values corresponding to the estimated parameters whose values are similar; This indicates that the temperature data decomposition results in similar hidden state variable estimations.

[Table 3 about here.]

For the empirical models corresponding to the extensometers, the regime transition probabilities for all three spans are fixed such that $z^{\mathrm{s,s}} = 0.9999$ and $z^{\mathrm{ns,ns}} = 0.9990$. Moreover, the observation standard deviation is fixed to $\sigma_v^{\mathtt{E}} = 0.001$ and the transition model standard deviation is set to $\sigma_{\mathrm{Tr}}^{\mathtt{LA}} = 1 \times 10^{-7}$. For the empirical models corresponding to the longitudinal inclinometers, The regime transition probabilities for all three spans are fixed such that $z^{\mathrm{s,s}} = 0.9999$ and $z^{\mathrm{ns,ns}} = 0.9990$. In addition, for these spans, the observation standard deviation is set to $\sigma_v^{\mathtt{I}} = 0.0001$ and the transition model standard deviation is fixed to $\sigma_{\mathrm{Tr}}^{\mathtt{LA}} = 1 \times 10^{-7}$. The fixed parameters corresponding to the lateral inclinometers are the same as of the longitudinal inclinometers except the transition probability $z^{\mathrm{ns,ns}} = 0.9995$. Table 4 shows the remaining optimized parameter values.

[Table 4 about here.]

The first two column identifies the sensor type and span names, and the following three columns show the optimized parameter values associated with the inclination data. The last two columns are the dependency values with the temperature data.

Appendix C. Reinforcement and Imitation Learning Setup

This section presents the configurations used in this work in order to detect damages using the reinforcement and imitation learning approaches.

Simulated Environments

We follow the procedure described in Section 4.2 in order to simulate environments. Rather than simulating an environment for each sensor, we simulate one environment for each set of sensors collecting the same measurement type across all three spans. This generalization of the environments across sensors is possible because the superstructures of the three spans and their subsequent structural responses are similar. As a result, there are four environments, and consequently, four agents, which correspond to the lateral inclinometers, the longitudinal inclinometers 1, the longitudinal inclinometers 3, and the extensometers. Table 5 shows the distributions and the parameters of the anomaly function for simulating new episodes for each sensor type. The duration of each episode is 10 years which starts in 2019 and ending in 2029. All probability distributions used for simulating the episodes have an uniform probability density with a lower a and an upper b bounds. The bounds for the anomaly magnitude depend on the type of the sensor, while they remain the same for other parameters associated with the anomaly function regardless of the sensor type; The duration for an anomaly is drawn from an uniform distribution with the upper bound $b = 2000$ days (≈ 5 years). The center of the anomaly is drawn from an uniform distribution for which the bounds are associated with the beginning and end times of an episode. Here, the anomaly tolerance ϵ_{a} depends on the realization of an anomaly; Its value is equivalent to 25% of the simulated anomaly's magnitude, i.e., $\epsilon_{\mathrm{a}} = 0.25 H_{\mathrm{a}}$.

[Table 5 about here.]

Training and Test Datasets

Table 6 summarizes the dataset properties used in order to train and test each agent corresponding to a sensor type. For the training dataset $\mathcal{D}^{\text{train}}$, we have simulated 400000 episodes associated with each sensor and then merged them into a single set of 1.2 million episodes for each sensor type. Therefore, the total number of continuous states exceeds 600 million during the training. The last column of Table 6 indicates the ratio of abnormal episodes. For instance, 0.9 indicates that 90% of the total number of episodes during the training contain an anomaly simulated with the configuration presented in Table 5, while 10% of the total number of episodes does not contain any anomaly.

[Table 6 about here.]

Moreover, simulating episodes for the test phase follows the same configurations as for the training phase shown in Table 5, except for the number and duration of episodes; In order to evaluate the annual false positive ratio, the test dataset $\mathcal{D}^{\text{test, FP}}$ consists of a total number of 3000 episodes with a duration of 10 years for each measurement type. Note that the episodes for the false positive test set do not contain anomalies, i.e., $h_{\text{a}} = 0$. Furthermore, the true positive test set $\mathcal{D}^{\text{test, TP}}$ consists of 6000 abnormal episodes for each measurement type. The duration of each episode is 60 years, for which the anomaly duration in days follows an uniform distribution $W_{\text{a}} \sim \mathcal{U}(w_{\text{a}}; 0, 8000)$.

Discretized Environment

Table 7 shows the configurations for the two-dimensional environments with a grid size of $\mathtt{M} \times \mathtt{N}$, which are used to train the agents. The states of each environment are bounded according to the sensor type. All states have the same lower and upper bounds with respect to the probability of a non-stationary regime, i.e., $0 \leq \pi_t^{\mathtt{ns}} \leq 1$. The states' bounds differ with respect to the trend expected values, that is $|\mu_t^{\mathtt{T}}| \leq 5 \times 10^{-5}$ for the inclinometers, and $|\mu_t^{\mathtt{T}}| \leq 1 \times 10^{-4}$ for the extensometer. In addition all the environments are discretized using a grid of 160×60.

[Table 7 about here.]

Agent-Environment Interaction

The hyperparameters for the exploration-exploitation method are $c_\alpha = c_\epsilon = 5$. The discount factor $\gamma = 0.97$ for all the agents, and the initial Q-value $Q_0 = 0$ for all the discretized states. Furthermore, for the training of each agent, we take the absolute value of the trend expected values, i.e., $|\mu_t^{\mathtt{T}}|$. This is because mapping the simulated episodes on the environment results in a symmetrical set of visited states with respect to the dimension associated with the degradation's speed $\mu_t^{\mathtt{T}}$ and therefore, training only the positive half of the environment is sufficient. Moreover, we remove the initial trend expected value $\mu_0^{\mathtt{T}}$ from each episode following

$$\mathbf{s}_t = \left[(\mu_t^{\mathtt{T}} - \mu_0^{\mathtt{T}})\ \pi_t^{\mathtt{ns}}\right]^{\mathtt{T}}. \tag{C1}$$

This modification ensures that each episode begins from a state with a zero trend expected value because: (i) the initial trend for inclinometers is not equal, and (ii) without such a modification, we would not be able to take the advantage of the above-mentioned symmetrical property for computational efficiency.

Reward Function Values

This section provides the reward function values in order to obtain the target annual false positive ratios ζ_{FP} of 0.05 and 0.1 per year for the reinforcement learning agents. Table 8 shows the reward values for each sensor type where it can be noted that the false positive reward value r^{FP} associated with $\zeta_{\text{FP}} = 0.05$ is less than the one for $\zeta_{\text{FP}} = 0.1$. For instance, for the lateral inclinometer, these values are -21 and -15 for the annual false positive ratio of 0.05 and 0.1, respectively.

[Table 8 about here.]

Appendix D. Probability of True Positive Detections

This section presents the probability of true positive detections for all sensors using the reinforcement and imitation learning approaches. Figures 17 and 20 shows the isosurfaces corresponding to a 90% probability of true positive detections extracted from the probability of true positive detections obtained by TAGI. The three axes in each figure represent the covariates vector $\mathbf{x} = [|h_{\text{a}}| \; w_{\text{a}} \; \mathtt{N}_{t>t_{\text{s}}}]^{\intercal}$ consisting in the absolute value of the anomaly height $|h_{\text{a}}|$, its duration w_{a}, and the number of time steps $\mathtt{N}_{t>t_{\text{s}}}$ required for an agent to trigger an alarm after its starts. The blue plane quantifies the minimum detectable anomaly magnitude $|h_a|_{\min}$. The yellow plane quantifies the minimum detectable anomaly growth rate per year $\psi_{\min} = \psi_0 \times |h_a|_{\min}$ as a function of its corresponding $|h_a|_{\min}$.

[Figure 17 about here.]

[Figure 18 about here.]

[Figure 19 about here.]

[Figure 20 about here.]

