buildings

MDPI

Article

Bayesian Filtering for Temperature-Compensated Damage Detection in Hysteretic Structures: A Monte Carlo Comparison of Particle and Kalman-Type Filters

Lucio Pinello * and Luca Lomazzi

Department of Mechanical Engineering, Politecnico di Milano, via La Masa 1, 20156 Milano, Italy; luca.lomazzi@polimi.it

* Correspondence: lucio.pinello@polimi.it

Abstract

Online estimation of stiffness parameters from vibration data is a common building block of temperature-compensated damage detection. This study compares Bayesian filters for this task on a three-degree-of-freedom structure with a Bouc–Wen hysteretic element under varying temperatures: sequential importance resampling particle filters (SIR PFs), an action-aided particle filter (A-SIR PF) and its reformulations, a damage-factor PF, and extended, unscented and ensemble Kalman filters. Tracking, detection and localisation are assessed over paired Monte Carlo simulations of step and progressive degradation, at equal false-alarm rates. With the same model, the Kalman-type filters give the lowest errors, and adaptive systematic resampling brings the PFs to within 2–6% of them. Prior knowledge helps when it enters the transition model: a damage-factor PF, which includes a temperature–stiffness law learned from healthy data and the constraint that damage cannot increase the stiffness, is the most accurate filter, with errors on the healthy springs about 20% lower than those of the Kalman-type filters. The least-action principle, used as the objective of trajectory estimation, leads to a fixed-lag smoother that reduces the errors of the best filter by 25–62%, with a 1 s delay. At large thermal excursions, the environmental normalisation is necessary, and an autoencoder with standardised inputs and a conditional regression of the stiffness on the temperature are equally effective.

Keywords: Bayesian filtering; particle filter; damage factor; Kalman filter; smoothing; action-aided particle filter; Bouc–Wen hysteresis; environmental normalisation; damage detection

Check for updates

Academic Editors: Gaetano Miraglia and Stefania Coccimiglio

Received: 23 September 2026
Revised: 3 November 2026
Accepted: 12 November 2026
Published: 17 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Structural and mechanical systems progressively deteriorate during their service life because of aging, accumulated cyclic loading, fatigue, creep, corrosion, accidental events, and changes in operating conditions. Conventional inspection and scheduled-maintenance strategies provide only intermittent information about structural condition and may require costly shutdowns or intrusive access. Structural Health Monitoring (SHM) addresses these limitations by integrating sensing technologies, signal processing, system identification, statistical inference, and decision-support methodologies. The objective of SHM is not limited to the detection of anomalous structural behaviour, but also encompasses damage localisation, severity assessment, prognosis, and the generation of information supporting risk-informed operation and maintenance decisions [1,2].

Vibration-based SHM is particularly attractive because changes in structural properties, including stiffness, damping, and boundary conditions, are reflected in the measured dynamic response [3]. Numerous methods have therefore been proposed using time-domain signals, modal parameters, frequency-response quantities, transmissibility functions, wavelet coefficients, cepstral features, and other time-frequency representations [4].

More recently, machine-learning methods have enabled damage-sensitive representations to be learned directly from high-dimensional monitoring data [1]. However, classical feature-based methods are often unable to distinguish between damage-induced variations and those caused by environmental and operational changes, such as temperature fluctuations or varying loads, potentially leading to false alarms [5–7].

Environmental and operational variability represents one of the main obstacles to the transition of SHM methods from controlled laboratory experiments to long-term field applications. Consequently, a healthy structure operating under conditions that were not adequately represented in the reference data may be incorrectly classified as damaged [8]. Long-duration monitoring studies have demonstrated that uncontrolled environmental effects may significantly distort damage-sensitive signals and increase false-alarm rates [9]. Hybrid numerical and experimental frameworks have therefore been developed to represent a broader range of structural and environmental conditions during the design and validation of diagnostic algorithms [10]. In addition, gradual changes in the environmental distribution may produce concept drift, rendering a previously established healthy baseline progressively unrepresentative of the monitored structure [8].

To mitigate this issue, data-normalisation methods, such as principal component analysis [11,12] and cointegration [13–15], have been proposed, although they are often problem-specific and limited in their generalisation capability. For the same purpose, temperature-compensated autoencoders have also demonstrated the ability to detect and localise damage using time-domain signals acquired under varying thermal conditions [16]. Recent studies have introduced generalised and mechanics-informed autoencoders to improve robustness against excitation variability and to exploit known relationships among structural responses [17,18]. Similarly, approaches based on Bayesian frameworks have recently been employed for structural parameter estimation and damage diagnosis [19].

Among these, particle filters (PFs) are particularly attractive due to their ability to handle non-linear dynamics and non-Gaussian noise. Within a Bayesian state-space formulation, the PF approximates the posterior probability density function of the unknown system state through a finite set of weighted Monte Carlo samples. As new observations become available, these particles are sequentially propagated, weighted, and resampled, enabling the online assimilation of monitoring data without requiring the restrictive linearity and Gaussian assumptions underlying conventional filtering techniques [20]. This flexibility makes PF particularly suitable for prognostics and health management applications, in which the quantities of interest, such as the degradation state and the parameters governing its evolution, are generally uncertain and cannot be directly measured [21–26]. This capability has supported the application of PF to different degradation phenomena and engineering systems, including fatigue-damage evolution, structural monitoring, and energy-storage systems [21–25]. In SHM applications, PF has been employed to identify both time-dependent states and nominally time-invariant parameters in nonlinear and non-Gaussian systems. Nevertheless, the repeated resampling of particles may progressively reduce the diversity of the population, especially when parameters with weak or artificial dynamics must be estimated. To address this limitation, mutation operators have been incorporated into the filtering procedure to restore particle diversity and improve parameter identification under demanding noise conditions [27]. The performance of

PF-based prognosis also depends strongly on the formulation of the state-transition model and on the stochastic terms introduced to represent modelling uncertainty. Consequently, the process noise must not be regarded solely as a numerical perturbation, but should be formulated consistently with the expected characteristics and constraints of the degradation process [22].

More generally, probabilistic life-prediction approaches highlight the importance of correctly representing the uncertainty associated with initial conditions, degradation parameters, and interactions among multiple evolving defects [28]. Particle filters have also been combined with data-driven models to reduce the dependence of prognostic frameworks on a fully specified physics-based representation. In battery prognosis, for example, radial-basis-function and multilayer-perceptron neural networks have been trained online within the particle-filtering framework [23,24]. In these approaches, the particles represent the parameters of the surrogate model, which are progressively adapted as new observations become available. This integration provides the flexibility required to follow changes in the monitored dynamics and to update the corresponding end-of-discharge or remaining-life predictions under anomalous or previously unobserved operating conditions [23,24]. Similar motivations arise when degradation takes place under time-varying operating conditions, for which the particle filter can be used to infer evolving degradation-model parameters and update prognostic estimates as the operational regime changes [26].

Also, PFs have been successfully applied to identify stiffness variations in multiple-degree-of-freedom (MDOF) systems subjected to seismic-like excitations [29,30]. Specifically, ref. [30] proposed the combination of PF-based stiffness estimation with autoencoders (AEs) trained on healthy-state data to neutralise temperature effects in the damage diagnosis of MDOF systems. The approach was subsequently investigated for an MDOF structure containing a nonlinear hysteretic component and subjected to changing temperature conditions [31]. This capability to model hysteretic behaviour is especially relevant to civil engineering applications, where structures such as reinforced concrete frames may exhibit pronounced hysteresis under dynamic loading due to components with highly non-linear characteristics. Such behaviour is commonly described by the Bouc–Wen model [32,33], widely adopted in parameter identification studies because it provides a compact phenomenological representation of smooth hysteresis. It can also be extended to describe stiffness or strength degradation, pinching, asymmetric response, and other complex hysteretic phenomena [34]. The associated parameters and internal hysteretic variables are generally not directly observable, and may exhibit strong correlations and limited identifiability. Their online estimation therefore constitutes a demanding inverse problem [35,36].

Recent studies have addressed hysteretic parameter identification using attention-based recurrent neural networks and dimensionality-reduction techniques [37], while physics-enhanced recurrent architectures have embedded state-space equations and numerical integration schemes into data-driven structural-response models [38]. Thus, combining domain knowledge, sequential inference, and data-driven representations appears to be a promising approach for improving the identification of nonlinear structural behaviour from noisy and incomplete measurements.

Kalman-type filters are the main alternative to particle filters for the joint estimation of states and parameters in structural dynamics. The extended Kalman filter (EKF) linearises the model around the current estimate. Instead, the unscented Kalman filter (UKF) propagates a deterministic set of sigma points through the nonlinear model [39], while the ensemble Kalman filter (EnKF) estimates the covariances required by the Kalman update from a random ensemble of model runs [40,41]. In structural system identification, the UKF was found to be more accurate than the EKF for strongly nonlinear systems [42], and Kalman-type and particle filters have been compared for hysteretic and other nonlinear

structural models [29,43]. Particle filters are asymptotically exact for nonlinear and non-Gaussian models, but their performance depends on design choices that are often reported only implicitly, namely: (i) the resampling scheme and the frequency of resampling [44–46], (ii) the treatment of static parameters, for which artificial dynamics or kernel shrinkage are introduced [47,48], and (iii) the flattening of the likelihood when the weights degenerate [49]. A further modification of the particle weights was proposed by the authors in Ref. [50], namely the Action-aided Particle Filter (A-SIR PF). Inspired by the principle of least action, the A-SIR PF divides the likelihood-based weights by an action score accumulated along the trajectory of each particle, with the aim of promoting dynamically coherent trajectories and penalising non-physical evolutions.

A further way of improving a filter is to include prior knowledge of the structure in the state-space model itself. In vibration-based monitoring, damage is commonly represented as a reduction in the stiffness parameters with respect to those of the undamaged structure [51,52], and the environmental variability of the healthy structure can be learned from data acquired in the undamaged state. For instance, this is the case for the temperature dependence of the natural frequencies of the Z24 bridge [53]. In addition, the stiffness parameters of updated finite element models have been shown to correlate with temperature, and removing this effect reduces their variability [54]. Similarly, hierarchical Bayesian model updating accounts for the variability of the parameters under changing environmental conditions [55]. Knowledge of this kind often takes the form of constraints on the state, such as a stiffness that cannot increase because of damage or a monotonic degradation [22]. Constraints are information that the Kalman filter does not incorporate. They can be included in Kalman-type filters in several ways, e.g., by projecting the estimate onto the admissible set, but these approaches are in general not optimal when the system or the constraints are nonlinear [56]. In contrast, particle filters can impose the constraints on the individual particles, e.g., through their propagation or their weights [56–58].

Comparisons between filters in SHM studies are frequently based on one or a few simulations, i.e., on a few draws of the measurement noise and of the filter randomness. In addition, the damage-detection performance is often assessed with thresholds that differ between the compared methods. However, the output of a particle filter is itself random, and rare divergent runs dominate thresholded metrics. Thus, such comparisons may lead to conclusions that do not survive repetition.

The present manuscript extends two previous works. Ref. [31] introduced the combination of PF-based stiffness estimation and AEs for the hysteretic benchmark considered here, using a conventional SIR PF, while ref. [50] proposed the action-aided weighting and illustrated it on a single simulation of a progressive-damage scenario. In this work, the question is broadened from whether the action-aided weighting improves the SIR PF to which Bayesian filter should provide the stiffness estimates of a temperature-compensated damage-detection chain for a hysteretic structure, and to how such filters should be compared. The main contributions of this work are as follows:

- A paired Monte Carlo comparison of SIR PFs with resampling at every step and with adaptive systematic resampling, of the A-SIR PF with both schemes and two reformulations of the action-aided idea, of tempered and Liu–West particle filters, and of the EKF, UKF and EnKF, for step and progressive degradation. All the filters use the same data, model and noise covariances.
- A complete detection and localisation chain, in which the normalisation models are trained on the estimates provided by each filter on independent healthy simulations. The thresholds are calibrated at equal false-alarm rates, and the comparison is complemented by ROC analysis and paired statistical tests.

- An analysis of the action-aided weighting and three reformulations of the least-action idea: an incremental weighting, an action-consistent prior on the stiffness trajectories and a least-action smoother. The last two are derived from the path-integral form of the Bayesian posterior.
- A damage-factor particle filter, in which a temperature–stiffness law learned from healthy data and the physical constraint that damage cannot increase the stiffness enter the transition model. This filter is compared with an EKF that uses the same model.
- An assessment of the effect of the resampling scheme, of the excitation (harmonic or broadband) and of the environmental normalisation (AE or conditional regression), together with robustness studies with respect to the number of particles, measurement noise, model-form mismatch, sensor layout, damage location, temperature range and drift rate.

The manuscript is organised as follows. Section 2 describes the model, the filters, the environmental normalisation, the damage indicators and the assessment criteria. Section 3 presents the case study and the Monte Carlo design. Section 4 discusses the results, including the limitations of the study. Finally, Section 5 summarises the conclusions and outlines directions for future work.

2. Methodology

2.1. State-Space Model

This study refers to a multi-degree-of-freedom (MDOF) mass–spring–damper system, as depicted in Figure 1 and presented in Ref. [31].

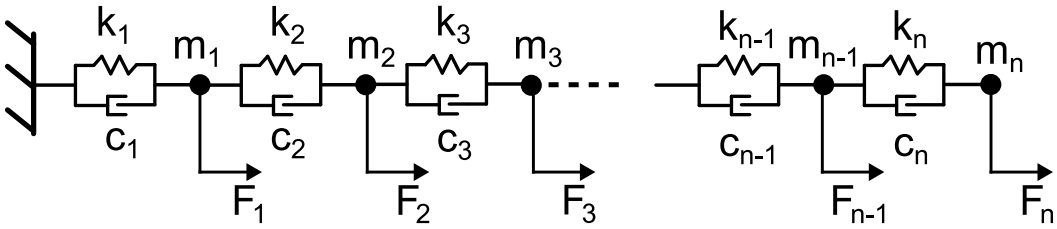


Figure 1. Schematic of a generic MDOF mass–spring–damper system.

The following assumptions are made:

- Damping is proportional to a reference stiffness value through a coefficient β [59], as commonly done in the literature [31,60].
- The first degree-of-freedom is affected by a degrading hysteretic behaviour, described by the Bouc–Wen formula reported in Equation (1).

$$\dot{r}(t) = \dot{x} - \beta_{BW}|\dot{x}||r|^{n-1}r - \gamma\dot{x}|r|^{n} \tag{1}$$

The variables of the MDOF system are the position x and speed $\dot{x}$ of each mass, the stiffnesses k, the proportionality damping constant β, the Bouc–Wen parameters β_{BW}, γ, and n, and the hysteretic behaviour r. Thus, the vector $\mathbf{x}^{\tau}$ containing the state variables can be defined as in Equation (2), where x_i^{τ} is the position of the $i-th$ mass at the generic time step τ, v_i^{τ} is the speed of the $i-th$ mass at the generic time step τ, k_i^{τ} is the stiffness of the $i-th$ spring at the generic time step τ, β is the proportionality coefficient of damping, according to the relationship $c_i = \beta k_i$, β_{BW}, γ, and n are the Bouc–Wen parameters, and r is the hysteretic displacement.

$$\mathbf{x}^{\tau} = [x_1^{\tau}, \ldots, x_n^{\tau}, v_1^{\tau}, \ldots, v_n^{\tau}, k_1^{\tau}, \ldots, k_n^{\tau}, \beta, \beta_{BW}, \gamma, n, r]^{T} \tag{2}$$

The evolution of the system over time is described by the hidden Markov process described in Equation (3) at the generic time step τ.

$$\mathbf{x}^{\tau} = g(\mathbf{x}^{\tau-1}, \mathbf{w}^{\tau-1}) \tag{3}$$

The function $g(\cdot)$ in Equation (3) is a generic function, while $\mathbf{w}^{\tau-1}$ is the process noise vector at the time step $\tau - 1$. Noise is considered to be Gaussian according to the multivariate distribution $N(0, \Sigma_w)$, where $\Sigma_w = diag(\sigma)$. σ is the process noise vector, whose elements σ are the process noise of the variables contained in the state-space vector $\mathbf{x}$.