Table 1. Minimum detectable anomaly magnitudes $|h_a|_{\min}$ and the constant ψ_0 corresponding to the minimum detectable anomaly growth rates in a year such that $\psi_{\min} = \psi_0 \times |h_a|_{\min}$. The first column shows the name of the sensors. The second column is associated with the agents with the annual false positive ratios of $\zeta_{\mathrm{FP}} = 0.05$ and 0.1 per year. The following columns correspond to the metrics obtained from the imitation and reinforcement learning approaches; They are the minimum detectable anomaly magnitude $|h_a|_{\min}$, the constant ψ_0, and the minimum detectable anomaly growth rates $\psi_{\min}$ in years as a function of $|h_a|_{\min}$, i.e., $\psi_{\min} = |h_a|_{\min}$ per $\mathtt{N}_\psi$ years.

Sensor ID and name	ζ_{FP}	Imitation learning			Reinforcement learning	
		$\lvert h_a\rvert_{\min}$	ψ_0	$\mathtt{N}_\psi$ [year]	$\lvert h_a\rvert_{\min}$	$\mathtt{N}_\psi$ [year]
$\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer	0.05	0.0030 °d	0.08	13	0.0017 °d	7
	0.10	0.0035 °d	0.07	15	0.0020 °d	8
$\mathtt{I}_1$: Longitudinal inclinometer 1	0.05	0.0013 °d	0.19	6	0.0013 °d	6
	0.10	0.0013 °d	0.19	6	0.0013 °d	6
$\mathtt{I}_3$: Longitudinal inclinometer 3	0.05	0.0011 °d	0.23	5	0.0010 °d	4
	0.10	0.0012 °d	0.21	5	0.0010 °d	4
$\mathtt{E}$: Extensometer	0.05	0.0100 mm	0.15	7	0.0100 mm	7
	0.10	0.0090 mm	0.17	6	0.0090 mm	6

Table 2. Changes in structural responses, i.e, $\Delta\theta$ and $\Delta\ell$ due to different damage scenarios, associated with sensors installed on the bridge. The damage scenario names consist of two or three parts separated by dashes. They represent the damage's type, setup, and severity such that `COR`: corrosion, `BDL`: bearing device lockage, `PST`: partial settlement, `U`: uniform, and `NU`: non-uniform. Also, $\mathtt{I_L}$: Lateral inclinometer, $\mathtt{I_1}$: Longitudinal inclinometer 1, $\mathtt{I_3}$: Longitudinal inclinometer 3, and `E`: Extensometer.

Damage Scenario			$\Delta\theta$, °d			$\Delta\ell$, mm
ID	Type	Abbreviation	$\mathtt{I_L}$	$\mathtt{I_1}$	$\mathtt{I_3}$	`E`
1	Corrosion	`COR` − 2%	+0.0001	+0.0013	+0.0005	+0.0125
2		`COR` − 5%	+0.0003	+0.0033	+0.0013	+0.0321
3	Uniform bearing device lockage	`BDL` − `U` − 5.4%	0.0000	−0.0026	−0.0004	−0.0244
4		`BDL` − `U` − 14.6%	+0.0001	−0.0072	−0.0012	−0.0660
5	Non-uniform bearing device lockage	`BDL` − `NU` − 10.2%	+0.0013	+0.0002	+0.0003	−0.0135
6		`BDL` − `NU` − 22.2%	+0.0030	+0.0005	+0.0007	−0.0301
7		`BDL` − `NU` − 36.3%	+0.0051	+0.0007	+0.0012	−0.0512
8		`BDL` − `NU` − 74.1%	+0.0116	+0.0017	+0.0026	−0.1164
9	Uniform partial settlement	`PST` − `U` − 1 mm	0.0000	−0.0013	−0.0013	0.0000
10	Non-uniform partial settlement	`PST` − `NU` − 1 mm	−0.0030	−0.0010	−0.0009	+0.0021
11		`PST` − `NU` − 2 mm	−0.0059	−0.0020	−0.0018	+0.0041
12		`PST` − `NU` − 5 mm	−0.0147	−0.0050	−0.0045	+0.0103
		$\lvert h_a \rvert_{\min}$	0.0020	0.0013	0.0010	0.0090

Table 3. Optimized parameters for the BDLM models associated with the temperature data across all three spans.

Span	$\ell^{\mathtt{KR,T}}$	$\phi^{\mathtt{AR,T}}$	$\sigma^{\mathtt{AR,T}}$	$\sigma_v^{\mathtt{T}}$
$4-5$	0.46	0.07	2.61	0.10
$5-6$	0.46	0.08	2.65	0.10
$6-7$	0.45	0.07	2.64	0.10

Table 4. Optimized parameters for the BDLM models. The sensor names are $\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer, $\mathtt{I}_1$: Longitudinal inclinometer 1, $\mathtt{I}_3$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer

Sensor	Span	$\ell^{\mathtt{KR,I}}$	$\phi^{\mathtt{AR,I}}$	$\sigma^{\mathtt{AR,I}}$	Dependency	
					$\beta^{\mathtt{AR,I\vert T}} \times 10^{-4}$	$\beta^{\mathtt{KR,I\vert T}} \times 10^{-4}$
$\mathtt{I}_{\mathrm{L}}$	$4-5$	0.86	0.11	0.0010	-2.12	-2.96
	$5-6$	5.21	0.55	0.0010	-4.40	-1.39
	$6-7$	3.85	0.58	0.0020	-6.33	-2.89
$\mathtt{I}_1$	$4-5$	1.67	0.68	0.0004	-3.51	-3.77
	$5-6$	4.42	0.43	0.0009	-3.64	-4.16
	$6-7$	0.81	0.52	0.0005	-1.56	-2.63
$\mathtt{I}_3$	$4-5$	2.51	0.48	0.0004	$+5.11$	$+3.28$
	$5-6$	0.36	0.67	0.0003	-3.06	-4.42
	$6-7$	4.57	0.74	0.0004	-1.00	-1.64
					$\beta^{\mathtt{AR,E\vert T}}$	$\beta^{\mathtt{KR,E\vert T}}$
$\mathtt{E}$	$4-5$		0.82	0.0040	$+0.02$	$+0.02$
	$5-6$		0.82	0.0030	$+0.02$	$+0.02$
	$6-7$		0.80	0.0030	$+0.02$	$+0.02$

Table 5. Probability distributions and their parameters associated with the anomaly function for simulating new episodes. The only difference in values is related to the anomaly magnitude which depends on the sensor type.

Sensor type	Parameter	Value		Unit
		a	b	
Inclinometer	$H_\mathrm{a} \sim \mathcal{U}(h_\mathrm{a}; a, b)$	-0.005	$+0.005$	°d
Extensometer		-0.030	$+0.030$	mm
All types	$W_\mathrm{a} \sim \mathcal{U}(w_\mathrm{a}; a, b)$	0	2000	day
	$T_\mathrm{c} \sim \mathcal{U}(t_\mathrm{c}; a, b)$	Aug 15, 2019	Aug 13, 2029	date
	w_0	constant:	90	day
	ϵ_a	constant:	$0.25H_\mathrm{a}$	mm or °d
	ε	constant:	1×10^{-7}	mm or °d

Table 6. Summary of the datasets used in order to train ($\mathcal{D}^{\text{train}}$) and test ($\mathcal{D}^{\text{test, FP}}$ and $\mathcal{D}^{\text{test, TP}}$) each agent corresponding to its sensor type. The superscripts FP and TP are respectively associated with the evaluation of the annual false positive ratio and the probability of the true positive detections.

Data type	Notation	Episode			
		Total number	Duration, year	Anomaly	
				Duration, day	Ratio
Train	$\mathcal{D}^{\text{train}}$	1.2×10^5	30	2000	0.9
Test	$\mathcal{D}^{\text{test, FP}}$	3000	30	–	0.0
	$\mathcal{D}^{\text{test, TP}}$	6000	60	8000	1.0

Table 7. Configurations of the discretized environments with a grid size of $\mathtt{M} \times \mathtt{N}$ that are used to train the agents corresponding to the for inclinometers and extensometers.