Considering the MDOF affected by the degraded hysteretic behaviour, the position of each degree of freedom can be described in time as in Equation (4), where $w^{\tau}_{position}$ is the process noise added to a given position.

$$x_i^{\tau+1} = v_i^{\tau} dt + x_i^{\tau} + w^{\tau}_{position} \qquad \forall i = 1, \ldots, n \tag{4}$$

Thus, the state-space equation governing the evolution of the velocity of each degree of freedom shown in Equation (5) for a generic time step τ can be derived where F_i represents the generic force acting on the *i-th* mass m_i, k_i is the stiffness of the *i-th* spring, c_i is the damping of the *i-th* damper, r is the hysteretic displacement of the first spring, $\delta_{i,1}$ is the Kronecker delta used to represent the hysteretic behaviour of the first degree of freedom, and w_{speed} is the speed process noise. For the first mass, the restoring force of the first spring is $k_1 r$ instead of $k_1 x_1$, i.e., the first spring is the hysteretic element. Equation (5) is the explicit (forward) Euler discretisation of the equations of motion with time step dt.

$$\begin{aligned} v_i^{\tau+1} = \frac{F_i^{\tau}}{m_i} dt + v_i^{\tau}\left(1 - \frac{dt}{m_i}(c_i^{\tau} + c_{i+1}^{\tau})\right) + v_{i-1}^{\tau}\frac{dt}{m_i}c_i^{\tau} + v_{i+1}^{\tau}\frac{dt}{m_i}c_{i+1}^{\tau} + \\ -\frac{dt}{m_i}\left(\left[(1-\delta_{i,1})\, x_i^{\tau} + \delta_{i,1}\, r^{\tau}\right]k_i^{\tau} + x_i^{\tau}k_{i+1}^{\tau}\right) + \\ +x_{i-1}^{\tau}\frac{dt}{m_i}k_i^{\tau} + x_{i+1}^{\tau}\frac{dt}{m_i}k_{i+1}^{\tau} + w^{\tau}_{speed} \qquad \forall i = 1, \ldots, n \end{aligned} \tag{5}$$

Equation (5) must be adjusted for the first and last degrees of freedom, as it is necessary to cancel out

- The terms reporting the subscript $i - 1$, when $i = 1$.
- The terms reporting the subscript $i + 1$, when $i = n$.

The behaviour of the Bouc–Wen hysteretic displacement r is governed by Equation (6), where w_r^{τ} represents the hysteretic behaviour process noise

$$r^{\tau+1} = r^{\tau} + dt\left(v_1 - \beta_{BW}|v_1||r^{\tau}|^{n-1}r^{\tau} - \gamma(v_1)|r^{\tau}|^{n}\right) + w_r^{\tau} \tag{6}$$

Equation (6) is valid only for MDOF systems with a degrading behaviour on the first degree of freedom.

For a healthy system, variables such as stiffness or damping are assumed to be invariant over time, and the effect of varying operating conditions, e.g., temperature changes, is modelled as an augmented state at a given time, to which the process noise is added to represent the effects of unmodelled dynamics. Equations from (7) to (12) are used to describe the governing evolution over time of the $i - th$ stiffness k_i, $i - th$ damping c_i, proportionality coefficient β, and Bouc–Wen parameters β_{BW}, γ, and n, respectively, with their respective process noise w at the generic time step τ.

$$k_i^{\tau+1} = k_i^{\tau} + w^{\tau}_{spring} \qquad \forall i = 1, \ldots, n \tag{7}$$

$$c_i^{\tau+1} = c_i^{\tau} + w_{damper}^{\tau} \qquad \forall i = 1, \ldots, n \tag{8}$$

$$\beta_i^{\tau+1} = \beta_i^{\tau} + w_{\beta}^{\tau} \tag{9}$$

$$\beta_{BW,i}^{\tau+1} = \beta_{BW,i}^{\tau} + w_{\beta_{BW}}^{\tau} \tag{10}$$

$$\gamma_i^{\tau+1} = \gamma_i^{\tau} + w_{\gamma}^{\tau} \tag{11}$$

$$n_i^{\tau+1} = n_i^{\tau} + w_n^{\tau} \tag{12}$$

Additionally, in this work the stiffness is assumed to be the only parameter that varies with temperature T according to Equation (13), obtained by rescaling the relationship reported in ref. [61], where T is expressed in °C and $k_{0,i} = 9$ N/m is the reference stiffness. Over the operational range of the main case study (40 to 10 °C) the law produces a monotonic stiffness increase of about 4%, from 8.50 to 8.85 N/m.

$$k_i(T) = \frac{0.1T^2 - 18T + 10{,}000}{10{,}000} k_{0,i} \qquad \forall i = 1, \ldots, n \tag{13}$$

The equations above describe the MDOF system dynamics, modelled as hidden Markov states. In this work, the states are estimated using a PF based on the measurements $\mathbf{y}$. Assuming that only acceleration measurements are available, the measurements $\mathbf{y}$ can be defined for the *i-th* degree of freedom in Equation (14), where ω_i^{τ} represents the acceleration measurement noise at the generic time step τ, so that $\Sigma_{\omega} = diag(\sigma_{y_1}, \ldots, \sigma_{y_n})$. As done for Equation (5), also Equation (14) must be modified when $i = 1$ and $i = n$ by removing the terms with subscript $i - 1$ and $i + 1$, respectively. The external forces F_i^{τ} are assumed to be known exactly, and the measurement noise is zero-mean Gaussian with variance R on every channel.

$$\begin{aligned} y_i^{\tau} = {} & \frac{F_i^{\tau}}{m_i} - \frac{v_i^{\tau}}{m_i}\left(c_i^{\tau} + c_{i+1}^{\tau}\right) + \frac{v_{i-1}^{\tau}}{m_i} c_i^{\tau} + \frac{v_{i+1}^{\tau}}{m_i} c_{i+1}^{\tau} + \\ & - \frac{1}{m_i}\left(\left[\left(1 - \delta_{i,1}\right) x_i^{\tau} + \delta_{i,1}\, r^{\tau}\right] k_i^{\tau} + x_i^{\tau} k_{i+1}^{\tau}\right) + \frac{x_{i-1}^{\tau}}{m_i} k_i^{\tau} + \\ & + \frac{x_{i+1}^{\tau}}{m_i} k_{i+1}^{\tau} + \omega_i^{\tau} \qquad \forall i = 1, \ldots, n \end{aligned} \tag{14}$$

2.2. Bayesian Filtering of the Augmented State

The formulation of a PF first requires defining a state-space model describing the evolution of the system under investigation. Let $\mathbf{x}_k$ denote the system state, $\mathbf{y}_k$ the available observations, and $\boldsymbol{\theta}_k$ the parameters governing the state evolution. The index k identifies the discrete time instant. The evolution of the state, commonly referred to as the process equation, is expressed in Equation (15). Without loss of generality, the function governing the process may be derived either from a physics-based model or from an empirical relationship.

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \boldsymbol{\theta}_k) \tag{15}$$

The process parameters $\boldsymbol{\theta}_k$ depend on the characteristics of the structure or system being monitored and, in general, cannot be deterministically specified for every application. For this reason, they are frequently treated as uncertain variables associated with suitable probability distributions. Their distributions are progressively updated together with the state $\mathbf{x}_k$, allowing the model to adapt online and provide more reliable estimates [21,25,26,28]. Accordingly, the augmented state vector $\mathbf{z}_k$, which contains both the system state and the model parameters to be estimated, can be written as in Equa-

tion (16), where $\boldsymbol{\omega}_k$ and $\boldsymbol{\omega}_{\theta,k}$ represent the process noise associated with the state $\mathbf{x}_k$ and the parameters $\boldsymbol{\theta}_k$, respectively.

$$\mathbf{z}_k = \left[\boldsymbol{\theta}_k, \boldsymbol{x}_k\right] = \left[\boldsymbol{\theta}_{k-1} + \boldsymbol{\omega}_{\theta,k}\, f(\mathbf{x}_{k-1}, \boldsymbol{\theta}_k, \boldsymbol{\omega}_k)\right] \tag{16}$$

In many SHM applications, the state of interest cannot be measured directly [21–26]. Thus, the available observations $\mathbf{y}_k$ are actually indirect observations. The relationship between the system state and the corresponding observation is described through the measurement function $h(\cdot)$, as reported in Equation (17), where $\boldsymbol{\nu}_k$ denotes the measurement noise.

$$\mathbf{y}_k = h(\mathbf{x}_k) + \boldsymbol{\nu}_k \tag{17}$$

The conventional state-space formulation adopted for PF-based state estimation and prognosis is summarised in Equation (18).

$$\mathbf{z}_k = \left[\boldsymbol{\theta}_k, \boldsymbol{x}_k\right] = \left[\boldsymbol{\theta}_{k-1} + \boldsymbol{\omega}_{\theta,k}, f(\mathbf{x}_{k-1}, \boldsymbol{\theta}_k, \boldsymbol{\omega}_k)\right] \qquad y_k = h(\mathbf{x}_k) + \nu_k \tag{18}$$

In Equation (18), $\boldsymbol{\omega}_k$ and $\boldsymbol{\omega}_{\theta,k}$ account for uncertainty in the state-transition model and in the evolution of the model parameters. These stochastic terms may be assumed to follow known probability distributions, namely $\boldsymbol{\nu}_k \sim p_{\boldsymbol{\nu}_k}(\cdot)$, $\boldsymbol{\omega}_k \sim p_{\boldsymbol{\omega}_k}(\cdot)$, and $\boldsymbol{\omega}_{\theta,k} \sim p_{\boldsymbol{\omega}_{\theta,k}}(\cdot)$. A frequently adopted choice consists of modelling each contribution through an independent zero-mean Gaussian distribution.

The PF algorithm performs Bayesian inference through Monte Carlo sampling. The posterior probability distribution of each unknown quantity is approximated by combining prior information with the agreement between predicted and observed measurements, as quantified through a likelihood function. The likelihood model is selected according to the assumed characteristics of the measurement uncertainty. In the present work, a Gaussian likelihood is adopted under the assumption that the measurement system is affected by unbiased uncertainty. Regardless of its specific mathematical form, the likelihood measures the consistency between the observation collected at time k and the measurement predicted by each candidate state.

Once the state-space representation has been defined, the filtering problem consists of estimating the conditional probability density function of the state vector given the available observations. Within the Bayesian framework, the prediction distribution of $\mathbf{z}_k$ is obtained using the measurements collected up to time $k-1$, denoted by $\mathbf{y}_{1:k-1}$. The subsequent assimilation of the observation $\mathbf{y}_k$ produces the posterior distribution at time k. These two stages are expressed in Equations (19) and (20), where $p(\mathbf{z}_k \mid \mathbf{z}_{k-1})$ denotes the state-transition distribution, while $p(\mathbf{y}_k \mid \mathbf{z}_k)$ is the likelihood function.

$$p(\mathbf{z}_k \mid \mathbf{y}_{1:k-1}) = \int p(\mathbf{z}_k \mid \mathbf{z}_{k-1}) p(\mathbf{z}_{k-1} \mid \mathbf{y}_{1:k-1}) d\mathbf{z}_{k-1} \tag{19}$$

$$p(\mathbf{z}_k \mid \mathbf{y}_{1:k}) \propto p(\mathbf{y}_k \mid \mathbf{z}_k), p(\mathbf{z}_k \mid \mathbf{y}_{1:k-1}) \tag{20}$$

The terms $p(\mathbf{z}_k \mid \mathbf{y}_{1:k-1})$ and $p(\mathbf{z}_k \mid \mathbf{y}_{1:k})$ represent the prior and posterior probability density functions of the augmented state vector, respectively.

All the filters compared in this work approximate the recursion of Equations (19) and (20) for the same augmented state (Equation (2)), process model, process-noise covariance $\mathbf{Q}$ and measurement-noise variance R, and they are run on the same data. They differ in how the predictive and posterior distributions are represented: by weighted particles (Section 2.3), by a Gaussian approximation (EKF and UKF), or by an ensemble corrected through a Kalman gain (EnKF; Section 2.4).

2.3. Particle Filters

2.3.1. Sequential Importance Resampling Particle Filter

Equations (19) and (20) define the optimal Bayesian filtering solution. In practical applications, however, their analytical evaluation is generally infeasible because the functions $f(\cdot)$ and $h(\cdot)$ appearing in Equation (18) may be non-linear and the associated probability distributions may be non-Gaussian. Particle filtering addresses this limitation by approximating the Bayesian distributions through a finite ensemble of weighted samples. The implementation considered here is the sequential importance resampling particle filter, commonly referred to as the SIR-PF [27]. The SIR-PF procedure comprises four main stages: (i) initialisation, (ii) prediction, (iii) weight updating, and (iv) resampling. During initialisation, N samples, or particles, denoted by $\mathbf{z}_0^i : i = 1, 2, \ldots, N$, are generated from the initial probability distribution $p(\mathbf{z}_0)$. For each subsequent time instant $k > 0$, the prediction, updating, and resampling operations are repeated sequentially. In the prediction stage, each particle is propagated according to the transition distribution $p(\mathbf{z}_k \mid \mathbf{z}_{k-1})$, resulting in the predicted ensemble $\mathbf{z}_k^i : i = 1, 2, \ldots, N$. The consistency of each predicted particle with the current observation is then quantified through its weight w_k^i. In the update stage, this weight is proportional to the likelihood associated with the corresponding particle, as stated in Equation (21).

$$w_k^i \propto p(y_k \mid \mathbf{z}_k^i) \qquad \forall i = 1, 2, \ldots, N \tag{21}$$

Equation (17) implies that the observation acquired at time k depends only on the state at the same time instant. Moreover, according to Equation (15), the current state is conditionally dependent only on the state at time $k-1$. This approach allows the importance-weight update to be expressed without explicitly considering the complete state history. The general importance-sampling formulation of the particle weight is reported in Equation (22).

$$w_k^i \propto \frac{p(y_k \mid \mathbf{z}_k^i)\, p(z_k^i \mid z_{k-1}^i)}{q(z_k^i \mid z_{k-1}^i, y_k)} \qquad \forall i = 1, 2, \ldots, N \tag{22}$$

In Equation (22), $q(z_k^i \mid z_{k-1}^i, y_k)$ is the importance density from which particles are sampled. Although different importance densities may be selected, a standard choice is to adopt the transition distribution itself [22]. Under this assumption, the transition and proposal terms cancel out, and the recursive weight update reduces to Equation (23).

$$w_k^i = w_{k-1}^i\, p(y_k \mid z_k^i) \tag{23}$$

The resulting weights are subsequently normalised so that their sum is equal to one. The normalised weight $\tilde{w}_k^i$ associated with the i^{th} particle is calculated according to Equation (24).

$$\tilde{w}_k^i = \frac{w_k^i}{\sum_{i=1}^{N} w_k^i} \tag{24}$$

The normalised weights provide a discrete approximation of the posterior distribution and quantify the relative contribution of each particle to the state estimate. In the final stage, a new particle ensemble is generated by resampling $\mathbf{z}_k^i : i = 1, 2, \ldots, N$ according to the probabilities defined by $\tilde{w}_k^i : i = 1, 2, \ldots, N$. Particles characterised by higher weights are therefore more likely to be replicated, whereas particles with negligible weights tend to be discarded. After resampling, all particle weights are reset to the uniform value $w_k^i = 1/N$.

Resampling Schemes

Two resampling schemes are compared. First, multinomial resampling is performed at every time step, so that $w_{k-1}^i = 1/N$ in Equation (23) and the weights reduce to the current likelihood, as reported in Algorithm 1. In the second scheme, the weights are accumulated recursively according to Equation (23), and resampling is performed only when the effective sample size (ESS), defined in Equation (25), drops below $N/2$. In this case, systematic resampling [62] is used, as it has a lower resampling variance than the multinomial scheme [46].

$$ESS_k = \frac{1}{\sum_{i=1}^{N} \left(\tilde{w}_k^i\right)^2} \tag{25}$$

Between two resampling events, no particle is discarded, which preserves the diversity of the parameter values, and the likelihood of several consecutive observations is accumulated before the particles are selected. With this scheme, the state estimates and the uncertainty bands are computed as weighted means and weighted quantiles of the particles. The two variants are denoted SIR PF (multinomial resampling at every step) and SIR-ESS PF (adaptive systematic resampling) in the following.

Algorithm 1 SIR PF and A-SIR PF as implemented in this work (one filtering run).

Require: measurements $\mathbf{y}_{0:N_t-1}$, inputs $\mathbf{F}_{0:N_t-1}$, number of particles N, process-noise covariance $\mathbf{Q}$, measurement-noise variance R, set of measured DOFs $\mathcal{M}$ ($n_m = |\mathcal{M}|$), coefficient m, floor ϵ, flag ACTION

Initialisation ($k = 0$): $x_0^i, v_0^i \sim \mathcal{N}(0, P_0)$; $\boldsymbol{\theta}_0^i = \boldsymbol{\alpha} \odot \boldsymbol{\theta}_{\text{ref}}$ with $\alpha_j \sim \mathcal{U}(0.95, 1.05)$ drawn once per parameter

for $k = 0, \ldots, N_t - 1$ **do**

  if $k > 0$ **then** ▷ prediction (Equations (4)–(12))

    $\mathbf{z}_k^i = f(\mathbf{z}_{k-1}^i, \mathbf{F}_{k-1}) + \boldsymbol{\omega}_k^i, \quad \boldsymbol{\omega}_k^i \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \quad i = 1, \ldots, N$

  end if

  $\hat{\mathbf{y}}_k^i = h(\mathbf{z}_k^i, \mathbf{F}_k)$ ▷ predicted accelerations (Equation (14))

  $\ell_k^i = \sum_{j \in \mathcal{M}} \log \mathcal{N}\left(y_{k,j}; \hat{y}_{k,j}^i, R\right)$ ▷ log-likelihood

  $e_k^i = \frac{1}{n_m} \sum_{j \in \mathcal{M}} \left(y_{k,j} - \hat{y}_{k,j}^i\right)$ ▷ scalar residual in measurement space [m/s^2]

  if $k = 0$ **then**

    $S_0^i = |e_0^i|$

  else

    $T_k^i = \frac{1}{2}\, m \left(\frac{e_k^i - e_{k-1}^i}{\Delta t}\right)^2, \qquad V_k^i = m\,|e_k^i|$ ▷ Equations (29) and (30)

    $S_k^i = \left|S_{k-1}^i + T_k^i - V_k^i\right|$ ▷ Equation (27)

  end if

  $\log w_k^i = \ell_k^i - \text{ACTION} \cdot \log \max\left(S_k^i, \epsilon\right)$ ▷ Equation (32) with $w_{k-1}^i = 1/N$

  $\tilde{w}_k^i = \exp(\log w_k^i - \max_j \log w_k^j) \,/\, \sum_l \exp(\log w_k^l - \max_j \log w_k^j)$

  draw ancestor indices $a_i \sim \text{Categorical}(\tilde{w}_k^1, \ldots, \tilde{w}_k^N)$ ▷ multinomial resampling at every step

  $\mathbf{z}_k^i \leftarrow \mathbf{z}_k^{a_i}, \quad S_k^i \leftarrow S_k^{a_i}, \quad e_k^i \leftarrow e_k^{a_i}$ ▷ action and last residual inherited from the ancestor

  $\hat{\mathbf{z}}_k = \frac{1}{N} \sum_i \mathbf{z}_k^i$; empirical 2.5% and 97.5% quantiles of the stiffness components

end for

2.3.2. Action-Aided Particle Filter

The A-SIR PF, proposed in ref. [50], extends the conventional SIR-PF by introducing an additional trajectory-dependent contribution into the particle-weighting procedure.