Sensor type	Lower bound		Upper bound		Grid size ($\mathtt{M} \times \mathtt{N}$)	
	$\mu_t^{\mathtt{T}}$ ($\times 10^{-4}$)	$\pi_t^{\mathtt{ns}}$	$\mu_t^{\mathtt{T}}$ ($\times 10^{-4}$)	$\pi_t^{\mathtt{ns}}$	$\mu_t^{\mathtt{T}}$ ($\mathtt{M}$)	$\pi_t^{\mathtt{ns}}$ ($\mathtt{N}$)
Inclinometer	-0.5	0.0	$+0.5$	$+1.0$	160	60
Extensometer	-1.0	0.0	$+1.0$	$+1.0$		

Table 8. Reward values used to train the agents. Two reward sets are defined for each sensor type in order to obtain the annual false positive ratios of $\zeta_{\mathrm{FP}} = 0.05$ and 0.1 per year.

Sensor name	ζ_{FP}	Reward values			
		r^{TP}	r^{FP}	r^{TN}	η
$\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer	0.05	+4	−21	+1	+4
	0.10	+4	−15	+1	+4
$\mathtt{I}_1$: Longitudinal inclinometer 1	0.05	+4	−15	+1	+4
	0.10	+4	−9	+1	+4
$\mathtt{I}_3$: Longitudinal inclinometer 3	0.05	+4	−16	+1	+4
	0.10	+4	−10	+1	+4
$\mathtt{E}$: Extensometer	0.05	+4	−59	+1	+4
	0.10	+4	−26	+1	+4

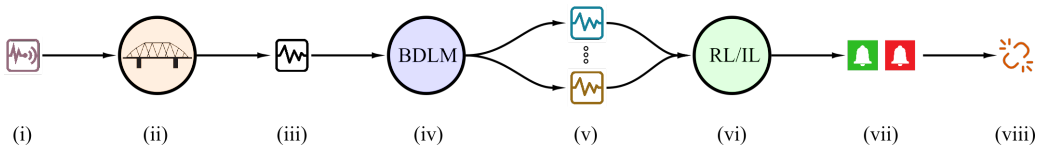


Figure 1. The general flow of this paper composed of eight distinct elements, namely (i) sensor technology, (ii) structural characteristics and sensors configuration, (iii) measurements, (iv) Bayesian dynamic liner models (BDLM), (v) measurement decomposition, (vi) reinforcement learning (RL) or limitation learning (IL), (vii) alarm triggering policies, and (viii) damage detection.

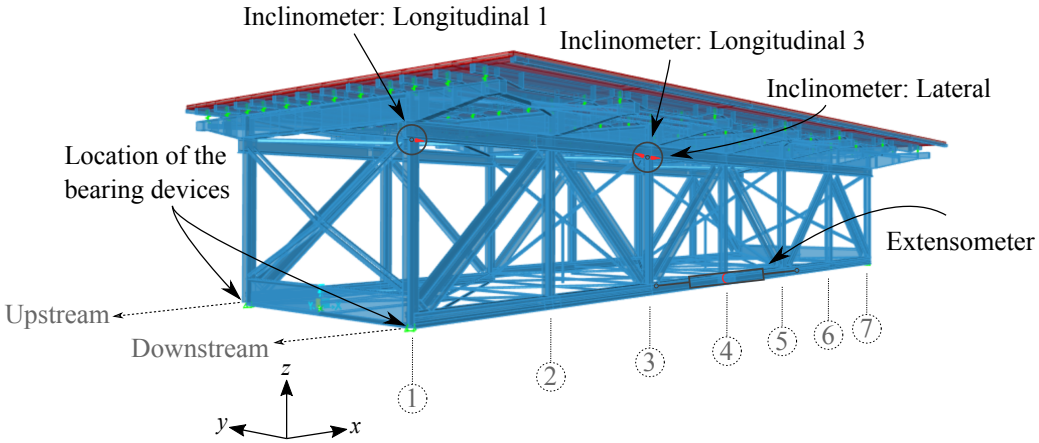


Figure 2. The finite element model and sensor configuration for each span: an extensometer on the section 4, two longitudinal inclinometers on the sections 1 and 3, a lateral inclinometer on the section 3, and a thermometer on the section 4. All the sensors are installed on the downstream side of the span. Note that the icons used for the sensor are only for illustrating the location of the sensors.

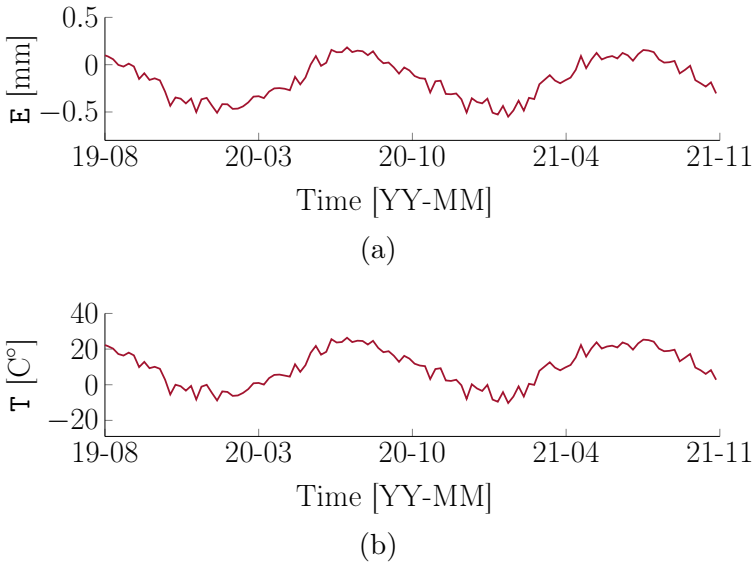


Figure 3. Illustration of the elongation and temperature data collected on the span $5-6$: (a) weekly averaged elongation data and (b) weekly averaged temperature data.

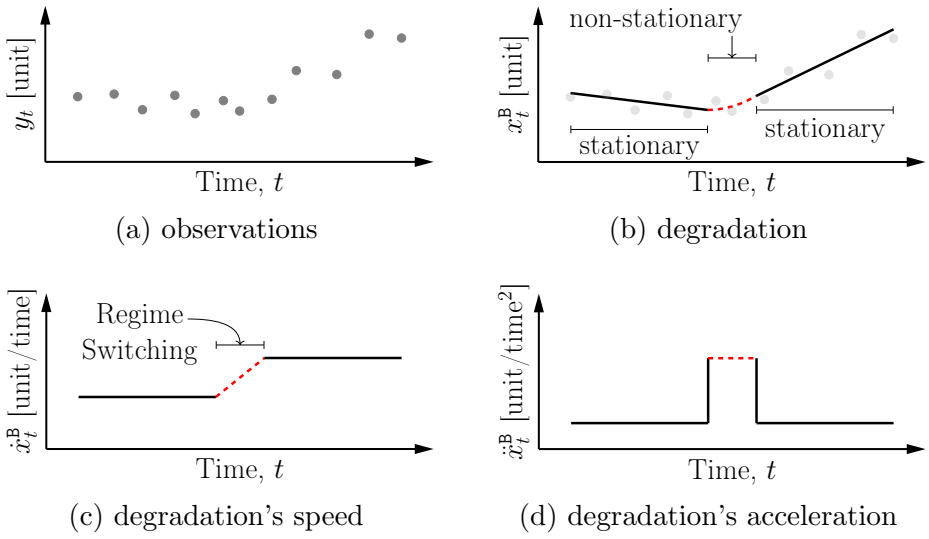
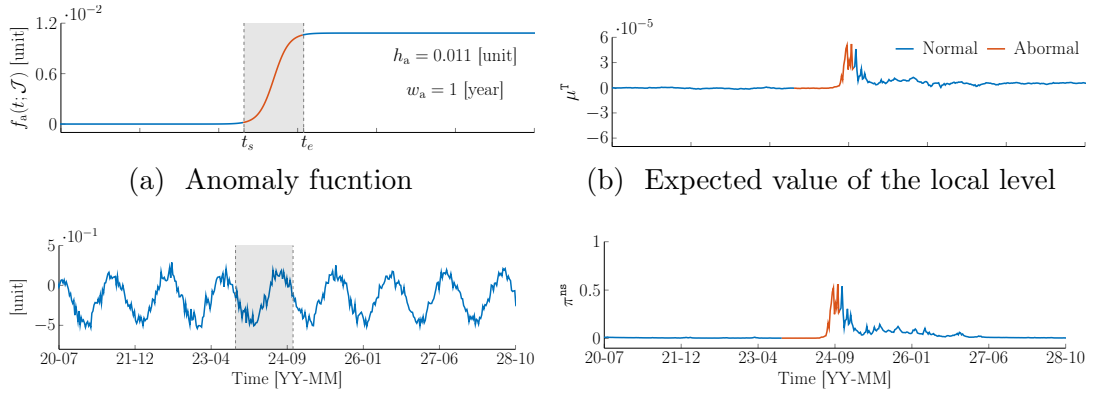


Figure 4. Illustration of the kinematics of two dynamic models including (a) observation, (b) structure's degradation (c) structure's degradation's speed, and (d) structure's degradation's acceleration.