In the A-SIR PF, the standard Bayesian update is retained but complemented by an action-based term inspired by the Lagrangian formulation used in mechanics. Rather than

evaluating particles exclusively on the basis of their instantaneous agreement with the current observation, the action-aided formulation also considers the temporal behaviour of their trajectories.

In classical mechanics, the action S is a scalar functional describing the evolution of a system over a time interval. It is defined as the time integral of the Lagrangian L, as shown in Equation (26), which is conventionally expressed as the difference between the kinetic energy T and the potential energy V.

$$S[x(t)] = \int_{t_1}^{t_2} L(x, \dot{x}, t), dt = \int_{t_1}^{t_2} (T - V), dt \tag{26}$$

The action has the physical dimensions of energy multiplied by time, expressed in Js. Hamilton's principle, also known as the principle of stationary action, states that the trajectory followed by a physical system between two configurations makes the action stationary and, in many common cases, locally minimal. A related interpretation can be found in optics. According to Fermat's principle, light follows a path for which the travel time is stationary. In the path-integral interpretation, all admissible paths contribute through a phase factor of the form $e^{iS/\hbar}$. Contributions associated with paths far from the stationary-action trajectory tend to cancel because of destructive interference, whereas neighbouring paths combine more coherently. In the classical limit, the dominant contribution therefore corresponds to the stationary-action trajectory.

An analogous rationale is adopted for the action-aided PF formulation. Particle trajectories that remain consistent with the observed system evolution should contribute more strongly to the state estimate, while trajectories exhibiting persistent deviations or irregular temporal behaviour should be progressively penalised. The optical analogy is therefore used only as a conceptual motivation: paths that are far from the stationary trajectory contribute weakly to the final physical response, while particles that are closely aligned with the observed evolution should receive larger weights. On this basis, an action-dependent regularisation term is introduced to favour trajectories that are simultaneously compatible with the measurements and temporally coherent. The resulting mechanism penalises erratic particle paths, particularly when they diverge from the acquired observations, and establishes a continuous balance between stochastic exploration and consistency with the observed dynamics. The action contribution is not intended to replace the Bayesian likelihood: it sustains the conventional weighting process by introducing information associated with the complete particle trajectory. For each particle, the action-aided PF evaluates an action score S_k^i describing the cumulative evolution of that particle relative to the observations. As specified in Equation (27), the action score is recursively updated by adding the current Lagrangian contribution L_k^i to the action accumulated up to the previous time step.

$$S_k^i = |S_{k-1}^i + L_k^i| \qquad \forall i = 1, 2, \ldots, N \tag{27}$$

The absolute value in Equation (27) guarantees that the quantity subsequently introduced into the weight formulation remains non-negative. Although the signed Lagrangian contributions may vary during particle propagation, the corresponding action score is therefore constrained to positive values. The particle Lagrangian L_k^i is constructed by analogy with its classical-mechanics counterpart. As reported in Equation (28), it is expressed as the difference between two terms, T_k^i and V_k^i, which play roles analogous to kinetic and potential energy, respectively.

$$L_k^i = T_k^i - V_k^i \qquad \forall i = 1, 2, \ldots, N \tag{28}$$

The potential-like term V_k^i measures the discrepancy between the particle prediction and the observation. It is defined through the magnitude of the particle error e_k^i, as shown in

Equation (29), where m is a coefficient representing the mass in the classical formulation of potential and kinetic energy. Since the state $\mathbf{z}_k^i$ and the observation $\mathbf{y}_k$ do not live in the same space, the error is defined in the measurement space, through the measurement function of Equation (17), as in the signed mean of the acceleration residuals over the set $\mathcal{M}$ of the n_m measured degrees of freedom:

$$e_k^i = \frac{1}{n_m} \sum_{j \in \mathcal{M}} \left(y_{k,j} - h_j(\mathbf{z}_k^i) \right), \qquad V_k^i = m|e_k^i| \qquad \forall i = 1, 2, \ldots, N \tag{29}$$

so that e_k^i is a scalar expressed in $\mathrm{m/s^2}$. The kinetic-like contribution T_k^i instead characterises the temporal variation of the particle error. It is defined as half of the squared finite-difference derivative of the error, as given in Equation (30), where Δt denotes the time interval between two consecutive observations. Equation (30) resembles the classical formulation of the kinetic energy, with the variation of the particle error over the time steps used as the velocity; since the error is considered as the displacement in the potential-like term, the derivative of the error over time can be seen as the velocity in the kinetic-like term. The idea is to promote temporal coherence of the particles by also considering how fast the particle fluctuates to adapt to observations. Thus, particles with smoother trajectories are associated with a lower kinetic energy.

$$T_k^i = \frac{1}{2} m \left(\frac{e_k^i - e_{k-1}^i}{\Delta t} \right)^2 \qquad \forall i = 1, 2, \ldots, N \tag{30}$$

Through these definitions, the action score accounts for both the instantaneous particle-to-observation discrepancy and the evolution of that discrepancy over time. Four remarks on these definitions are worth making. First, T_k^i and V_k^i do not have the same physical units, i.e., $\mathrm{m^2/s^6}$ and $\mathrm{m/s^2}$, respectively. Thus, the Lagrangian of Equation (28) is an analogy rather than a dimensionally homogeneous energy difference. Second, since the same coefficient m multiplies both terms, the recursion of Equation (27) can be written as $S_k^i = m\tilde{S}_k^i$ with $\tilde{S}_k^i = |\tilde{S}_{k-1}^i + T_k^i/m - V_k^i/m|$. Thus, the factor m cancels when the weights are normalised (Equation (24)) and only enters through the initial value $S_0^i = |e_0^i|$, whose influence vanishes after the first step. In other words, m has no practical effect on the filter, as verified numerically in Section 4.10, and a tunable trade-off between the two terms would require two different coefficients or an exponent applied to S_k^i. Third, the residual can be written as the sum of the channel-averaged measurement noise $\bar{v}_k$, which is white with variance R/n_m and common to all particles, and of a particle-specific model error δ_k^i. Thus, the expected value of the noise-driven part of the kinetic-like term is $mR/(n_m \Delta t^2)$, i.e., it grows as Δt^{-2}, whereas for a smooth model error the term $\frac{m}{2}(\Delta\delta_k^i/\Delta t)^2$ tends to a finite limit. At the sampling interval used in this work, T_k^i exceeds V_k^i by about four orders of magnitude, and it is dominated by the increments of the measurement noise, which are common to all particles and cancel in the normalisation. Hence, only the particle-specific part of S_k^i affects the weights, and its relative influence decreases as S_k^i accumulates over time. Finally, S_k^i is accumulated as a sum over time steps rather than as a time integral, and a factor Δt would again cancel in the normalisation. Particles associated with persistent (high potential-like terms) or rapidly varying errors (high kinetic-like term) consequently tend to accumulate less favourable action values. The action score is then incorporated into the importance-weight update. The general action-aided weighting expression is reported in Equation (31).

$$w_{S,k}^i \propto \frac{1}{S_k^i} \frac{p(y_k \mid \mathbf{z}_k^i) p(z_k^i \mid z_{k-1}^i)}{q(z_k^i \mid z_{k-1}^i, y_k)} \qquad \forall i = 1, 2, \ldots, N \tag{31}$$

When the transition density is selected as the importance density, the formulation simplifies to Equation (32).

$$w_{S,k}^{i} = w_{S,k-1}^{i} \frac{p\left(y_k \mid z_k^i\right)}{S_k^i} \tag{32}$$

The subscript S identifies the action-aided particle weight. Equation (31) therefore replaces the conventional importance-weight expression in Equation (22). The inverse dependence on S_k^i reflects the least-action rationale adopted in the action-aided formulation: particle trajectories associated with larger action values are penalised through smaller weights, while trajectories with smaller action values are favoured. The use of the absolute value in Equation (27) is consistent with this objective. The weighting mechanism is designed to reward action values close to zero, independently of whether the cumulative signed Lagrangian would otherwise be positive or negative.

A distinguishing feature of the action-aided approach is that the weight assigned at time k depends not only on the latest observation, but also on the previous evolution of the particle. Since the action is recursively accumulated, as indicated in Equation (27), information from the particle and observation histories is retained in the weighting process. This differs from the standard formulations in Equations (22) and (23), in which the recursive update is Markovian and the current likelihood is evaluated primarily from the state and observation available at the current time instant [20–27].

Weight normalisation, resampling, and future-state propagation can be performed according to the procedures used in the conventional SIR-PF. The action-aided weights alter the probability with which particles contribute to the posterior distribution and are selected during resampling. At resampling, each new particle inherits, together with the state, the accumulated action $S_k^{a_i}$ and the last residual $e_k^{a_i}$ of its ancestor a_i; no reconstruction from the observation history is performed. This operation ensures that the subsequent weighting stage continues to account for the trajectory information accumulated before resampling. The complete procedure, as implemented, is summarised in Algorithm 1.

With adaptive resampling (A-SIR-ESS PF), the weights are accumulated according to Equation (32), the action is updated at every step for all particles, and the ancestor draw of Algorithm 1 is performed only when the ESS of the A-SIR weights drops below $N/2$, with systematic instead of multinomial sampling.

From a computational point of view, the introduction of the action term does not significantly increase the overall cost of the filtering process, since the SIR PF and the A-SIR PF have the same dominant asymptotic complexity. Let N_p be the number of particles and N_t the number of time steps. For each filtering step, the computational cost of the standard PF is mainly associated with the propagation of the particles through the non-linear state-space model, the evaluation of the likelihood from the measured dynamic response, the normalisation of the particle weights, and the possible resampling step. Therefore, the cost per time step can be written as in Equation (33), where C_{dyn} denotes the cost of propagating one particle through the non-linear state-space model and C_ℓ is the cost of evaluating the likelihood for one particle.

$$\mathcal{O}\Big(N_p(C_{\text{dyn}} + C_\ell + 1)\Big) \tag{33}$$

Since the system dimension is fixed, both C_{dyn} and C_ℓ can be regarded as constants with respect to N_p, leading to a per-step complexity of $\mathcal{O}(N_p)$ and to an overall filtering cost of $\mathcal{O}(N_t N_p)$. The A-SIR PF introduces an additional action-based weighting term, computed from the particle error and its temporal variation. If the cumulative action and the previous error are stored for each particle, as done in this study, this operation requires only a constant number of additional operations per particle. Consequently, the A-SIR PF also scales as $\mathcal{O}(N_p)$ per time step and as $\mathcal{O}(N_t N_p)$ over the full simulation, with only a

slightly larger multiplicative constant than the standard PF. In terms of memory, both filters require $\mathcal{O}(N_p d)$ storage for the particles and weights, where d is the dimension of the state vector, while the A-SIR PF requires an additional $\mathcal{O}(N_p)$ storage term for the cumulative action and previous error. Thus, the introduction of the action term does not change the computational complexity class of the standard SIR PF; it only introduces a low-constant linear overhead. The inheritance of S_k^i and e_k^i at resampling adds the copy of two vectors of length N_p. The measured computational times are reported in Section 4.11.

2.3.3. Reformulations of the Action-Aided Weighting

Two reformulations of the action-aided idea are also considered. In the first one (A-SIR-FK PF), the residual is normalised by the standard deviation of the channel-averaged measurement noise, $\tilde{e}_k^i = e_k^i / \sqrt{R/n_m}$, and the path penalty is applied incrementally in exponential form,

$$w_k^i \propto p(\mathbf{y}_k \mid \mathbf{z}_k^i)\, \exp\!\big(-\lambda P_k^i\big), \qquad P_k^i = \tfrac{1}{2}\big(\tilde{e}_k^i - \tilde{e}_{k-1}^i\big)^2 + \kappa\, |\tilde{e}_k^i|, \tag{34}$$

so that, with resampling at every step, the lineage weights correspond to the path measure tilted by $\exp(-\lambda \sum_j P_j^i)$, i.e., to a Feynman–Kac formulation [63]. Compared with Equation (32), this formulation has four properties: (i) common-mode contributions cancel exactly, (ii) the influence of the past does not vanish as the action accumulates, (iii) the strength λ does not cancel in the normalisation, and (iv) both persistent ($\kappa = 1$) and rapidly varying errors are penalised.

The second reformulation starts from the observation that both the A-SIR and the incremental weightings re-use the measurements already contained in the likelihood. In the Euclidean (imaginary-time) form of the path integral, the weight of a trajectory is $\exp(-S_E)$, with

$$S_E[\mathbf{z}_{0:k}] = \sum_{j=1}^{k} \Big[-\log p(\mathbf{z}_j \mid \mathbf{z}_{j-1}) - \log p(\mathbf{y}_j \mid \mathbf{z}_j) \Big] + \text{const}, \tag{35}$$

which is exactly the negative logarithm of the posterior density of the path. In this form the kinetic-like term is the transition density (the prior on the trajectory), the potential-like term is the likelihood, and both are dimensionless. The SIR PF already samples paths according to $\exp(-S_E)$. Therefore, a formulation of the least-action idea that is consistent with Bayesian filtering must act on the transition model of the parameters. In the action-consistent SIR PF (AC-SIR PF), each stiffness is augmented with a degradation rate u_i evolving as an Ornstein-Uhlenbeck process:

$$k_i^{\tau+1} = k_i^\tau + \Delta t\, u_i^\tau + w_k^\tau, \qquad u_i^{\tau+1} = \rho\, u_i^\tau + w_u^\tau, \qquad \rho = e^{-\Delta t/\tau_u}, \tag{36}$$

with $w_k \sim \mathcal{N}(0, (1-\varphi)Q_k)$ and $w_u \sim \mathcal{N}(0, \sigma_u^2(1-\rho^2))$, $2\sigma_u^2\tau_u = \varphi Q_k/\Delta t$. With this choice the long-term diffusion of the stiffnesses equals that of the SIR PF, whereas at time scales shorter than τ_u the stiffness paths are smoother, i.e., of lower kinetic-like action. An optional jump component (probability p_J per step of a jump $\mathcal{N}(0, 0.8^2)$ N/m on one randomly chosen stiffness) accommodates abrupt damage. The weights are the plain likelihood, and the filter reduces to the SIR PF for $\varphi = 0$, $p_J = 0$. The tunable parameters of both reformulations are selected on independent tuning simulations (Section 2.8).

2.3.4. Other Particle-Filter Variants

Two established remedies against weight degeneracy and loss of parameter diversity are included as references. In the tempered SIR PF, the likelihood is raised to the power $\alpha_T = 0.5$, which is equivalent to doubling the measurement-noise variance in the weights and flattening them [49]. In the Liu–West SIR PF [47], after each resampling the parameters

$\boldsymbol{\theta}$ (stiffnesses, damping coefficient and Bouc–Wen parameters) are moved with a kernel-shrinkage step,

$$\boldsymbol{\theta}^i \leftarrow a\,\boldsymbol{\theta}^i + (1-a)\,\bar{\boldsymbol{\theta}} + h\,\boldsymbol{\varepsilon}^i, \qquad \boldsymbol{\varepsilon}^i \sim \mathcal{N}\big(\mathbf{0}, \mathrm{diag}(\mathbf{V})\big), \tag{37}$$

where $\bar{\boldsymbol{\theta}}$ and $\mathbf{V}$ are the mean and the variance of the particles, $a = (3\delta - 1)/(2\delta)$, $h^2 = 1 - a^2$ and $\delta = 0.99$. This step restores the diversity while preserving the mean and the variance of the particle cloud. Both variants use multinomial resampling at every step, as in the SIR PF.

2.3.5. Damage-Factor Particle Filter

The equations of motion are not the only physical knowledge available on the monitored structure. In the damage-factor particle filter (DF-PF), two further pieces of prior knowledge on the structure are used, and both enter the transition model rather than the weights. First, following the common representation of damage as a reduction in the stiffness parameters with respect to those of the undamaged structure [51,52], the stiffness of the j-th spring $k_j(t)$ is written as the product of a healthy temperature–stiffness law $k_{h,j}\big(\hat{T}(t)\big)$ and of a damage factor $d_j(t)$,

$$k_j(t) = k_{h,j}\big(\hat{T}(t)\big)\, d_j(t), \tag{38}$$

where $\hat{T}$ is the measured temperature smoothed by a causal local linear fit over the last 0.5 s. The law $k_{h,j}$ is not taken from the simulator, but it is learned offline before the monitoring, as in the data-normalisation approaches that model the environmental variability of the healthy structure from data [53]. Specifically, $k_{h,j}$ is a quadratic polynomial in $\hat{T}$ fitted to the stiffness estimates of the EKF on the independent healthy simulations, with the same five-fold cross-fitting used for the normalisation models, as reported in Section 2.6. The EKF is used for this calibration because it is among the most accurate filters and among the least expensive. The error of the learned law with respect to the exact law of Equation (13) is about 0.02 N/m (0.2%). The thermal variation is therefore part of the transition model, $k_j(t) = \big[k_{h,j}(\hat{T}(t))/k_{h,j}(\hat{T}(t-\Delta t))\big]\, k_j(t-\Delta t) + w_k$ with $w_k \sim \mathcal{N}(0, s^2 Q_k)$, and the random walk only has to describe the damage. Second, damage cannot increase the stiffness: the damage factor is constrained to $d_j \leq 1 + \delta$, where the margin δ accounts for the uncertainty of the learned law, and particles that cross the bound are reflected. The monitoring starts from the calibrated healthy state, $d_j(0) = 1$, and adaptive systematic resampling is used as in the SIR-ESS PF. A hard constraint is natural for a particle filter, which can impose it on each particle during the propagation [56–58] and thus represent the truncated posterior. For comparison, the same model is given to the EKF (DF-EKF). In this case, the constraint is imposed by projecting the updated mean onto the admissible set, which is one of the ways of including constraints in Kalman-type filters, although it is in general not optimal when the system is nonlinear [56]. The scale $s \in \{0.5, 1\}$ and the margin $\delta \in \{0, 0.02, 0.05\}$ are selected in the tuning simulations (Section 2.8). The posterior of d_j also provides directly the probability that the stiffness is below a given fraction of its healthy value, i.e., the damage measure proposed in Ref. [51]. However, for comparability, the damage-factor filters are evaluated in this work with the same detection chain as the other filters.

2.4. Kalman-Type Filters

The EKF, UKF and EnKF use the same augmented state, discrete-time process model (Equations (4)–(12)), process-noise covariance $\mathbf{Q}$ and measurement model (Equation (14)) as the particle filters, and their initial parameter estimates are drawn in the same way, i.e., with a random offset within ±5% of the true initial values. The initial covariance is diagonal, with the variance of a uniform distribution of ±5% for the parameters and P_0

for the kinematic states. At each step, given the predicted mean $\hat{\mathbf{z}}_k^-$ and covariance $\mathbf{P}_k^-$, the predicted measurement $\hat{\mathbf{y}}_k$, the cross-covariance $\mathbf{P}_{zy}$ and the innovation covariance $\mathbf{P}_{yy}$, the update is

$$\mathbf{K}_k = \mathbf{P}_{zy}\mathbf{P}_{yy}^{-1}, \qquad \hat{\mathbf{z}}_k = \hat{\mathbf{z}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k), \qquad \mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k\mathbf{P}_{yy}\mathbf{K}_k^T. \tag{39}$$

The three filters differ in how these moments are computed:

- EKF: the process and measurement functions are linearised around the current estimate, with Jacobians computed by central finite differences, and the covariance update is written in the Joseph form for numerical robustness [64];
- UKF: the moments are computed by propagating $2n + 1 = 31$ sigma points through the nonlinear functions, with the scaled unscented transform ($\alpha = 1$, $\beta = 2$, $\kappa = 0$) [39];
- EnKF: an ensemble of $N_e = 200$ members is propagated through the nonlinear model with independent samples of the process noise, the moments of Equation (39) are estimated from the ensemble, and each member is updated with an independently perturbed observation (stochastic EnKF [40,41]); no inflation or localisation is used.

Unlike the particle filters, which can only select among the propagated particles, these filters correct all the states, including positions and velocities, through the gain $\mathbf{K}_k$. However, they are exact only for linear-Gaussian models. The 95% bands of the EKF and UKF and, for consistency, of the EnKF are computed as the mean ± 1.96 standard deviations. The cost per step is dominated by $2n$ (EKF) or $2n + 1$ (UKF) evaluations of the process and measurement models and by $\mathcal{O}(n^3)$ matrix operations, with $n = 15$, and by N_e model evaluations for the EnKF. Unlike that of the particle filters (Equation (33)), this cost does not depend on a number of particles.

2.5. Least-Action Trajectory Estimation

Equation (35) suggests a third, more direct, use of the least-action idea. Since $\exp(-S_E)$ is the posterior density of the path, the trajectory of minimum Euclidean action based on the measurements is the maximum a posteriori trajectory. Thus, when formulated consistently with Bayesian inference, the least-action principle is an estimation principle for the whole trajectory, rather than a re-weighting of the particles or a prior. For the Gaussian process and measurement noise of the benchmark, the Euclidean action is a sum of quadratic forms of the process and measurement residuals,

$$S_E[\mathbf{z}_{0:k}] = \frac{1}{2}\sum_{j=1}^{k}\left[(\mathbf{z}_j - f(\mathbf{z}_{j-1}))^T\mathbf{Q}^{-1}(\mathbf{z}_j - f(\mathbf{z}_{j-1})) + \frac{\|\mathbf{y}_j - h(\mathbf{z}_j)\|^2}{R}\right] + \text{const}, \tag{40}$$

where the kinetic-like term penalises the departures of the trajectory from the dynamics and the potential-like term the misfit to the measurements, both in dimensionless form. After linearisation of f and h around the EKF estimates, the minimiser of Equation (40) is given by the Rauch-Tung-Striebel backward recursion [65,66] applied to the EKF of Section 2.4:

$$\mathbf{C}_j = \mathbf{P}_j\mathbf{F}_{j+1}^T(\mathbf{P}_{j+1}^-)^{-1}, \qquad \hat{\mathbf{z}}_j^s = \hat{\mathbf{z}}_j + \mathbf{C}_j(\hat{\mathbf{z}}_{j+1}^s - \hat{\mathbf{z}}_{j+1}^-), \qquad \mathbf{P}_j^s = \mathbf{P}_j + \mathbf{C}_j(\mathbf{P}_{j+1}^s - \mathbf{P}_{j+1}^-)\mathbf{C}_j^T, \tag{41}$$

where $\mathbf{F}_{j+1}$ is the Jacobian of the transition from j to $j + 1$ and the superscript s denotes the smoothed estimate. For online monitoring, a fixed-lag version is used: every 0.1 s the backward recursion is run over the last $L + 0.1$ s, so that the stiffness at time t is estimated with the data up to $t + L$ and becomes available with a delay L. This least-action smoother (LAS) is used with $L = 1$ s, which is short compared with the 2.78 s window of the damage indicator, while lags of $L = 0.5$ and 2 s are considered in a sensitivity study. In the detection

analysis, the lag is added to the detection delays, so that all the methods are compared at the time at which an alarm can actually be raised.

2.6. Environmental Normalisation

2.6.1. Autoencoders

AEs have been implemented using the scikit-learn (version 1.70) Python (version 3.13.2) library [67] with the architecture reported in Table 1 and hyperbolic-tangent activation function, since a ReLU neuron in the single-neuron bottleneck can be inactive over part of the input range, where the reconstruction is then constant. One AE is trained for each spring. The input layer of the AEs has two neurons: one receives the measured temperature, and the other the stiffness estimation obtained by the PF based on acceleration measurements. The output layer reconstructs the input temperature and stiffness. The AE is trained exclusively on healthy-state data, for which the relationship between temperature and stiffness is not affected by damage. After training, the reconstruction error can therefore be used as a damage-sensitive indicator: damage-induced changes in the temperature–stiffness relationship cause the input data to deviate from the patterns learned during training, resulting in larger reconstruction errors.

Table 1. Architecture of the AEs.

#	Layer	Neurons
1	Input	2
2	Mapping Layer	10
3	Bottleneck	1
4	Demapping layer	10
5	Output	2

Training was performed for 2000 epochs with the Adam optimisation using an early stopping criterion if the validation error does not decrease for 20 epochs. A penalty term of 0.001 was set for regularisation by constraining the size of the weights to avoid overfitting. The seed was fixed to ensure reproducibility between different trainings. The network parameters were optimised by minimising the mean square reconstruction error between the input and output vectors of the AE. The training data are the stiffness estimates provided by the same filter (the filter under assessment, with the same settings used in the tests) on independent simulations of the healthy structure, together with the corresponding measured temperatures, whereas the true stiffness is never used. Separate AEs are therefore trained for each filter. To avoid any overlap between training, threshold calibration and evaluation, the healthy simulations are split into five folds: for each fold used for evaluation, the AEs are trained on two other folds and the alarm thresholds, which are described in Section 2.7, are calibrated on the remaining two folds. For each run, one sample every 10 time steps is retained after the first 2.78 s (5000 steps) to account for the initial convergence of the filter, giving 50,000 training samples per spring from 20 healthy training simulations, of which 20% are used as validation set for early stopping. One AE per spring is used because each network has to learn the healthy temperature–stiffness relationship of one spring, and the comparison of the per-spring reconstruction errors provides the damage localisation. Data have been standardised before training, as described in Section 3.

2.6.2. Conditional Normalisation

As an alternative to the AE, a conditional model of the healthy relationship between temperature and estimated stiffness is considered. For each spring, a multilayer perceptron (MLP) with the measured temperature as the only input, two hidden layers of 10 neurons

with ReLU activation and the estimated stiffness as output is trained on the same healthy estimates, with the same input standardisation, regularisation, early-stopping and cross-fitting settings as the AE. The squared error of Equation (42) is then replaced by the squared residual, computed on the standardised variables, $SE_i^\tau = \left(\hat{k}_i^\tau - \hat{k}_{reg,i}(\hat{T}^\tau)\right)^2$. Unlike the AE, which learns the healthy relationship implicitly by reconstructing both of its inputs, the conditional model treats the temperature as an exogenous input and attributes any deviation of the stiffness from its healthy prediction to the structure. Two further indicators are used as references: the squared deviation of $\hat{k}_i$ from its mean over the baseline window, without any normalisation (raw), and the squared deviation from the exact thermal law of Equation (13) (oracle), which is not available in practice. Unless otherwise stated, the results refer to the AE, as in Refs. [31,50].

2.7. Diagnostic Indicators

Diagnostic indicators are used to trigger an alarm based on the discrepancies between the observed behaviour and the expected one. However, the diagnostic indicator should be informative not only about the presence of damage, but also about where the damage occurred. Thus, each spring is associated with a diagnostic indicator and, therefore, with an alarm.

The indicators are based on the reconstruction error of the AE, as this error should be significant only when associated with a damaged spring. Thus, the first damage indicator adopted is the squared error SE_i^τ between the input temperature $\hat{T}_i^\tau$ and stiffness $\hat{k}_i^\tau$ and the output temperature $\hat{T}_{rec,i}^\tau$ and stiffness $\hat{k}_{rec,i}^\tau$ of the AE for the *i-th* spring at the τ time step, as in Equation (42), where all the quantities are standardised as described in Section 3. The input stiffness is the one estimated by the filter under assessment.

$$SE_i^\tau = (\hat{T}_i^\tau - \hat{T}_{rec,i}^\tau)^2 + (\hat{k}_i^\tau - \hat{k}_{rec,i}^\tau)^2 \tag{42}$$

The alarms are given when SE_i^τ reaches a predefined threshold. However, threshold definition is arbitrary and can lead to false alarms if the threshold has not been properly defined or in case of noise. Thus, the method proposed in ref. [31] is adopted in this study: each spring is associated with a threshold defined as in Equation (43), where $ATSE_i^{99}$ is the threshold of the *i-th* spring evaluated as the $99th$ percentile of the SE of the *i-th* spring over the baseline window $[\tau_b, k_0)$, i.e., from the end of a burn-in of $\tau_b = 5000$ time steps (2.78 s), which excludes the initial convergence of the filters, to the end of the baseline at $k_0 = 15{,}000$ steps (8.33 s). The moving-average window has width $\tau_0 = 5000$ steps (2.78 s).

$$ATSE_i^{99} = p_{99}\left(SE_i^{\tau_b:k_0}\right) \tag{43}$$

Since thresholds computed in this way differ between filters, they do not guarantee that different filters operate at the same false-alarm rate. For the comparisons reported in Section 4, a calibrated threshold h_i is therefore also used: h_i is the 99th percentile of the moving average μ_{SE_i} over the monitoring period (after the baseline) of independent healthy calibration runs processed with the same filter. By construction, all the filters then have the same 1% per-sample false-alarm rate on healthy data. An alarm is triggered for the *i-th* spring when the moving average $\mu_{SE_i^{\tau-\tau_0+1,\tau}}$ of the squared errors SE_i over a sliding window of width τ_0 reaches the threshold $ATSE_i^{99}$ (or h_i). Thus, the deterministic damage indicator $Id_{SE_i}^{\tau-\tau_0+1,\tau}$ associated with this procedure is defined according to Equation (44).

$$Id_{SE_i}^{\tau-\tau_0+1,\tau} = \begin{cases} 1 & \text{if } \mu_{SE_i^{\tau-\tau_0+1,\tau}} \geq ATSE_i^{99}, \\ 0 & \text{otherwise} \end{cases} \tag{44}$$

This damage indicator is a binary indicator as it can only be associated with no damage (0) or with damage (1). The threshold $ATSE_i^{99}$ and the moving average $\mu_{SE_i^{\tau-\tau_0+1,\tau}}$ are computed on the same time window and, the wider the window, the more robust the algorithm is to false alarms due to noise or error fluctuations. However, the drawback of increasing the time window is the delayed detection of damages.

An additional damage indicator, called probabilistic damage indicator $Ip_{SE_i}^{\tau-\tau_0+1,\tau}$ is defined in Equation (45), where $Id_{SE_i}^{\tau}$ is the instantaneous equivalent of $Id_{SE_i}^{\tau-\tau_0+1,\tau}$ computed at time step τ, and defined in Equation (46).

$$Ip_{SE_i}^{\tau-\tau_0+1,\tau} = \frac{\sum_{j=\tau-\tau_0}^{\tau} Id_{SE_i}^{j}}{\tau_0} \tag{45}$$

$$Id_{SE_i}^{\tau} = \begin{cases} 1 & \text{if } SE_i^{\tau} \geq ATSE_i^{99} \\ 0 & \text{otherwise} \end{cases}, \tag{46}$$

Three detection metrics are computed for each run: (i) the detection delay, defined as the time between the damage onset and the first instant after which $Id_{SE_3} = 1$ for at least 100 consecutive steps (0.056 s), (ii) the false-alarm duration, defined as the total time during which $Id_{SE_i} = 1$ for a healthy spring during the monitoring period, and (iii) the area under the receiver operating characteristic curve (AUC), as a threshold-free measure. The AUC is obtained by comparing, for each spring, the maximum of μ_{SE_i} after the damage onset in the damaged runs with the same quantity in independent healthy runs.

2.8. Assessment Criteria

The filters are compared according to the following criteria, which separate the systematic differences between filters from the randomness of the data and of the filters themselves:

- For each simulation (seed), the process noise of the ground truth, the measurement noise and the temperature-measurement noise are generated once, and all the filters are run on exactly the same measurements. The stochastic filters use the same filter seed, so that the differences between filters can be analysed in a paired fashion. The healthy simulations are used to train the normalisation models and to calibrate the thresholds, with the five-fold cross-fitting scheme of Section 2.6, and as negative cases in the ROC analysis.
- Tracking is assessed, for each stiffness, through the RMSE and the MAE over [2.78, 16.67] s, i.e., after the initial filter convergence, and through the time-averaged width and the empirical coverage of the 95% bands. In addition, a run is considered lost if at least one stiffness has an RMSE larger than 0.5 N/m, i.e., about 6% of the nominal stiffness.
- Three threshold definitions are compared: (i) the per-run thresholds of Equation (43), (ii) a common multiplier of the calibrated thresholds h_i, chosen for each filter so that 10% of the healthy simulations raise a persistent alarm on any spring during the monitoring period (equal run-level false-alarm rate), and (iii) a robust calibration, used as a sensitivity analysis. In the robust calibration, h_i is the median plus three scaled median absolute deviations of the per-run 99th percentiles of μ_{SE_i} on the healthy calibration simulations, which makes it insensitive to a few divergent healthy runs. Localisation is quantified by the fraction of simulations in which the first persistent alarm after the onset is raised on the damaged spring, and by the fraction of the post-onset time in which the damaged spring has the largest normalised indicator μ_{SE_i}/h_i.