(a) Anomaly fucntion (b) Expected value of the local level

(c) Simulated time series (d) Probability of the non-stationary regime

Figure 5. Example of an abnormal episode used by an expert to build a demonstration. The realization of the anomaly function results in an anomaly magnitude of $h_a = 0.11$ and a duration of $w_a = 1$ year. The shaded area represents the anomaly duration with the start t_s and end t_e times. The red and blue lines indicate the abnormal and normal time steps, respectively.

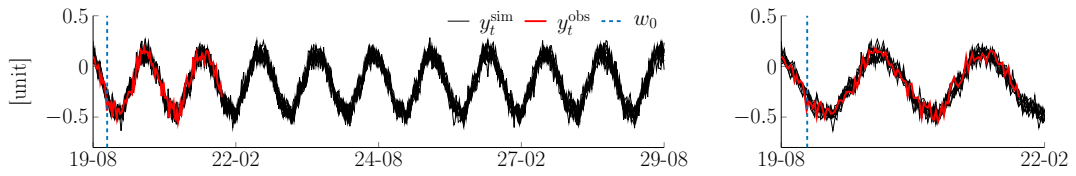


Figure 6. Example of simulated time series $\mathbf{y}_t^{\text{sim}}$ for an extensometer over the span of ten years (left) and the same simulated time series for a shorter period of time (right). The real measurement time series $\mathbf{y}_t^{\text{obs}}$ is indicated by the red line. The vertical blue dashed line corresponds to the number of days without anomaly, i.e., w_0.

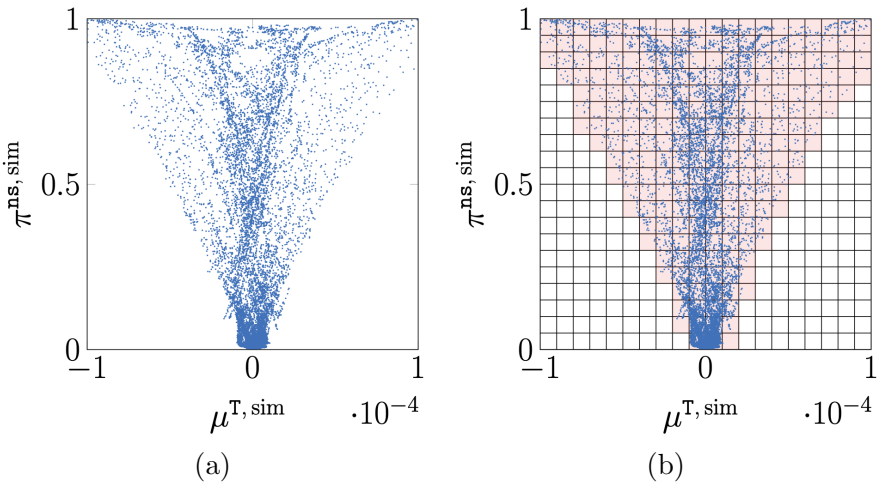


Figure 7. Example of an environment's state space along with the data points (blue dots) corresponding to the $(\mu_t^{\mathrm{T}}, \pi_t^{\mathrm{ns}})$ pairs obtained from the BDLM analysis: (a) the continuous states, and (b) the 20×20 grid along with the mapped data into the 400 induced discretized states $\hat{\mathbf{s}}$. The shaded regions represent the visited states $\hat{\mathbf{s}}$, while the white ones are associated with non-visited states $\neg\hat{\mathbf{s}}$.

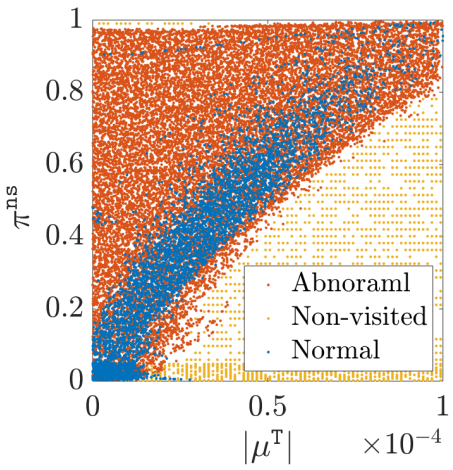
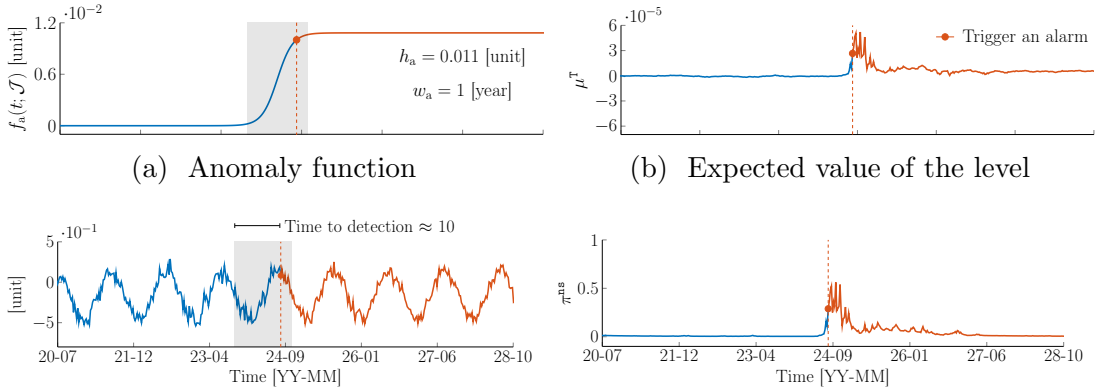


Figure 8. Illustrations of a training demonstration set. The red dots are associated with the abnormal visited states corresponding to the time steps during the anomaly for which the expert's decision is a^0 : trigger an alarm. The blue dots are associated with the normal visited states corresponding to the remaining time steps for which the expert's decision is a^1 : not trigger an alarm. The yellow dots represent the non-visited states for which the agent is forced to take the action of triggering an alarm by imitating the expert's decision.



(a) Anomaly function

(b) Expected value of the level

(c) Simulated time series

(d) Probability of the non-stationary regime

Figure 9. Example of an abnormal episode used to build a test dataset $\mathcal{D}$. The realization of the anomaly function results in an anomaly magnitude of $h_a = 0.11$ and a duration of $w_a = 1$ year. The shaded area represents the anomaly duration for which the agent triggers an alarm approximately 10 months $\equiv 0.83$ year after the anomaly starts. The point where the alarm is triggered is identified by a circle marker and a vertical dashed line. The covariates vector is $\mathbf{x}_i = [0.011\ 1\ 0.83]^\intercal$.