- Paired differences of continuous metrics are tested with the Wilcoxon signed-rank test [68], paired binary outcomes with the exact McNemar test [69], and AUC differences with a paired bootstrap over the simulations (2000 resamples). The Holm correction [70] is applied when several comparisons are reported together.
- In a control test, the SIR PF is re-run on the same data with a different filter seed, which quantifies the variability due to the randomness of the filter alone.
- The reformulated action-aided filters (Section 2.3.3) and the damage-factor filters (Section 2.3.5) were tuned on four additional simulations not used elsewhere, with selection rules fixed in advance. For the reformulated filters, the selected configuration is the one with the smallest geometric mean of the RMSE ratio with respect to the SIR PF, while for the damage-factor filters the ratio is computed with respect to the same filter without the damage-factor model, i.e., the SIR-ESS PF and the EKF, respectively. The selected configurations were then evaluated on the simulations of the main campaign.

3. Case Study

The filters are compared in a numerical case study of a three-degree-of-freedom (3-DOF) MDOF system, represented in Figure 2, which relies on the MDOF equations described in Section 2.1. The system parameters are selected according to a benchmark configuration available in the literature [31]. Each mass is set to $m_i = 1$ kg, while the reference stiffness is $k_{0,j} = 9$ N/m for $j = 1, \ldots, n+1$. The proportional damping coefficient is $\beta = 2 \times 10^{-3}$ s, i.e., $c_j = \beta k_j \approx 0.018$ N · s/m (damping ratio of the first mode of about 0.3%). The undamped natural frequencies of the system are 0.37, 0.68 and 0.88 Hz. The first degree of freedom is assumed to exhibit non-linear hysteretic behaviour, modelled through the Bouc–Wen formulation with parameters $\beta_{BW} = 2$, $\gamma = 1$, and $n = 2$. All state variables are affected by Gaussian process noise with zero mean and prescribed variances, accounting for modelling uncertainties and unmodeled dynamics. The system is excited by a harmonic force $F_1(t) = 10 \sin(2\pi f t)$ N with $f = 30$ Hz applied to the first mass, which is assumed to be known exactly by the filters. Since the excitation frequency lies far above the natural frequencies, the forced response is small, and the stiffness-dependent part of the measured accelerations (i.e., the acceleration minus the known term F_i/m_i) has an RMS value of 0.07–0.09 m/s^2, only 4–5 times the standard deviation of the measurement noise. A change of 1% in any stiffness modifies the accelerations by about 3% of the noise standard deviation per sample, so that the stiffnesses are only weakly observable from individual samples and the information is accumulated over many time steps. With this excitation level the hysteretic variable remains small ($|r| < 1.4 \times 10^{-2}$ m) and the nonlinear terms of Equation (6) are at most 5×10^{-4} times the linear term: the Bouc–Wen element therefore operates in its linear range. A broadband excitation is considered as an additional scenario, and it is described in Section 3.2. The total simulation time is $T_{\text{hor}} = 16.7$ s, discretized into $N_t = 3 \times 10^4$ time steps with sampling frequency $f_s = 1800$ Hz, ensuring compliance with the Shannon sampling criterion. Environmental effects are introduced through a time-varying temperature profile. The temperature is assumed uniform across the structure and varies linearly from 40°C to 10°C over the simulation horizon. Structural stiffness is assumed to depend on temperature, thus introducing confounding effects in the damage diagnosis task. Two damage scenarios have been analysed:

1. A progressive degradation of the stiffness of the third spring starting at time $t = 11.1$ s and evolving linearly until the end of the simulation, reaching a final value equal to 70% of the corresponding healthy stiffness at the same temperature. This scenario is called Progressive Damage.

2. A sudden degradation of the stiffness of the third spring occurring at time $t = 11.1$ s, represented by a step reduction in the stiffness down to 90% of the corresponding healthy stiffness at the same temperature. This scenario is called Step Damage.

The particle filters use 3×10^4 particles to estimate the full state vector at each time step. The initial particles are sampled such that kinematic variables are initialised around zero with small variance ($P_0 = 10^{-8}$), while structural parameters are initialised with a random offset within $\pm 5\%$ of their true initial values. The offset is drawn once per parameter and is common to all particles, so that the initial parameter spread is generated by the process noise during the first steps. The process noise variances used in the PF implementation are: position $Q_x = 1 \times 10^{-16}$, velocity $Q_v = 1 \times 10^{-16}$, stiffness $Q_k = 1.62 \times 10^{-5}$, damping: $Q_c = 8 \times 10^{-10}$, Bouc–Wen parameters: $Q_{\beta_{BW}} \approx 9.60 \times 10^{-7}, Q_\gamma \approx 2.40 \times 10^{-7}, Q_n \approx 9.60 \times 10^{-7}, Q_r = 0$. The observation noise variance is $R = 2.5 \times 10^{-4}$ $(\mathrm{m/s^2})^2$, i.e., a standard deviation of 0.0158 $\mathrm{m/s^2}$ on each of the three measured accelerations. The resampling scheme and the specific settings of each filter are given in Section 3.1.

For the A-SIR PF, the Lagrangian is evaluated using $m = 1.5$ and $\epsilon = 10^{-12}$.

The temperature measured by the monitoring system is affected by Gaussian noise with a standard deviation of about 1 °C. The AEs are trained on the stiffness estimates obtained by the filters on independent simulations of the healthy (undamaged) system with the same temperature profile, as described in Section 2.6. Thus, the tests concern interpolation within the training temperature range. Before AEs training, data arestandardised according to Equations (47) and (48) for the input temperature T_{input} and stiffness k_{input}, respectively, where $\bar{T}$ and σ_T are the mean and the standard deviation of the measured temperature, and $\bar{k}_i$ and σ_{k_i} are those of the stiffness of the i-th spring estimated by the filter, all computed on the healthy training simulations of each fold. Standardisation gives the two inputs the same weight in the reconstruction error. In contrast, with fixed scaling constants, the scaled stiffness varies over a narrow interval, and the AE with a single-neuron bottleneck can converge to a constant output equal to the mean of the training data, so that no environmental normalisation is performed. The same standardisation is used for the conditional model of Section 2.6.

$$T_{input} = \frac{\hat{T}_i^\tau - \bar{T}}{\sigma_T} \tag{47}$$

$$k_{input} = \frac{\hat{k}_i^\tau - \bar{k}_i}{\sigma_{k_i}} \tag{48}$$

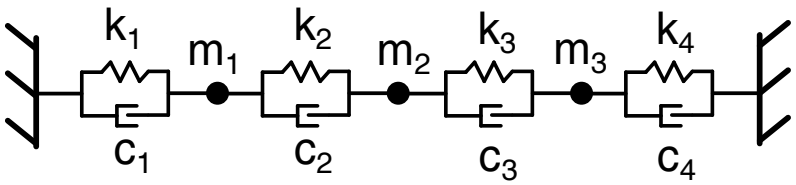


Figure 2. Schematic of the 3-DOF system implemented in the case study.

3.1. Filter Settings

Table 2 lists the filters compared in this work and their specific settings. All of them use the model, the process-noise variances, the measurement-noise variance and the initialisation given above, and they are run on the same data. The parameters of the A-SIR-FK PF, of the AC-SIR PF and of the damage-factor filters are those selected on the tuning simulations and presented in Sections 4.5 and 4.7.

Table 2. Filters and smoother compared in this work. All of them share the model, process-noise covariance, measurement-noise variance, initialisation and data, except the damage-factor filters (DF), which add the learned thermal law and the damage constraint of Section 2.3.5 and start from the calibrated healthy state. Multinomial stands for multinomial resampling at every time step. Adaptive means systematic resampling when $ESS < N/2$ (Equation (25)).

Filter	Representation	Weighting or Update	Resampling	Specific Settings
SIR PF	3×10^4 particles	likelihood (Equation (23))	multinomial	-
SIR-ESS PF	3×10^4 particles	likelihood, accumulated	adaptive	-
A-SIR PF	3×10^4 particles	likelihood divided by the action (Equation (32))	multinomial	$m = 1.5$, $\epsilon = 10^{-12}$
A-SIR-ESS PF	3×10^4 particles	as A-SIR PF, accumulated	adaptive	$m = 1.5$, $\epsilon = 10^{-12}$
A-SIR-FK PF	3×10^4 particles	likelihood $\times$ incremental penalty (Equation (34))	multinomial	$\lambda = 0.3$, $\kappa = 1$
AC-SIR PF	3×10^4 particles	likelihood; degradation-rate prior (Equation (36))	multinomial	$\varphi = 0.5$, $\tau_u = 0.5$ s, $p_J = 0$
Tempered SIR PF	3×10^4 particles	likelihood raised to α_T	multinomial	$\alpha_T = 0.5$
Liu–West SIR PF	3×10^4 particles	likelihood; kernel shrinkage (Equation (37))	multinomial	$\delta = 0.99$
DF-PF	3×10^4 particles	likelihood, accumulated; learned thermal law and $d_j \leq 1 + \delta$ in the transition (Equation (38))	adaptive	$s = 1$, $\delta = 0.02$
EKF	Gaussian	Kalman update, finite-difference Jacobians	-	Joseph form
UKF	Gaussian, 31 sigma points	Kalman update, unscented transform	-	$\alpha = 1$, $\beta = 2$, $\kappa = 0$
EnKF	200 members	Kalman update, perturbed observations	-	no inflation or localisation
DF-EKF	Gaussian	as EKF, with the model of the DF-PF; constraint by projection	-	$s = 0.5$, $\delta = 0.02$
LAS	Gaussian	EKF + fixed-lag RTS recursion (Equation (41))	-	lag $L = 1$ s

3.2. Broadband Excitation

To assess whether the conclusions depend on the weak observability of the stiffnesses under the harmonic input, the filters are also compared under a band-limited white-noise force (0–30 Hz) applied to the first mass, with the same RMS value as the harmonic force. This excitation drives the structural modes: the stiffness-dependent part of the accelerations increases from 0.07–0.09 to 2.0–3.8 m/s^2, a 1% stiffness change modifies the accelerations by 1.3–1.5 standard deviations of the measurement noise per sample (against 0.03 with the harmonic force), and the hysteretic element becomes strongly nonlinear (the nonlinear terms of Equation (6) reach 94% of the linear term). All the other settings, including the process noise of the filters, are unchanged.

3.3. Monte Carlo Design

The main campaign comprises 50 paired simulations for each of three scenarios: healthy structure, Step Damage and Progressive Damage. The SIR PF, SIR-ESS PF, A-SIR PF, A-SIR-ESS PF, A-SIR-FK PF, AC-SIR PF, DF-PF, EKF, UKF, EnKF, DF-EKF and the LAS are run on all of them, and the healthy simulations are used for training, calibration and as negative cases, following Section 2.8. Additional studies, each with 10 paired simulations per scenario and $N = 3 \times 10^4$ particles unless otherwise stated, are summarised in Table 3.

Table 3. Additional studies. Unless stated otherwise, each study uses 10 paired simulations for each of the healthy, step and progressive scenarios and $N = 3 \times 10^4$ particles.

Study	Modification with Respect to the Main Campaign
Control test	SIR PF re-run on the first 30 simulations with a different filter seed
Action ablation	A-SIR PF with $L = T$ only (30 simulations) and with $L = -V$ only (damaged scenarios)
Tempered SIR PF	damaged scenarios; also broadband excitation
Liu–West SIR PF	damaged scenarios
Coefficient m	A-SIR PF with $m \in \{0.015, 15\}$ (progressive scenario)
Number of particles	SIR PF and A-SIR PF with $N \in \{1000, 2000, 5000, 10{,}000\}$ ($N = 5000$: 30 simulations); SIR-ESS PF with the same N
Measurement noise	noise standard deviation multiplied by 2 (damaged scenarios) and by 4
Model mismatch	ground truth integrated with a fourth-order Runge-Kutta scheme (4 sub-steps) and Bouc–Wen parameters perturbed by $+20\%$ (β_{BW}), -20% (γ), $+10\%$ (n)
Reduced sensor set	accelerations of masses 1 and 2 only
Damage location	damage on Spring 1 (hysteretic element) or Spring 2 (damaged scenarios; healthy simulations of the main campaign)
Slow drift	time horizon three times longer (50 s), same temperature excursion and damage severity, 8 simulations
Broadband excitation	band-limited white-noise force (Section 3.2); all filters except the Liu–West SIR PF
Extended temperature	temperature decreasing from 40 to -40 °C (stiffness change of about 15%)
A-SIR-FK PF	also run in the measurement-noise, model-mismatch and reduced-sensor studies
Smoother lag	LAS with $L \in \{0.5, 2\}$ s

4. Results and Discussion

All the quantitative results reported in this section are based on the paired Monte Carlo campaign described in Section 3.3, while single simulations are shown only to illustrate the typical behaviour of the filters. Unless otherwise stated, the results refer to $N = 3 \times 10^4$ particles and 50 paired simulations per scenario, and paired differences are computed with respect to the SIR PF.

4.1. Overall Behaviour of the Filters

Figures 3 and 4 show the stiffness estimates of the four springs for a representative simulation of the Progressive Damage and Step Damage scenarios, respectively. The simulation was chosen as the one whose Spring-3 RMSEs are closest to the Monte Carlo medians of the displayed filters. With resampling at every step (left column), the SIR PF and the A-SIR PF follow the slow thermal increase of the healthy stiffnesses with wandering fluctuations of up to 0.3–0.5 N/m, which are often not covered by their narrow bands. Moreover, the two filters differ from each other by as much as they differ from the truth. After the onset of the progressive damage, they track the decreasing stiffness of Spring 3 with a lag that grows to about 1 N/m at the end of the record, whereas after the step damage they converge to the new stiffness level within about 1–2 s. In contrast, with adaptive resampling (right column), the SIR-ESS PF is practically superimposed on the UKF in both figures. After the initial transient, both filters follow the healthy stiffnesses with fluctuations of about 0.1 N/m and with bands that contain the true value. They also track the progressive degradation with a much smaller lag and converge after the step damage as quickly as the filters in the left column.

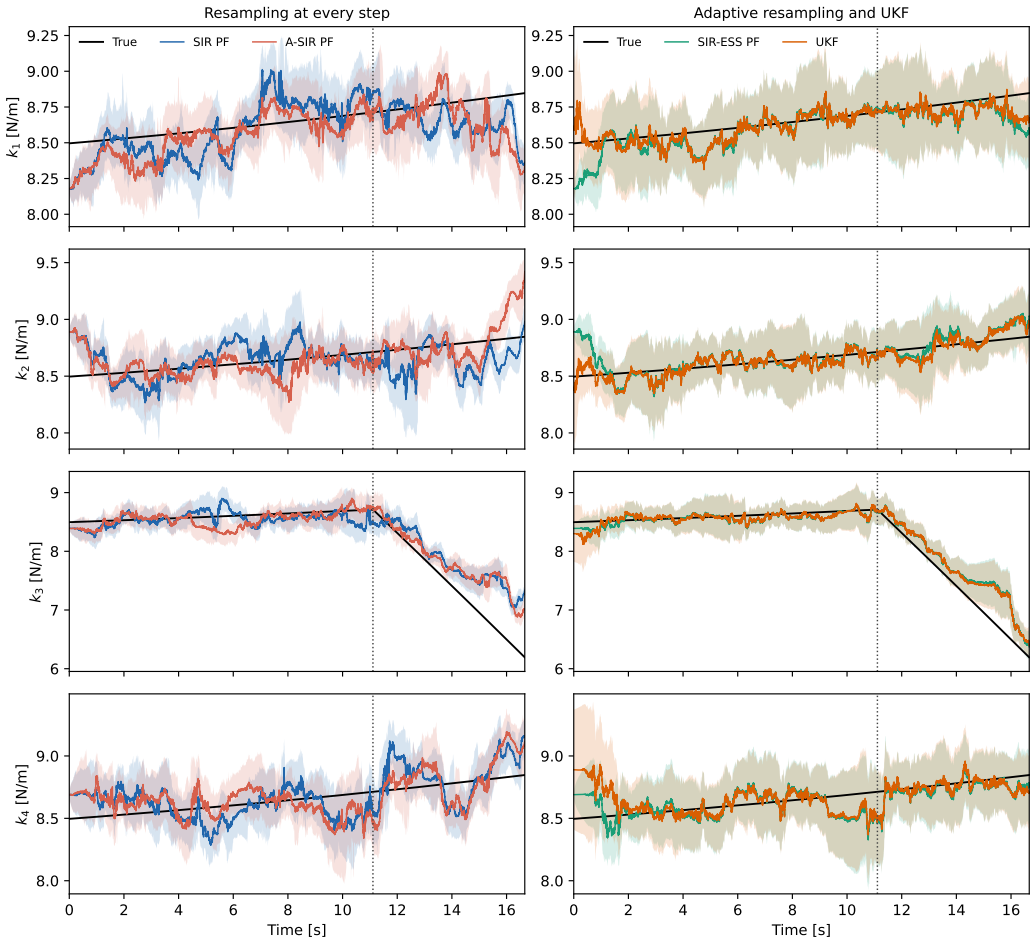


Figure 3. Stiffness estimates with 95% bands for a representative simulation of the Progressive Damage scenario. (**Left**): particle filters with multinomial resampling at every step (SIR PF and A-SIR PF); (**right**): SIR PF with adaptive systematic resampling (SIR-ESS PF) and UKF. Black: true stiffness; dotted vertical line: damage onset.

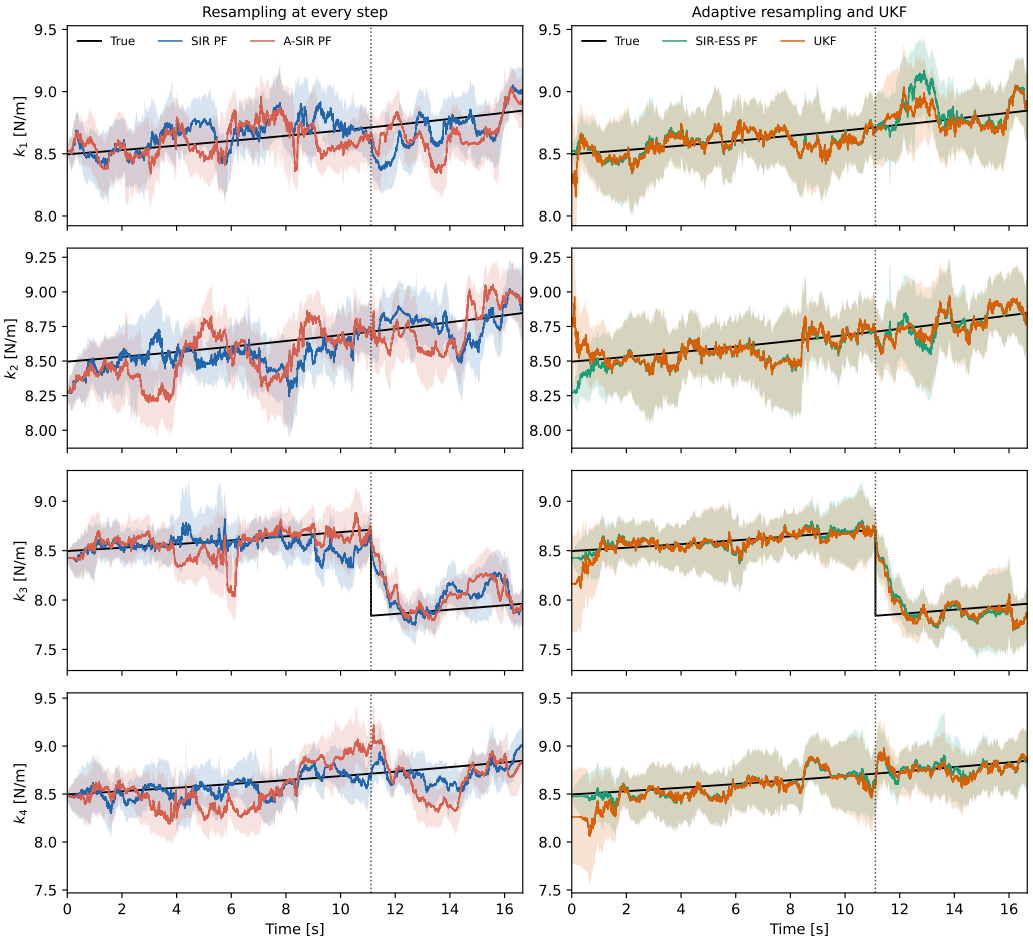


Figure 4. Stiffness estimates with 95% bands for a representative simulation of the Step Damage scenario. (**Left**): particle filters with multinomial resampling at every step (SIR PF and A-SIR PF); (**right**): SIR PF with adaptive systematic resampling (SIR-ESS PF) and UKF. Black: true stiffness; dotted vertical line: damage onset.

4.2. Stiffness Tracking

Table 4 summarises the tracking accuracy of the eleven filters and of the least-action smoother in the main campaign, while Figure 5 shows the distributions of the RMSE. The per-spring RMSE, MAE, bandwidth and coverage are reported in Tables A1–A3 in Appendix A. Four groups of methods can be identified, which are separated by horizontal lines in Table 4 and by the background colour in Figure 5: (i) particle filters with resampling at every step (white background), (ii) particle filters with adaptive resampling, including the DF-PF (green background), (iii) Kalman-type filters, including the DF-EKF (grey background), and (iv) the LAS (purple background). The LAS and the damage-factor filters are discussed separately in Sections 4.6 and 4.7, respectively.

Starting from the particle filters with resampling at every step, the SIR PF, the A-SIR PF and the two reformulations have errors 1.3–2.3 times larger than those of the Kalman-type filters. Also, they are over-confident, with a coverage of the 95% bands between 0.53 and 0.89, and they lose at least one stiffness in up to 14% (SIR PF and A-SIR PF) and 38% (AC-SIR PF) of the Progressive Damage simulations. Within this group, the paired differences between the A-SIR PF and the SIR PF are not statistically significant for any spring and any scenario, according to the Holm-corrected Wilcoxon signed-rank tests. Specifically, for Spring 3, the mean paired difference is $+0.002$ N/m (95% confidence interval $[-0.068, 0.071]$) in the progressive scenario and -0.007 N/m ($[-0.069, 0.055]$) in the step scenario, and the A-SIR PF gives the lower error in only 44% and 38% of the simulations,

respectively. Moreover, the A-SIR PF shows a heavier upper tail of the error distribution: in 5 of the 50 healthy simulations, against 1 for the SIR PF, at least one stiffness estimate departs from the truth by more than 1 N/m. The results of the two reformulations are discussed in Section 4.5.

Moving on to the particle filters with adaptive resampling, the SIR-ESS PF and the A-SIR-ESS PF practically reach the accuracy of the Kalman-type filters. Their Spring-3 RMSE is 0.245–0.249 N/m in the Progressive Damage scenario and 0.134–0.136 N/m in the Step Damage scenario, i.e., only 2% and 6% above that of the EKF and the UKF, and the RMSE of the healthy springs is 0.08–0.09 N/m. In addition, no run is lost, and the coverage of the bands (0.77 for the progressively damaged spring, 0.94–1.00 otherwise) is the same as that of the Kalman-type filters. Finally, the A-SIR-ESS PF and the SIR-ESS PF are statistically indistinguishable, with paired Spring-3 differences of -0.004, -0.001 and 0.000 N/m in the three scenarios ($p \geq 0.19$).

Concerning the Kalman-type filters, the EKF, the UKF and the EnKF give the lowest errors in all scenarios. Their Spring-3 RMSE is 0.243–0.244 N/m in the Progressive Damage scenario and 0.127–0.131 N/m in the Step Damage scenario, against 0.385 and 0.214 N/m for the SIR PF, while the RMSE of the healthy springs is 0.07–0.09 N/m and no run is lost. The EKF and the UKF are practically indistinguishable, and the EnKF with 200 members is only 1–5% worse. Since the EKF performs as well as the UKF, the nonlinearity of the estimation problem is mild in the main case study: the Bouc–Wen element operates in its linear range, and the parameters enter the equations of motion bilinearly. The 95% bands of these filters are conservative for the healthy springs, with a coverage close to 1, and slightly too narrow for the progressively damaged spring, with a coverage of 0.77–0.78. This is due to the random-walk model of the stiffness, which lags behind a sustained degradation.

Table 4. Stiffness tracking in the main case study. Mean over 50 paired simulations per scenario of the RMSE [N/m] of Spring 3 (k_3) and of the healthy springs ($k_{1,2,4}$). Cov. stands for empirical coverage of the 95% bands (Spring 3 in the damaged scenarios, all springs in the healthy one). Lost means percentage of runs with RMSE > 0.5 N/m on at least one spring. Time is the mean wall-clock time of one run in the parallel campaign. In bold are reported the values for which the Holm-corrected Wilcoxon signed-rank test of the paired difference with respect to the SIR PF is significant at the 5% level ($^-$ lower, $^+$ higher error).

	Progressive Damage				Step Damage				Healthy			
Filter	k_3	$k_{1,2,4}$	**Cov.**	**Lost [%]**	k_3	$k_{1,2,4}$	**Cov.**	**Lost [%]**	**RMSE**	**Cov.**	**Lost [%]**	**Time [s]**
SIR PF	0.385	0.188	0.67	14	0.214	0.182	0.81	4	0.142	0.88	0	338
A-SIR PF	0.386	0.227	0.65	14	0.207	0.217	0.80	4	0.245	0.84	10	355
A-SIR-FK PF	0.323	0.172	0.66	10	0.207	0.180	0.79	4	0.211	0.85	6	349
AC-SIR PF	**0.553**$^+$	**0.326**$^+$	0.53	38	0.359	0.322	0.66	16	0.283	0.73	12	559
SIR-ESS PF	**0.249**$^-$	**0.088**$^-$	0.77	0	**0.136**$^-$	**0.093**$^-$	0.95	0	**0.079**$^-$	1.00	0	315
A-SIR-ESS PF	**0.245**$^-$	**0.088**$^-$	0.77	0	**0.134**$^-$	**0.091**$^-$	0.95	0	**0.079**$^-$	1.00	0	334
DF-PF	**0.248**$^-$	**0.069**$^-$	0.77	0	**0.126**$^-$	**0.065**$^-$	0.95	0	**0.060**$^-$	1.00	0	373
EKF	**0.243**$^-$	**0.085**$^-$	0.78	0	**0.127**$^-$	**0.085**$^-$	0.96	0	**0.077**$^-$	1.00	0	17
UKF	**0.243**$^-$	**0.085**$^-$	0.78	0	**0.128**$^-$	**0.084**$^-$	0.96	0	**0.077**$^-$	1.00	0	14
EnKF	**0.244**$^-$	**0.088**$^-$	0.77	0	**0.131**$^-$	**0.088**$^-$	0.96	0	**0.081**$^-$	1.00	0	15
DF-EKF	0.322	**0.059**$^-$	0.63	0	**0.148**$^-$	**0.055**$^-$	0.92	0	**0.046**$^-$	1.00	0	21
LAS ($L = 1$ s)	**0.091**$^-$	**0.059**$^-$	0.96	0	**0.086**$^-$	**0.064**$^-$	0.94	0	**0.052**$^-$	1.00	0	40

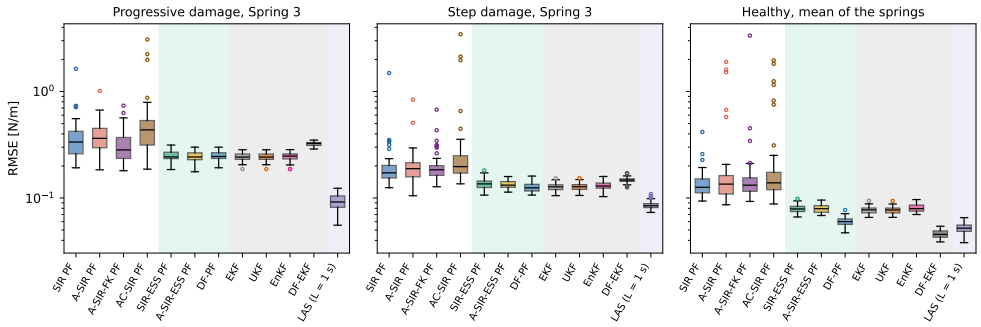


Figure 5. Boxplot of the Monte Carlo distributions (50 simulations, logarithmic scale) of the RMSE of Spring 3 in the damaged scenarios and of the mean RMSE of the four springs in the healthy scenario. Background colors highlight common methods: white is used for PFs, green for ESS PFs, and grey for KFs.

4.3. Effect of the Resampling Scheme

The comparison between the two groups of particle filters isolates the effect of the resampling scheme, since the SIR PF and the SIR-ESS PF, as well as the A-SIR PF and the A-SIR-ESS PF, share the model, the weights, the number of particles, the data and the random-number stream. In this problem, a single sample carries little information on the stiffnesses, as noted in Section 3. As reported in Table 4, compared with resampling at every step, the adaptive scheme reduces the Spring-3 RMSE by 35% and 37% in the progressive and step scenarios, respectively. The RMSE of the other springs is reduced by about 50%, with a lower error in 82–100% of the paired simulations. In addition, the lost runs and the over-confidence of the bands are eliminated, and the adaptive scheme is slightly cheaper (315 vs. 338 s per run), because resampling is performed in about 1% of the steps. The same improvement is obtained for the A-SIR PF, whose Spring-3 RMSE decreases from 0.386 to 0.245 N/m and from 0.207 to 0.134 N/m. Thus, the gain is entirely due to the resampling scheme, and it is much larger than the effect of the action term and about twice that of the best weighting modification, i.e., the A-SIR-FK PF (-0.062 N/m in the progressive scenario).

With adaptive resampling, the particle filters reach the accuracy of the Kalman-type filters. This means that, in this benchmark, the larger errors of the particle filters were due to the resampling noise rather than to the particle approximation itself [71]. Adaptive resampling also reduces the number of particles required from 3×10^4 to about 2000, as discussed in Section 4.10.

4.4. Damage Detection and Localisation

Detection and localisation were analysed with paired comparisons over the same simulations, reported in Table 5, and with the ROC curves shown in Figure 6, using the three threshold definitions of Section 2.8. Figure 6 shows the ROC curves of the AE-based indicator of Spring 3 for the main filters of each group. The curves of the SIR-ESS PF, the UKF and the LAS coincide with the axes, as their AUC is equal to 1.

With the per-run thresholds of ref. [31] (Equation (43)), the SIR PF and the A-SIR PF detect the damage in 90–100% of the simulations and localise it on Spring 3 in most of them. However, at equal run-level false-alarm rates, the A-SIR PF detects and localises the damage less often than the SIR PF, i.e., in 30% vs. 98% of the simulations in the step scenario and in 40% vs. 100% in the progressive scenario. This is because its occasional large excursions in the healthy simulations raise the thresholds required to keep the false-alarm rate fixed. Similarly, the A-SIR-FK PF (60%) and, above all, the AC-SIR PF (0%) are below the SIR PF. With the robust calibration, the SIR PF and the A-SIR PF detect the damage in a similar fraction of the simulations (100% and 90–96%, respectively), but the A-SIR PF remains slower, by about 0.8 s.

In contrast, the particle filters with adaptive resampling, the Kalman-type filters and the damage-factor filters detect and localise the damage in all simulations and at every operating point. Their median delays are similar, i.e., 0.43–0.57 s and 1.16–1.34 s at equal false-alarm rates in the step and progressive scenarios, respectively, against 1.59 and 2.13 s for the SIR PF. They also show larger localisation margins than the SIR PF. However, with the per-run thresholds, they raise a persistent alarm in 12–30% of the healthy simulations, and the SIR PF and the A-SIR PF in about half of them (52% and 48%, respectively). This is because the per-run thresholds are computed on the short baseline window of each run, so that they do not account for the variability of the indicator over the rest of the monitoring period. Thus, the comparisons between filters are based on the thresholds calibrated on independent healthy simulations.

The threshold-free AUC, reported in Table 6 and Figure 6, confirms this ranking. The AUC is equal to 1.00 for the SIR-ESS PF, the A-SIR-ESS PF, the Kalman-type filters, the damage-factor filters and the LAS, whereas it is 0.996/0.999 (step/progressive) for the SIR PF, 0.980/0.980 for the A-SIR-FK PF, 0.865/0.886 for the A-SIR PF and 0.868/0.878 for the AC-SIR PF.

The control test in Table 5, in which the SIR PF is re-run on 30 of the simulations with a different filter seed, shows that the operating-point metrics of the particle filters with resampling at every step are themselves very variable. The two runs of the same filter have AUCs that are not significantly different (0.968/0.968 for the re-run) but, at equal run-level false-alarm rate, the step damage is detected in 97% and 47% of the common simulations, respectively. Therefore, differences between these filters in thresholded detection and localisation metrics are dominated by rare divergent runs, which affect whichever filter they occur in. For instance, in the model-mismatch study of Section 4.10, it is the SIR PF that diverges in a healthy simulation, and the comparison is reversed.

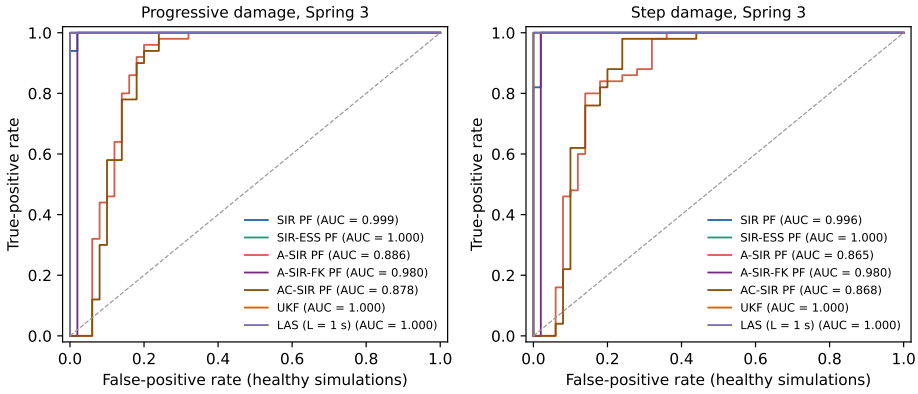


Figure 6. ROC curves of the AE-based indicator of Spring 3 for the main filters of each group, where score is the maximum of μ_{SE_3} after the damage onset, positives are the damaged simulations, while negatives are the healthy simulations. The dashed line represent the bisector (AUC = 0.5), blue is used for the SIR PF, green for the SIR-ESS PF, dark orange represent the A-SIR PF, purple the A-SIR-FK PF, brown refers to AC-SIR PF, orange to UKF, and indigo is used for the LAS. The curves of the SIR-ESS PF, the UKF and the LAS coincide with the axes, as their AUC is equal to 1.