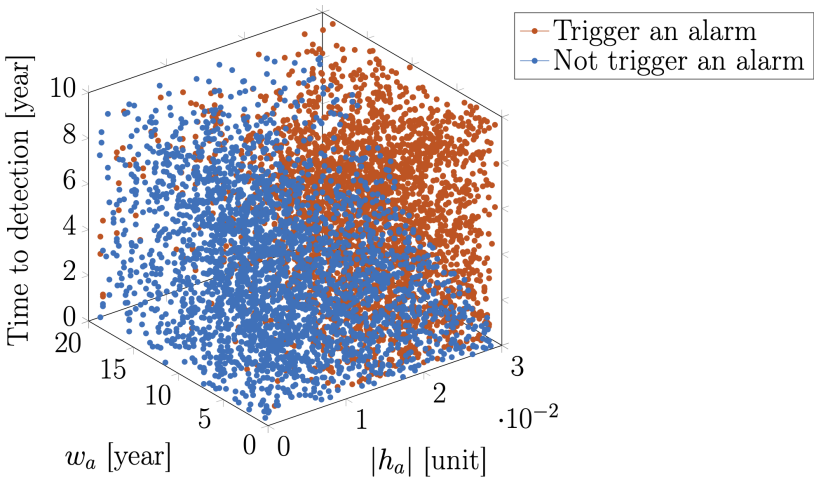
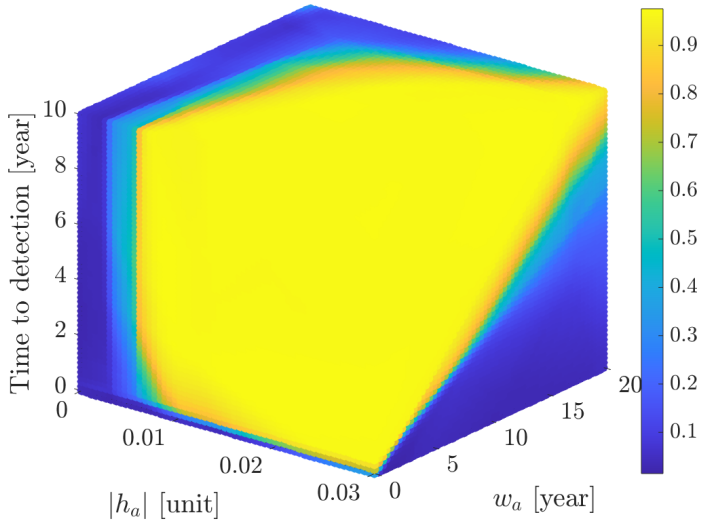
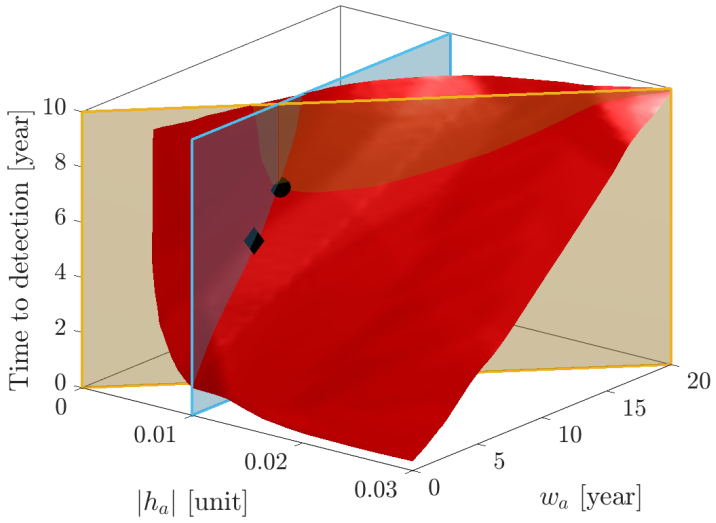


Figure 10. Example of the joint observations $\mathcal{D} = \{(\mathbf{x}_i, y_i), \forall i \in \{1 : \mathtt{D}\}\}$ from 1200 episodes, where the covariates vector $\mathbf{x}_i = [|h_{\mathrm{a}}| \ w_{\mathrm{a}} \ \mathtt{N}_{t>t_{\mathrm{s}}}]^{\intercal}$, and the classes y_i are $-1 \equiv$ false negatives (blue dots) and $+1 \equiv$ true positives (red dots).



(a)



(b)

Figure 11. Example of a model trained using TAGI: (a) probability of true positive detections with respect to prediction covariates vector $\mathbf{x}$, and (b) the isosurface corresponding to 90% probability of true positive detections. The vertical planes indicate the minimum detectable anomaly magnitude $|h_a|_{\min} = 0.01$ (blue plane) and the minimum detectable anomaly growth rate per year $\psi_{\min} = 0.15|h_a|_{\min}$ (yellow plane). The black circle and diamond identify two points located at the intersection of the isosurface and the vertical blue plane.

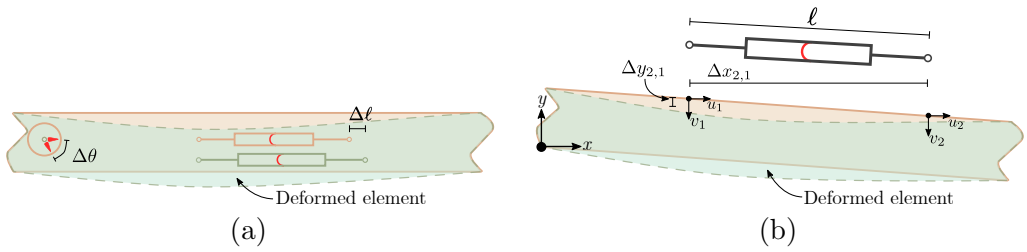
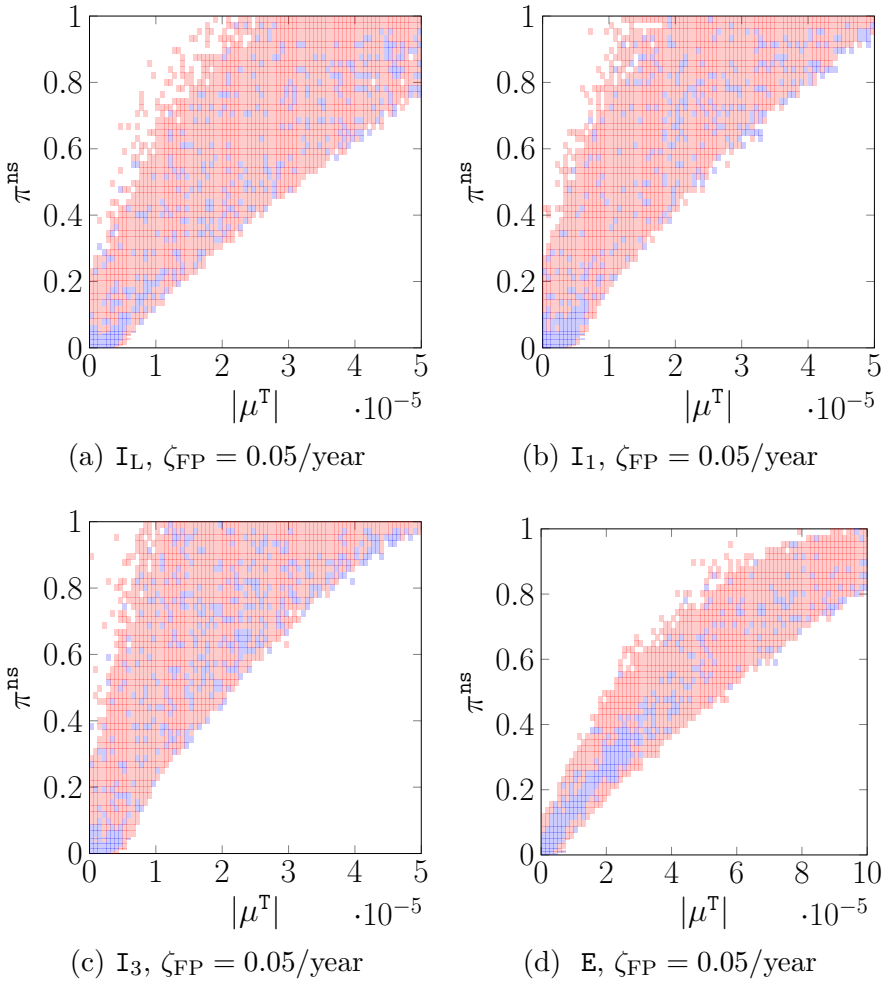


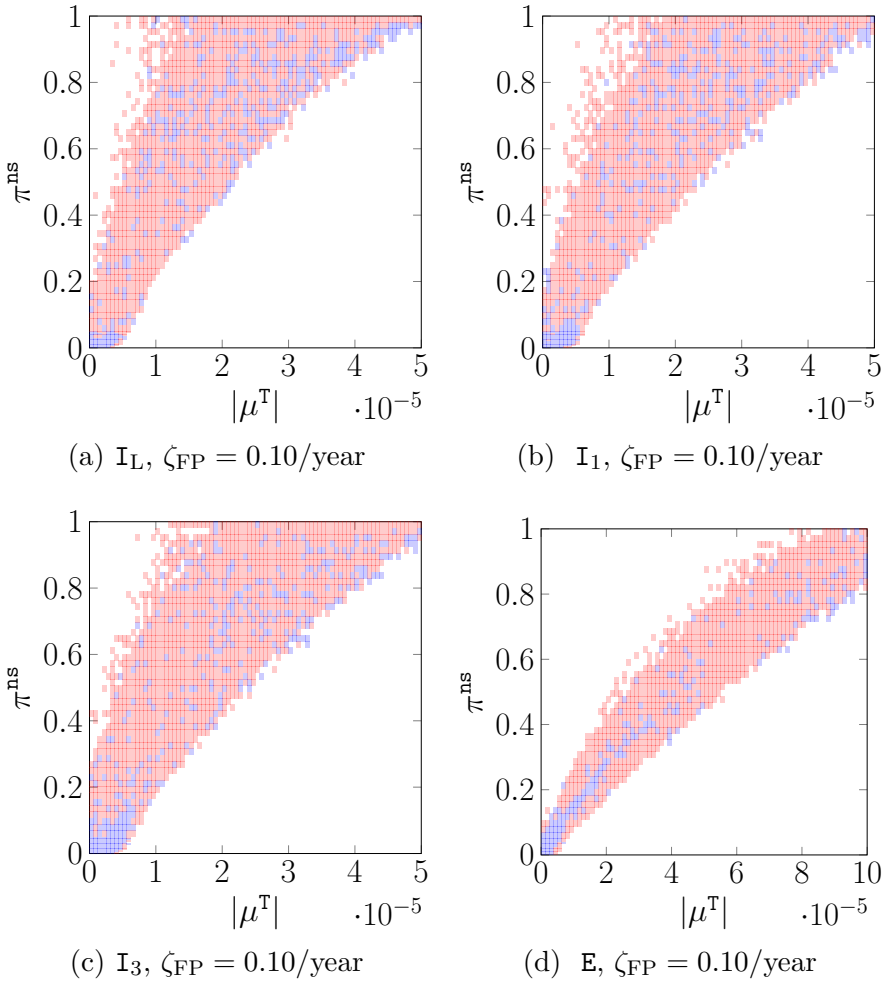
Figure 12. Illustration of (a) the deformation of an element and changes in the rotation of an inclinometer and the elongation of an extensometer, and (b) the displacements associated with the two nodes at the location of the supports of the extensometer on an structural element. $\Delta\theta$ and $\Delta\ell$ are the respective changes in structural responses at the location of an inclinometer and an extensometer. $\Delta x_{2,1} = x_2 - x_1$ and $\Delta y_{2,1} = y_2 - y_1$ represent the nodal coordinates differences in the respective x and y directions. u_i and vertical v_i are the respective nodal displacements at the location of the extensometer supports.



(a) $\mathtt{I}_{\mathrm{L}}$, $\zeta_{\mathrm{FP}} = 0.05$/year (b) $\mathtt{I}_1$, $\zeta_{\mathrm{FP}} = 0.05$/year

(c) $\mathtt{I}_3$, $\zeta_{\mathrm{FP}} = 0.05$/year (d) $\mathtt{E}$, $\zeta_{\mathrm{FP}} = 0.05$/year

Figure 13. The policy maps for the inclinometers and the extensometers for $\zeta_{\mathrm{FP}} = 0.05$/year. For a better illustration, the grid lines are not shown. The red colour indicates that the action is a^0: Trigger an alarm, and the blue one indicates that the action is a^1: Not trigger an alarm. The white areas are associated with the non-visited states. Also, $\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer, $\mathtt{I}_1$: Longitudinal inclinometer 1, $\mathtt{I}_3$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer.



(a) $\mathtt{I}_{\mathrm{L}}$, $\zeta_{\mathrm{FP}} = 0.10$/year (b) $\mathtt{I}_1$, $\zeta_{\mathrm{FP}} = 0.10$/year

(c) $\mathtt{I}_3$, $\zeta_{\mathrm{FP}} = 0.10$/year (d) $\mathtt{E}$, $\zeta_{\mathrm{FP}} = 0.10$/year

Figure 14. The policy maps for the inclinometers and the extensometers for $\zeta_{\mathrm{FP}} = 0.10$/year. For a better illustration, the grid lines are not shown. The red colour indicates that the action is a^0: Trigger an alarm, and the blue one indicates that the action is a^1: Not trigger an alarm. The white areas are associated with the non-visited states. Also, $\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer, $\mathtt{I}_1$: Longitudinal inclinometer 1, $\mathtt{I}_3$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer.

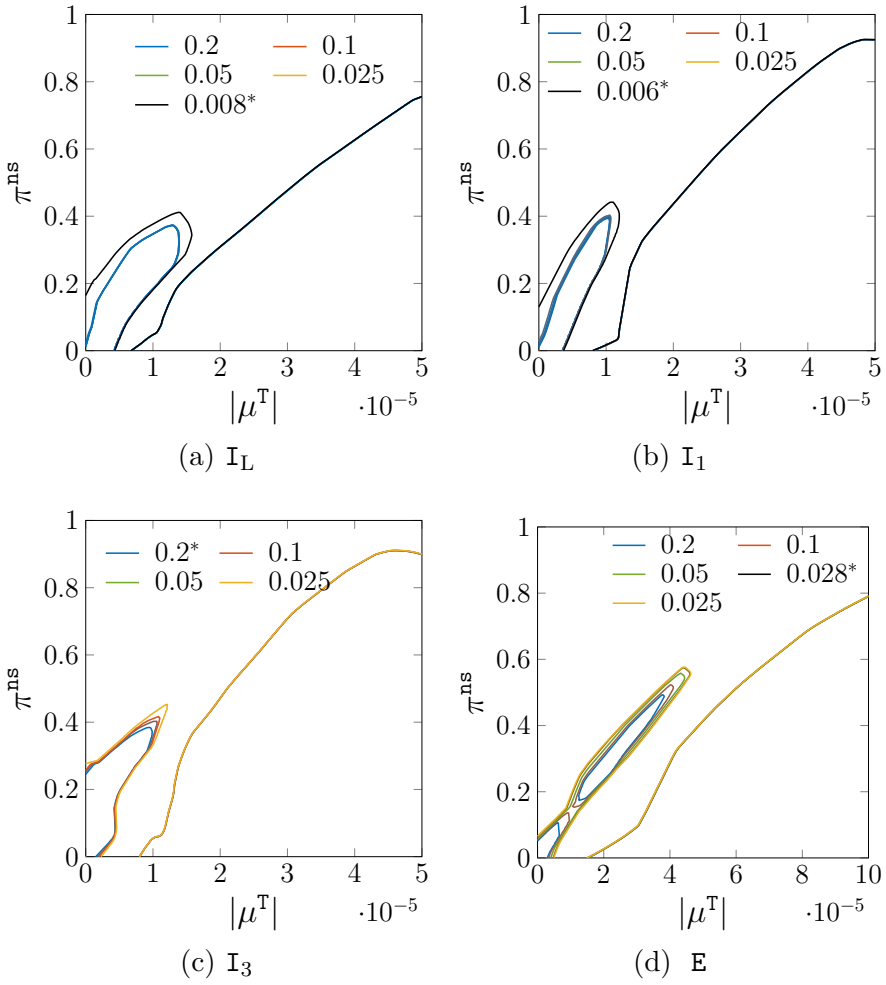
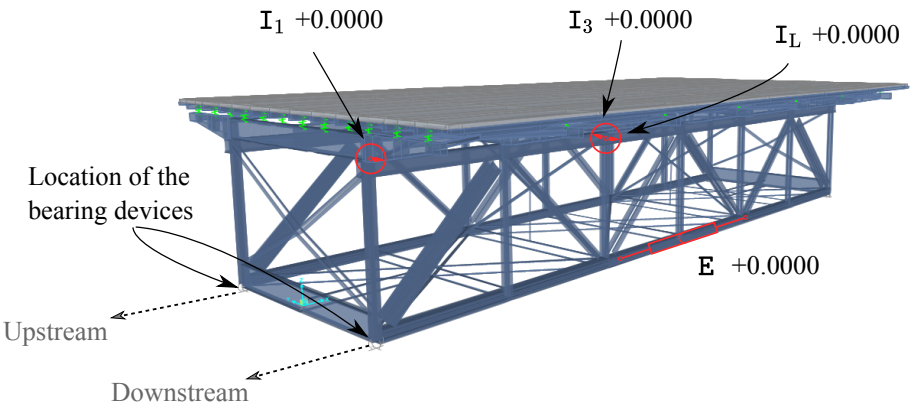
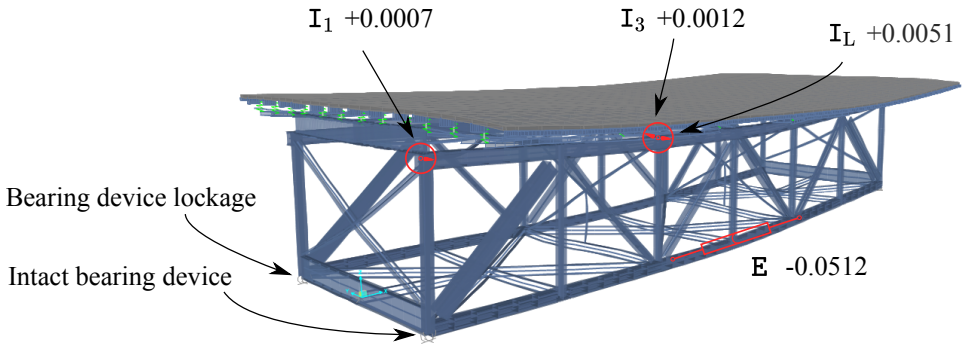