Table 5. Paired analysis of damage detection and localisation. The AE detector is used, Spring 3 is damaged, and 50 simulations per scenario are considered (30 for the control test). Det. stands for the fraction of damaged simulations with a persistent alarm on Spring 3. Delay is the median detection delay. 1st means the fraction of simulations in which the first persistent alarm after the onset is raised on Spring 3. Loc. is the fraction of the post-onset time in which Spring 3 has the largest normalised indicator. FA_0 stands for the fraction of healthy simulations with at least one persistent alarm. In bold are reported the values that differ from those of the SIR PF at the 5% level (McNemar test for proportions, Wilcoxon signed-rank test for delays and localisation fractions, uncorrected). The SIR PF′ (A/A) row is the control test, in which the SIR PF is re-run on the same data with a different filter seed.

		Step Damage				Progressive Damage				
Thresholds	**Method**	**Det.**	**Delay [s]**	**1st**	**Loc.**	**Det.**	**Delay [s]**	**1st**	**Loc.**	**FA_0**
Per-run $ATSE^{99}$ (original)	SIR PF	0.98	0.61	0.72	0.85	1.00	1.75	0.78	0.77	0.52
	A-SIR PF	0.90	0.72	0.58	**0.74**	1.00	1.59	**0.56**	0.69	0.48
	A-SIR-FK PF	0.98	0.67	0.84	0.86	1.00	1.46	0.80	0.80	0.44
	AC-SIR PF	0.94	**0.97**	0.62	**0.75**	0.90	1.75	0.58	0.67	0.68
	SIR-ESS PF	1.00	**0.38**	**0.90**	**0.96**	1.00	**1.21**	0.92	**0.87**	0.16
	A-SIR-ESS PF	1.00	**0.41**	0.90	**0.96**	1.00	**1.20**	0.92	**0.87**	0.16
	DF-PF	1.00	**0.33**	**0.90**	**0.97**	1.00	**1.12**	0.90	**0.88**	0.30
	EKF	1.00	**0.37**	**0.90**	**0.96**	1.00	**1.19**	0.94	**0.88**	0.14
	UKF	1.00	**0.38**	0.86	**0.96**	1.00	**1.18**	**0.94**	**0.87**	0.14
	EnKF	1.00	**0.41**	0.76	**0.95**	1.00	**1.24**	0.84	**0.86**	0.18
	DF-EKF	1.00	**0.45**	**0.92**	**0.95**	1.00	**1.22**	0.90	**0.87**	0.12
	LAS ($L = 1$ s)	1.00	**1.00**	0.88	**1.00**	1.00	1.53	0.88	**0.96**	0.32
	SIR PF′ (A/A)	0.97	1.00	0.73	0.79	0.97	1.52	0.73	0.74	0.37
Equal run-level false-alarm rates (10%)	SIR PF	0.98	1.59	0.96	0.86	1.00	2.13	0.98	0.77	0.10
	A-SIR PF	**0.30**	**2.34**	**0.30**	**0.48**	**0.40**	**3.28**	**0.36**	**0.40**	0.10
	A-SIR-FK PF	**0.60**	**0.91**	**0.56**	**0.65**	**0.60**	**1.56**	**0.58**	**0.58**	0.10
	AC-SIR PF	**0.00**	–	**0.00**	**0.25**	**0.00**	–	**0.00**	**0.27**	0.10
	SIR-ESS PF	1.00	**0.57**	0.98	**0.96**	1.00	**1.34**	0.98	**0.86**	0.10
	A-SIR-ESS PF	1.00	**0.48**	0.98	**0.96**	1.00	**1.31**	0.98	**0.86**	0.10
	DF-PF	1.00	**0.47**	0.98	**0.96**	1.00	**1.23**	0.98	**0.87**	0.10
	EKF	1.00	**0.45**	1.00	**0.97**	1.00	**1.29**	1.00	**0.89**	0.10
	UKF	1.00	**0.46**	1.00	**0.97**	1.00	**1.29**	1.00	**0.88**	0.10
	EnKF	1.00	**0.45**	0.96	**0.96**	1.00	**1.32**	0.94	**0.87**	0.10
	DF-EKF	1.00	**0.43**	1.00	**0.97**	1.00	**1.16**	0.98	**0.91**	0.10
	LAS ($L = 1$ s)	1.00	**1.00**	0.98	**1.00**	1.00	**1.67**	0.98	**0.96**	0.10
	SIR PF′ (A/A)	**0.47**	1.45	**0.37**	0.70	**0.60**	2.13	**0.47**	0.67	0.10
Robust calibration (median + 3 MAD)	SIR PF	1.00	1.07	0.86	0.85	1.00	1.89	0.88	0.72	0.28
	A-SIR PF	0.90	**1.93**	0.80	0.83	0.96	**2.65**	0.86	0.74	0.30
	A-SIR-FK PF	1.00	1.22	0.88	0.86	1.00	2.02	0.90	0.79	0.30
	AC-SIR PF	0.98	1.21	0.78	**0.75**	0.96	**2.07**	0.76	0.69	0.34
	SIR-ESS PF	1.00	**0.46**	0.94	**0.96**	1.00	**1.26**	**1.00**	**0.86**	0.22
	A-SIR-ESS PF	1.00	**0.43**	0.98	**0.96**	1.00	**1.27**	0.98	**0.87**	0.18
	DF-PF	1.00	**0.32**	0.94	**0.97**	1.00	**1.10**	0.92	**0.89**	0.32
	EKF	1.00	**0.45**	0.98	**0.96**	1.00	**1.26**	**1.00**	**0.86**	0.16
	UKF	1.00	**0.45**	0.98	**0.96**	1.00	**1.26**	**1.00**	**0.86**	0.16
	EnKF	1.00	**0.37**	0.84	**0.96**	1.00	**1.21**	0.88	**0.88**	0.22
	DF-EKF	1.00	**0.41**	**1.00**	**0.97**	1.00	**1.11**	**1.00**	**0.91**	0.22
	LAS ($L = 1$ s)	1.00	1.00	0.98	**1.00**	1.00	**1.61**	**1.00**	**0.96**	0.26
	SIR PF′ (A/A)	1.00	1.17	0.77	0.80	1.00	1.84	0.73	0.71	0.27

Table 6. Threshold-free detection performance. AUC of the AE-based indicator of Spring 3 and paired difference with respect to the SIR PF, with the 95% paired-bootstrap confidence interval (2000 resamples of the simulations). The score is the maximum of μ_{SE_3} after the onset, and the negatives are the healthy simulations.

	Step Damage		Progressive Damage	
Filter	**AUC**	**ΔAUC [95% CI]**	**AUC**	**ΔAUC [95% CI]**
SIR PF	0.996	–	0.999	–
A-SIR PF	0.865	−0.131 [−0.214, −0.062]	0.886	−0.113 [−0.193, −0.047]
A-SIR-FK PF	0.980	−0.016 [−0.060, +0.010]	0.980	−0.019 [−0.060, +0.004]
AC-SIR PF	0.868	−0.128 [−0.214, −0.050]	0.878	−0.120 [−0.206, −0.048
SIR-ESS PF	1.000	+0.004 [+0.000, +0.013]	1.000	+0.001 [+0.000, +0.005]
A-SIR-ESS PF	1.000	+0.004 [+0.000, +0.013]	1.000	+0.001 [+0.000, +0.005]
DF-PF	1.000	+0.004 [+0.000, +0.013]	1.000	+0.001 [+0.000, +0.005]
EKF	1.000	+0.004 [+0.000, +0.013]	1.000	+0.001 [+0.000, +0.005]
UKF	1.000	+0.004 [+0.000, +0.013]	1.000	+0.001 [+0.000, +0.005]
EnKF	1.000	+0.004 [+0.000, +0.013]	1.000	+0.001 [+0.000, +0.005]
DF-EKF	1.000	+0.004 [+0.000, +0.013]	1.000	+0.001 [+0.000, +0.005]
LAS ($L = 1$ s)	1.000	+0.004 [+0.000, +0.013]	1.000	+0.001 [+0.000, +0.005]
SIR PF′ (A/A)	0.968	−0.024 [−0.096, +0.021]	0.968	−0.029 [−0.097, +0.010]

4.5. The Action-Aided Weighting

Table 7 reports the modification of the particle weights introduced by the action term, measured directly along an A-SIR run with $N = 3 \times 10^4$ particles. The total-variation distance between the normalised likelihood-only weights and the normalised A-SIR weights is at most of the order of 10^{-5} per step, and it decreases over time to about 10^{-6}. Similarly, the coefficient of variation of $1/S_k^i$ across particles decreases from 10^{-4} to 3×10^{-6}. This is the expected consequence of the properties discussed in Section 2.3.2. As shown in the last column of Table 7, the kinetic-like term exceeds the potential-like term by about four orders of magnitude, with the latter being negligible. This term is dominated by the increments of the measurement noise, which are common to all particles, and it is accumulated over time, so that the relative differences of S_k^i between particles become negligible.

Thus, the action term acts mainly as a tiny perturbation of the resampling, which makes the two filters follow different but statistically equivalent trajectories, through the chaotic dependence of the particle system on its random inputs. Indeed, in the control test, in which the SIR PF is run twice on the same data with two different filter seeds, the paired differences are of the same magnitude as, or larger than, those between the SIR PF and the A-SIR PF. As reported in Table A4 and Figure 7, on the same 30 simulations, the standard deviation of the paired Spring-3 RMSE differences is 0.75 and 0.78 N/m in the progressive and step scenarios for the control test, against 0.17 and 0.13 N/m for the A-SIR PF. This means that the outcome of a single simulation is dominated by the randomness of the filter itself. The same conclusion holds with adaptive resampling. As reported in Table 4, the A-SIR-ESS PF and the SIR-ESS PF have practically identical errors, because the product of the factors $1/S_k^i$ accumulated between two resampling events remains a negligible perturbation of the accumulated likelihood.

The ablations reported in Table A4 are consistent with this interpretation. Using only the potential-like term ($L = -V$) makes the accumulated score oscillate around zero, so that $1/S$ becomes erratic and the filter loses track in all simulations. Instead, using only the kinetic-like term ($L = T$), which differs from the full A-SIR PF by a term about 10^4 times smaller, gives a mean Spring-3 RMSE difference with respect to the SIR PF of -0.064 N/m in the progressive scenario and $+0.003$ N/m in the step scenario. Given the negligible size of the potential-like term and the outcome of the control test, this difference is attributed to the variability of the filter. Finally, flattening the weights with a tempered likelihood gives a small, non-significant reduction in the errors, and the Liu–West kernel smoothing does not improve the SIR PF in this benchmark.

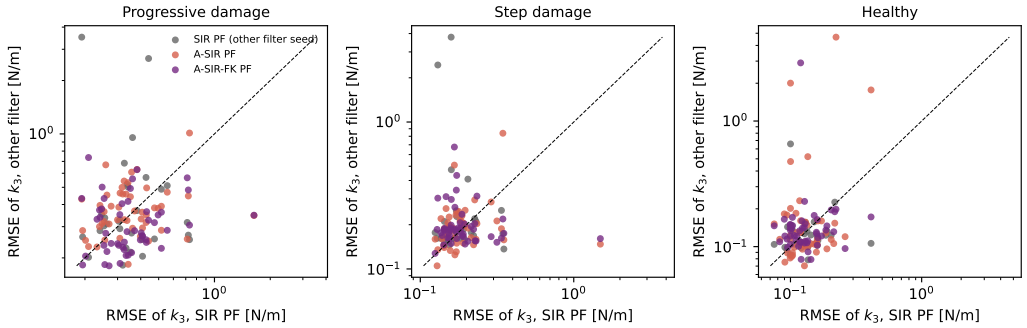


Figure 7. Paired Spring-3 RMSE of the SIR PF (horizontal axis) and of a second filter run on the same data (vertical axis): the SIR PF with a different filter seed (control test) in grey, the A-SIR PF in red, and the A-SIR-FK PF in purple. The dashed line is the bisector.

Table 7. Effect of the action term on the particle weights along an A-SIR PF run ($N = 3 \times 10^4$, Progressive Damage). Medians over the indicated time-step ranges are reported.TV stands for the total-variation distance between the normalised likelihood-only and A-SIR weights. CV is the coefficient of variation of $1/S_k^i$ across particles. σ_ℓ is the standard deviation of the log-likelihood across particles. ESS means the effective sample size (% of N) of the likelihood-only and A-SIR weights. Δa is the fraction of resampled ancestor indices that change when the action term is removed, with the same particles and the same uniform numbers. T/V is the ratio of the particle-averaged kinetic-like and potential-like terms.

Steps	TV	CV(1/S)	σ_ℓ	ESS$_\ell$ [%]	ESS$_S$ [%]	Δa [%]	T/V
1–100	5.9×10^{-5}	1.5×10^{-4}	0.108	98.9	98.9	0.76	1.6×10^4
100–1000	3.2×10^{-5}	8.0×10^{-5}	0.066	99.6	99.6	0.38	2.4×10^4
1000–10,000	6.3×10^{-6}	1.5×10^{-5}	0.076	99.4	99.4	0.07	2.4×10^4
10,000–30,000	1.2×10^{-6}	3.0×10^{-6}	0.072	99.5	99.5	0.02	2.4×10^4

Moving on to the incremental (Feynman–Kac) weighting, the penalty ($\kappa \in \{0, 1\}$) and the strength λ of the A-SIR-FK PF were selected on the tuning simulations. The selected configuration ($\kappa = 1$, $\lambda = 0.3$) reduced the errors by 6% on the tuning simulations, whereas the kinetic-only penalty did not improve on the SIR PF for any λ. In the 50 test simulations, the A-SIR-FK PF gives a Spring-3 RMSE of 0.323 N/m in the progressive scenario, against 0.385 N/m for the SIR PF. The paired difference is -0.062 N/m (95% interval $[-0.131, 0.007]$), with a lower error in 68% of the simulations and $p = 0.036$ before correction, which is not significant after the correction for multiple comparisons. In the step scenario, the Spring-3 RMSE is 0.207 against 0.214 N/m, and in the healthy simulations the errors are slightly larger, although not significantly. Therefore, the reformulated weighting shows a modest tendency to improve the tracking of the progressive degradation, which does not reach statistical significance. This tendency does not translate into better detection, and it is less than half of the effect of the resampling scheme. Moreover, the improvement observed in the tuning simulations shrank in the test simulations, which illustrates once more the importance of independent repeated evaluations.

Concerning the action-consistent prior, six configurations (φ, τ_u, p_J) of the AC-SIR PF were evaluated on the tuning simulations: $(0.5, 0.5\text{ s}, 0)$, $(0.5, 2\text{ s}, 0)$, $(0.9, 0.5\text{ s}, 0)$, $(0.9, 2\text{ s}, 0)$, $(0.5, 2\text{ s}, 5 \times 10^{-6})$ and $(0.5, 2\text{ s}, 2 \times 10^{-5})$. None of them outperformed the SIR PF. Specifically, the errors grew with the persistence of the degradation rate (φ, τ_u), and the jump component was particularly detrimental in the healthy simulations. On the 50 test simulations, the selected configuration $(0.5, 0.5\text{ s}, 0)$ gives Spring-3 RMSEs of 0.553 and 0.359 N/m in the progressive and step scenarios, respectively, and lower AUCs, so that the AC-SIR PF is significantly worse than the SIR PF. Since the stiffnesses are only

weakly observable from each sample, a prior with a persistent degradation rate lets the particle population drift in directions that the data correct too slowly, and jumps introduce excursions that are rejected late. Thus, in this benchmark, the limiting factor of the PFs is the combination of weak per-sample information and resampling noise, rather than a lack of temporal regularisation of the trajectories.

4.6. Least-Action Smoothing (LAS)

The LAS of Section 2.5 is the only formulation of the action idea that improves on the best filters. As reported in Table 4, with a lag of 1 s its Spring-3 RMSE is 0.091 and 0.086 N/m in the progressive and step scenarios, respectively, against 0.243 and 0.127 N/m for the EKF on which it is built. This corresponds to reductions of 62% and 33%, with a lower error in 100% and 98% of the simulations ($p < 0.001$), while the RMSE of the healthy springs decreases by 25–33%. The improvement is largest for the progressive degradation. A filter can follow a sustained degradation only with a delay, because the random-walk model initially attributes the deviations to noise, whereas the LAS uses the following second of data to place the degradation at the correct time. In addition, the coverage of the 95% bands of Spring 3 rises from 0.78 to 0.96, with bands about 40% narrower (0.33 vs. 0.54 N/m). For the step damage, instead, the smoother spreads the abrupt change over the lag window around the onset, which limits the gain on the damaged spring. Longer lags reduce the errors further: on the first 10 simulations, the Spring-3 RMSE in the progressive scenario is 0.138, 0.087 and 0.054 N/m for $L = 0.5$, 1 and 2 s, respectively, against 0.232 N/m for the EKF.