Figure 15. The policy map contours for the inclinometers and the extensometers. The contour lines represent the annual false positive ratios ζ_{FP} of 0.2, 0.1, 0.05, and 0.025 per year. The black contour lines indicate the probability of triggering alarms $\Pr(Y = -1 : a^0|\mathbf{s}) > 0.5$, where the corresponding annual false positive ratio values are indicated by a superscript $*$. Also, $\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer, $\mathtt{I}_1$: Longitudinal inclinometer 1, $\mathtt{I}_3$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer.

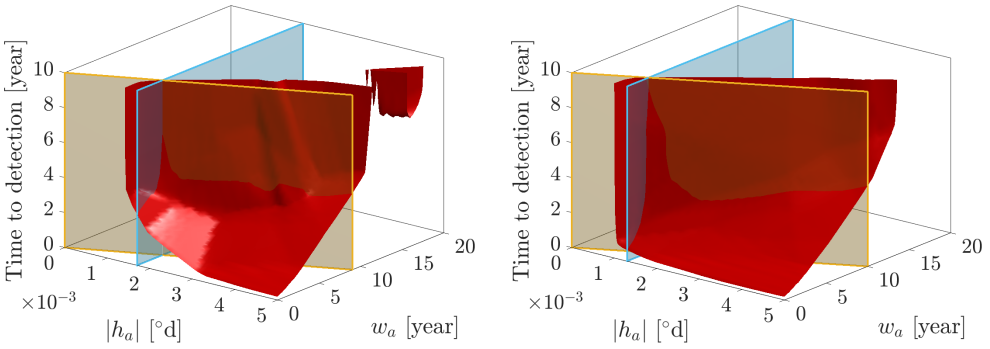


(a) Intact structure

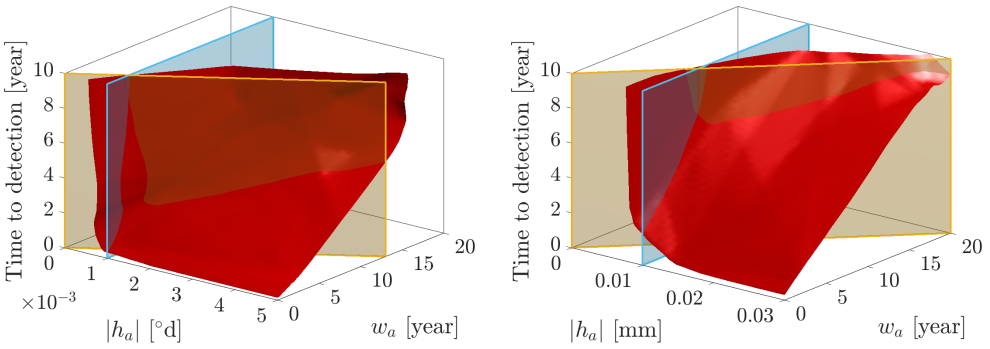


(b) Damaged structure

Figure 16. Effects of a 36.3% non-uniform bearing device lockage on the changes in relative structural responses $\Delta\theta$ and $\Delta\ell$ at the location of the inclinometers and the extensometer. Also, $\mathtt{I_L}$: Lateral inclinometer, $\mathtt{I_1}$: Longitudinal inclinometer 1, $\mathtt{I_3}$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer.

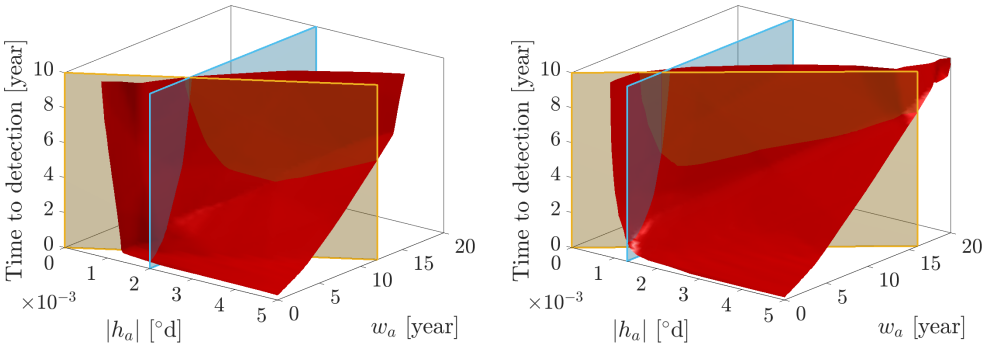


(a) $\mathtt{I}_{\mathrm{L}}$, $\zeta_{\mathrm{FP}} = 0.05$/year (b) $\mathtt{I}_1$, $\zeta_{\mathrm{FP}} = 0.05$/year

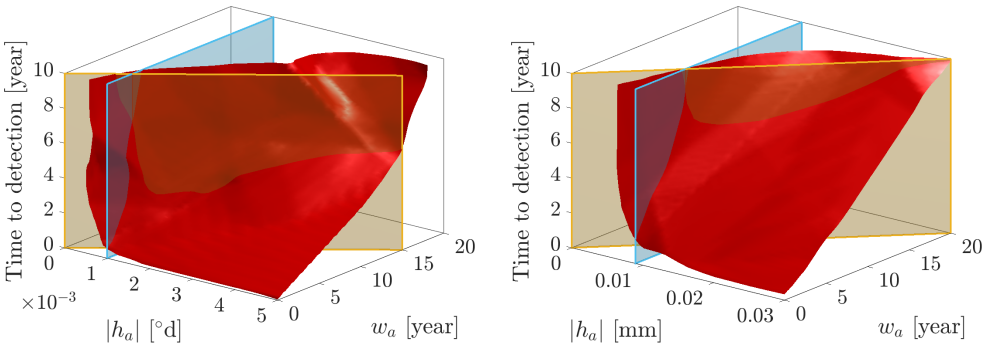


(c) $\mathtt{I}_3$, $\zeta_{\mathrm{FP}} = 0.05$/year (d) $\mathtt{E}$, $\zeta_{\mathrm{FP}} = 0.05$/year

Figure 17. Illustration of the 90% probability for true positive detections for $\zeta_{\mathrm{FP}} = 0.05$/year. The sensor names are $\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer, $\mathtt{I}_1$: Longitudinal inclinometer 1, $\mathtt{I}_3$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer.The blue plane quantifies the minimum detectable anomaly magnitude $|h_a|_{\min}$. The yellow plane quantifies the minimum detectable anomaly growth rate per year $\psi_{\min} = \psi_0 \times |h_a|_{\min}$ as a function of its corresponding $|h_a|_{\min}$.

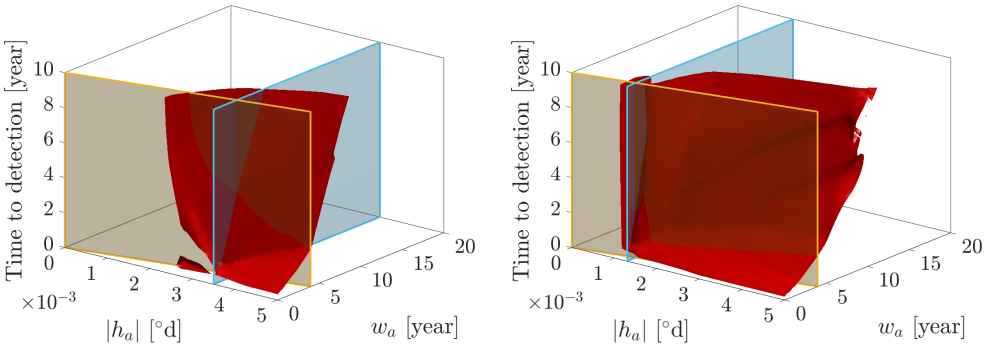


(a) $\mathtt{I}_{\mathrm{L}}$, $\zeta_{\mathrm{FP}} = 0.10$/year (b) $\mathtt{I}_1$, $\zeta_{\mathrm{FP}} = 0.10$/year

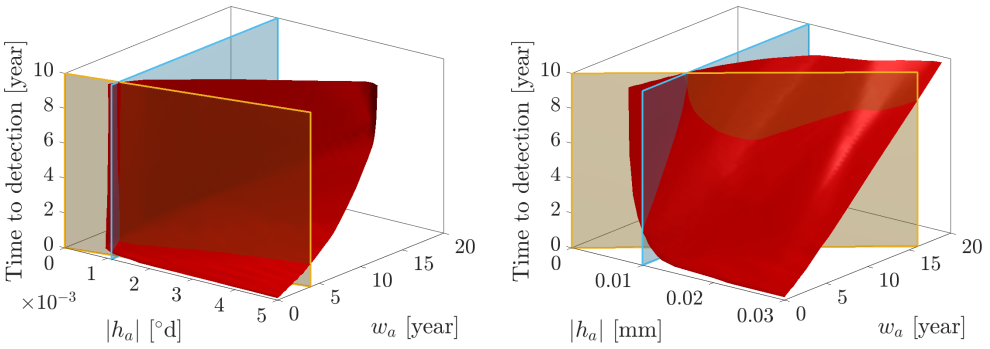


(c) $\mathtt{I}_3$, $\zeta_{\mathrm{FP}} = 0.10$/year (d) $\mathtt{E}$, $\zeta_{\mathrm{FP}} = 0.10$/year

Figure 18. Illustration of the 90% probability for true positive detections for $\zeta_{\mathrm{FP}} = 0.1$/year. The sensor names are $\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer, $\mathtt{I}_1$: Longitudinal inclinometer 1, $\mathtt{I}_3$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer.The blue plane quantifies the minimum detectable anomaly magnitude $|h_a|_{\min}$. The yellow plane quantifies the minimum detectable anomaly growth rate per year $\psi_{\min} = \psi_0 \times |h_a|_{\min}$ as a function of its corresponding $|h_a|_{\min}$.

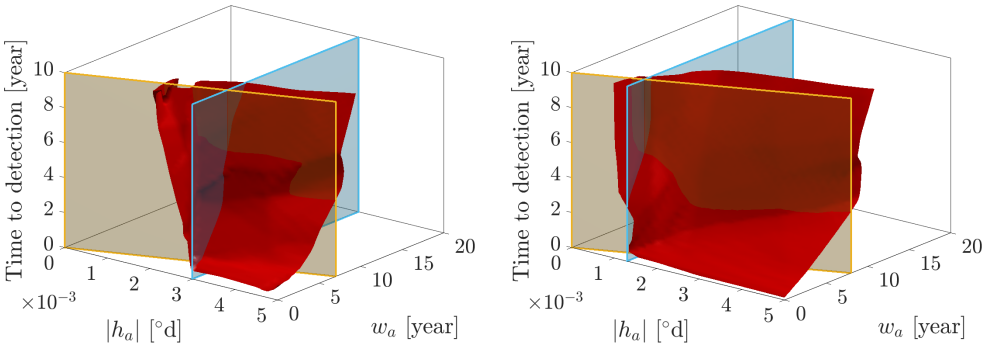


(a) $\mathtt{I}_{\mathrm{L}}$, $\zeta_{\mathrm{FP}} = 0.05$/year (b) $\mathtt{I}_1$, $\zeta_{\mathrm{FP}} = 0.05$/year

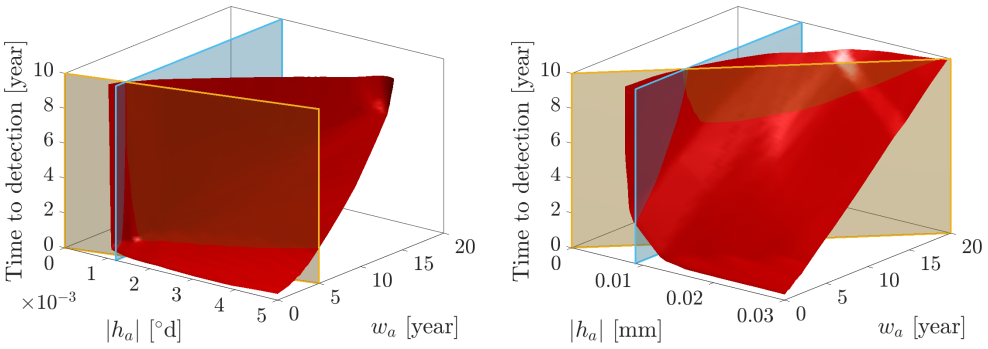


(c) $\mathtt{I}_3$, $\zeta_{\mathrm{FP}} = 0.05$/year (d) $\mathtt{E}$, $\zeta_{\mathrm{FP}} = 0.05$/year

Figure 19. Illustration of the 90% probability for true positive detections for $\zeta_{\mathrm{FP}} = 0.05$/year. The sensor names are $\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer, $\mathtt{I}_1$: Longitudinal inclinometer 1, $\mathtt{I}_3$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer.The blue plane quantifies the minimum detectable anomaly magnitude $|h_a|_{\min}$. The yellow plane quantifies the minimum detectable anomaly growth rate per year $\psi_{\min} = \psi_0 \times |h_a|_{\min}$ as a function of its corresponding $|h_a|_{\min}$.



(a) $\mathtt{I}_{\mathrm{L}}$, $\zeta_{\mathrm{FP}} = 0.10$/year (b) $\mathtt{I}_1$, $\zeta_{\mathrm{FP}} = 0.10$/year



(c) $\mathtt{I}_3$, $\zeta_{\mathrm{FP}} = 0.10$/year (d) $\mathtt{E}$, $\zeta_{\mathrm{FP}} = 0.10$/year

Figure 20. Illustration of the 90% probability for true positive detections for $\zeta_{\mathrm{FP}} = 0.1$/year. The sensor names are $\mathtt{I}_{\mathrm{L}}$: Lateral inclinometer, $\mathtt{I}_1$: Longitudinal inclinometer 1, $\mathtt{I}_3$: Longitudinal inclinometer 3, and $\mathtt{E}$: Extensometer.The blue plane quantifies the minimum detectable anomaly magnitude $|h_a|_{\min}$. The yellow plane quantifies the minimum detectable anomaly growth rate per year $\psi_{\min} = \psi_0 \times |h_a|_{\min}$ as a function of its corresponding $|h_a|_{\min}$.