Concerning detection, Table 5 shows that the LAS raises a persistent alarm on Spring 3 in all damaged simulations, with an AUC of 1.00, and that it localises the damage better than the EKF. Specifically, the fraction of the post-onset time in which Spring 3 has the largest indicator is 1.00/0.96 (step/progressive), against 0.97/0.89 for the EKF, and the localisation margin is larger. However, at equal latency, i.e., with the 1 s lag added, its median detection delay is about 0.4–0.6 s longer than that of the EKF, i.e., 1.00 vs. 0.45 s and 1.67 vs. 1.29 s at equal false-alarm rates. Under broadband excitation, the LAS further reduces the errors of the EKF to 0.013–0.042 N/m, as reported in Table 8. Its computational cost is about 2.4 times that of the EKF (40 vs. 17 s per run), i.e., about one eighth of that of a particle filter with $N = 3 \times 10^4$.

Table 8. Stiffness tracking under broadband excitation. Mean over 10 paired simulations per scenario of the RMSE [N/m] of Spring 3 (k_3) and of the healthy springs ($k_{1,2,4}$). Cov. stands for empirical coverage of the 95% bands (Spring 3 in the damaged scenarios, all springs in the healthy one). Lost means percentage of runs with RMSE > 0.5 N/m on at least one spring. Time is the mean wall-clock time of one run in the parallel campaign.

	Progressive Damage				**Step Damage**				**Healthy**			
Filter	k_3	$k_{1,2,4}$	**Cov.**	**Lost [%]**	k_3	$k_{1,2,4}$	**Cov.**	**Lost [%]**	**RMSE**	**Cov.**	**Lost [%]**	**Time [s]**
SIR PF	2.517	2.863	0.29	80	3.070	2.945	0.28	90	1.160	0.47	50	402
A-SIR PF	2.302	2.279	0.33	70	3.113	2.491	0.30	80	1.662	0.33	80	425
A-SIR-FK PF	1.447	1.218	0.42	60	1.844	1.441	0.41	70	2.624	0.29	80	412
AC-SIR PF	2.223	1.757	0.24	90	2.597	1.634	0.24	100	1.707	0.30	70	589
Tempered SIR PF	0.746	0.629	0.57	30	1.106	0.808	0.48	70	1.517	0.49	60	414
SIR-ESS PF	0.145	0.120	0.71	0	0.404	0.345	0.58	30	0.253	0.86	10	320
A-SIR-ESS PF	0.607	0.382	0.72	30	1.078	0.651	0.57	50	0.292	0.90	10	337
EKF	0.052	0.033	0.99	0	0.066	0.040	0.98	0	0.026	1.00	0	19
UKF	0.052	0.032	0.99	0	0.066	0.040	0.98	0	0.026	1.00	0	16
EnKF	0.128	0.132	0.64	10	0.145	0.131	0.64	0	0.104	0.75	0	16
LAS ($L = 1$ s)	0.016	0.016	1.00	0	0.042	0.025	0.97	0	0.014	1.00	0	40

It must be noted that the LAS is not a new algorithm, as it is the fixed-lag extended RTS smoother. Its interest here is that it is the form taken by the least-action idea once the action is defined consistently with the Bayesian formulation (Equation (40)), and that, in this form, the idea does improve the estimates, in contrast with its use as a weighting of the particles.

4.7. Damage-Factor Filtering

Moving on to the damage-factor filters, the hard constraint ($\delta = 0$) was the worst choice for both filters on the tuning simulations. The geometric-mean RMSE ratios with respect to the same filter without the damage-factor model were 1.66–2.76 for the DF-PF, against 0.76–0.97 with a margin, and 0.84–1.13 for the DF-EKF, against 0.71–0.98 with a margin. This is because, for a healthy spring, the true stiffness lies on the boundary of the admissible set, so that the truncation biases the estimates downwards.

On the 50 test simulations, the DF-PF reduces the RMSE of the healthy springs with respect to the SIR-ESS PF by 21–31%, with a lower error in all simulations, as reported in Table 4. On the damaged spring, it has the same RMSE in the progressive scenario (0.248 vs. 0.249 N/m) and a 7% lower one in the step scenario (0.126 vs. 0.136 N/m). The DF-PF is also more accurate than the EKF on the healthy springs, by 18–23% in all simulations, and is equally as accurate on the damaged spring (0.248 vs. 0.243 N/m and 0.126 vs. 0.127 N/m). Thus, it is the most accurate of the filters considered in this study, while the LAS remains more accurate on the damaged spring.

In detection, the DF-PF raises an alarm in all damaged simulations, with an AUC of 1.00, and its median delays are close to those of the Kalman-type filters, i.e., 0.47 and 1.23 s at equal false-alarm rates. Moreover, Tables 5 and 6 show that it localises the damage as well as the EKF, with a fraction of correctly localised time of 0.96 against 0.97 in the step scenario.

The DF-EKF selected by the same rule, with a halved process noise, reduces the errors on the healthy springs even more, by 30–43%. However, it tracks the degradation more slowly, with Spring-3 RMSEs of 0.322 and 0.148 N/m, although its median detection delays, i.e., 0.43 and 1.16 s, are not longer than those of the other filters. With the full process noise, its errors are close to those of the EKF (tuning ratio of 0.95), because the constraint imposed by projection is much less effective than in the particle filter.

Thus, this prior knowledge is most useful in combination with a particle representation, which handles the constraint exactly. It must be noted that part of the gain on the healthy springs derives from the learned thermal law and from the calibrated healthy initial state, which are not available to the other filters. Finally, the cost of the DF-PF is similar to that of the SIR-ESS PF (373 vs. 315 s per run with $N = 3 \times 10^4$), against 21 s for the DF-EKF.

4.8. Results for the Broadband Excitation

In the main case study, the superiority of the Kalman-type filters is not surprising. The Bouc–Wen element operates in its linear range, the noise is Gaussian and the parameters enter the equations of motion bilinearly, so that the posterior is close to Gaussian and a Gaussian filter is nearly optimal. The broadband excitation described in Section 3.2 provides a more demanding test, in which the stiffnesses are about 45 times more observable per sample and the hysteresis is strongly nonlinear. The results of this test are reported in Table 8.

In this regime, the SIR PF and the A-SIR PF lose track of the stiffnesses in most damaged simulations. With the settings of the main case study, the process noise of the kinematic states is negligible. Thus, when the response is large and nonlinear, the errors that accumulate on positions and velocities cannot be corrected, and the informative likelihood makes the weights degenerate. Consistently, the remedies that act against weight degeneracy help. The tempered SIR PF reduces the Spring-3 RMSE to 0.75 and 1.11 N/m in the progressive and step scenarios, respectively ($p = 0.020$ and $p = 0.049$). The SIR-ESS PF, which accumulates the weights and

resamples on average every 5 steps, reduces it further to 0.15 and 0.40 N/m, and to 0.25 N/m in the healthy scenario, in which tempering does not help. However, the SIR-ESS PF still loses track in 30% of the step simulations, and the A-SIR-ESS PF is less accurate, with RMSEs of 0.61 and 1.08 N/m and 30–50% of the damaged runs lost. The A-SIR-FK PF is better than the SIR PF in the damaged scenarios (1.45 and 1.84 N/m) but worse in the healthy one, and the AC-SIR PF is not significantly different from the SIR PF.

In contrast, the EKF and the UKF, which correct the kinematic states through the Kalman gain, track all stiffnesses with RMSEs of 0.02–0.07 N/m, i.e., more accurately than under the harmonic input. Also, the EKF is as accurate as the UKF, despite the strongly nonlinear hysteresis. The EnKF without inflation is less accurate (0.10–0.15 N/m) and over-confident, with a coverage of about 0.65–0.75, which is a known consequence of the underestimation of the ensemble spread when the observations are informative [41]. Hence, also in the informative regime, the action-based formulations do not provide an advantage over established remedies or over the Kalman-type filters. A proper treatment of this regime with particle filters would require a re-design of the kinematic process noise or of the proposal distribution, which is beyond the scope of this work.

4.9. Results of the Environmental Normalisation

Table 9 reports the AUC of Spring 3 obtained by replacing the AE reconstruction error with the conditional, raw and oracle indicators of Section 2.6. In the main case study, the thermal effect is small compared with the estimation noise, i.e., about 4% over the whole record and 0.06 N/m within the healthy monitoring interval. For instance, for the SIR PF, the AUC of Spring 3 (step/progressive) is:

- 0.975/0.991 without any normalisation;
- 0.996/0.999 with the AE;
- 0.994/0.998 with the conditional normalisation;
- 0.999/0.999 with the exact (oracle) thermal law (used as baseline, unfeasible for real experimental tests).

For the Kalman-type filters and the SIR-ESS PF, all four indicators give an AUC of 1.00. Instead, for the A-SIR PF, the two learned normalisations give a lower AUC than the indicator without normalisation, i.e., 0.865/0.886 with the AE and 0.670/0.762 with the conditional model, against 0.908/0.926. The indicator without normalisation is referred to the baseline mean of each run, whereas the learned models are referred to all the healthy training simulations, and they are therefore more affected by the large excursions of the estimates that occur in a few simulations of the A-SIR PF. This highlights that any normalisation can only be as good as the estimates it is trained on.

A larger thermal effect is obtained in the extended-temperature study, in which the temperature decreases from 40 to −40 °C and the stiffness changes by about 15%. In this scenario, the indicator without normalisation fails completely, with an AUC of Spring 3 of 0.00 and 0.36 in the step and progressive scenarios, respectively, and false alarms on all springs in all healthy simulations. In contrast, the AE, the conditional normalisation and the oracle law give an AUC of 1.00 in both scenarios and for both filters. Thus, the environmental normalisation is necessary when the thermal effect is comparable with the damage, and the AE and the conditional model are equally effective when they are trained on the estimates of the same filter over the same temperature range. However, the AE is effective only if it actually learns the healthy relationship between temperature and stiffness, which requires the standardisation of the inputs and a non-saturating bottleneck, as discussed in Sections 2.6 and 3. This can be verified on held-out healthy simulations through the fraction of the variance of the inputs explained by the reconstruction.

Table 9. AUC of the Spring-3 indicator (Step/Progressive Damage) with different environmental normalisations. None (raw) stands for the deviation from the baseline mean, without any normalisation. AE is the reconstruction error of the autoencoder. Conditional means the regression of the stiffness on the temperature. Oracle is the exact thermal law, which is not available in practice. The main case study uses 50 simulations per scenario, and the extended temperature range (40 to −40 °C) uses 10 simulations per scenario.

	Main Case Study					Extended Temperature	
Normalisation	**SIR PF**	**SIR-ESS PF**	**A-SIR PF**	**EKF**	**UKF**	**SIR PF**	**A-SIR PF**
None (raw)	0.975/0.991	1.000/1.000	0.908/0.926	1.000/1.000	1.000/1.000	0.000/0.360	0.130/0.320
AE	0.996/0.999	1.000/1.000	0.865/0.886	1.000/1.000	1.000/1.000	1.000/1.000	1.000/1.000
Conditional	0.994/0.998	1.000/1.000	0.670/0.762	1.000/1.000	1.000/1.000	1.000/1.000	1.000/1.000
Oracle law	0.999/0.999	1.000/1.000	0.933/0.939	1.000/1.000	1.000/1.000	1.000/1.000	1.000/1.000

4.10. Robustness and Sensitivity Studies

The results of the additional studies are summarised in Tables A4 and A5 in Appendix A. Each study uses 10 paired simulations per scenario and $N = 3 \times 10^4$ particles, unless otherwise stated.

4.10.1. Number of Particles

Table A5 reports the results obtained with different numbers of particles. With resampling at every step, the particle filters require a large number of particles. With $N = 1000$ and $N = 2000$, the stiffness RMSE is 1–4 N/m, i.e., the filters are unusable, and with $N = 5000$ both the SIR PF and the A-SIR PF lose track of at least one stiffness in 73–97% of the simulations. Using $N = 10^4$, the RMSE is still 0.25–1.2 N/m, and only with $N = 3 \times 10^4$ is it reduced to 0.13–0.39 N/m. In the damaged scenarios, the SIR PF and the A-SIR PF are not significantly different at any particle count.

With adaptive resampling, far fewer particles are needed, as the SIR-ESS PF does not lose any stiffness in any simulation already with $N = 1000$. Its Spring-3 RMSE with $N = 1000$, 2000, 5000 and 10^4 is 0.286, 0.250, 0.244 and 0.242 N/m in the progressive scenario, and 0.156, 0.148, 0.138 and 0.134 N/m in the step scenario, against 0.249 and 0.136 N/m with $N = 3 \times 10^4$ on the same simulations. Thus, with $N = 2000$, the SIR-ESS PF is practically as accurate as with $N = 3 \times 10^4$, and one run takes 21 s instead of 315 s, i.e., about as long as with the Kalman-type filters.

4.10.2. Measurement Noise

Doubling the noise standard deviation does not change the comparison between the SIR PF and the A-SIR PF significantly. The A-SIR PF has a lower Spring-3 RMSE in the step scenario in all 10 simulations, but not in the progressive one. Instead, with four times the nominal standard deviation, i.e., 70–90% of the RMS of the stiffness-dependent response, the two filters are equivalent, with an AUC of Spring 3 of 1.00 in both scenarios, because the flatter likelihood reduces the occurrence of divergent runs.

4.10.3. Model-Form Mismatch

With the ground truth integrated by a fourth-order Runge-Kutta scheme and a perturbed Bouc–Wen law, the stiffness-tracking errors of the SIR PF and the A-SIR PF are similar to those of the main campaign and not significantly different from each other. However, mainly because of one divergent healthy run, the AUC of the SIR PF is 0.85/0.86, against 1.00/1.00 for the A-SIR PF, which illustrates once more how strongly rare divergences affect the comparison of the detection performance.

4.10.4. Reduced Sensor Set

Without the accelerometer on the third mass, the errors on Springs 2 and 3 increase to about 1 N/m for both filters. The A-SIR PF is not better than the SIR PF in tracking, and it has a lower Spring-3 AUC (0.69/0.67 against 0.78/0.78), although the difference is not significant over 10 simulations.

4.10.5. Damage Location

With the damage on Spring 1, i.e., the hysteretic element, or on Spring 2, the two filters show no significant difference in tracking accuracy. In detection, the AUC of the A-SIR PF is lower in all four cases, by 0.03–0.15, and the difference is significant for the step damage on Spring 1 and for the progressive damage on Spring 2.

4.10.6. Drift Rate and Coefficient m

With a three times longer horizon, i.e., 50 s, a temperature rate of 36 °C/min and a progressive damage over 16.7 s (about six fundamental periods), the errors of both filters are larger than in the main campaign. This behaviour is due to the random-walk parameters, which wander over a longer time, and the two filters are not significantly different in the damaged scenarios over 8 simulations. Finally, with $m = 0.015$ and $m = 15$, the A-SIR PF gives Spring-3 RMSEs that are not significantly different from those of the SIR PF and of the A-SIR PF with $m = 1.5$, as expected from the invariance discussed in Section 2.3.2.

4.11. Computational Cost

The computational cost was measured on an Intel Core i7-8700K, with one filter run per core. With the implementation used for all the results, based on vectorised and compiled kernels, one time step with $N = 3 \times 10^4$ particles takes 15.1 ms for the SIR PF and 15.9 ms for the A-SIR PF (+5%). The extra cost is entirely in the weighting stage (3.8 vs. 3.1 ms). With $N = 5000$, the step times are 2.88 and 2.96 ms, and with $N = 1000$ they are 0.96 and 0.94 ms, respectively. With a plain NumPy implementation that follows the structure of the original script, the step time with $N = 3 \times 10^4$ is 43 and 54 ms, respectively. The peak memory of a filter run is 11.4 MB for both filters with $N = 3 \times 10^4$, since the A-SIR PF only stores two additional vectors of length N (0.48 MB).

The last column of Table 4 reports the mean time of one run of 3×10^4 steps in the parallel Monte Carlo campaign, with six runs at a time. One run required on average 338 s with the SIR PF, 355 s with the A-SIR PF, 349 s with the A-SIR-FK PF and 559 s with the AC-SIR PF, which propagates four additional states per particle. The SIR-ESS PF, the A-SIR-ESS PF and the DF-PF required 315, 334 and 373 s, respectively, the EKF, the UKF, the EnKF and the DF-EKF 17, 14, 15 and 21 s, and the least-action smoother 40 s. Thus, the A-SIR PF has the same complexity class as the SIR PF and a small constant overhead, as written in Section 2.3.2, whereas the Kalman-type filters are more than an order of magnitude cheaper than any particle filter with $N = 3 \times 10^4$. However, with adaptive resampling and $N = 2000$ particles, which give practically the same accuracy (Section 4.10), the SIR-ESS PF requires 21 s per run, i.e., about as much as the Kalman-type filters.

4.12. Limitations of the Present Study

The conclusions of this work must be read within the following limitations of the numerical benchmark:

- In the main campaign, the measurements are generated with the same forward-Euler discretisation, the same Bouc–Wen law and the same noise model used inside the filters. The model-mismatch study, in which the truth is integrated with a fourth-order

Runge-Kutta scheme with four sub-steps and the Bouc–Wen parameters are perturbed by $+20\%$, -20% and $+10\%$, only partially removes this favourable condition.

- The noise is Gaussian and, in the main campaign, the estimation problem is only mildly nonlinear. These conditions favour the Kalman-type filters, whereas particle filters are expected to be advantageous for non-Gaussian noise, multimodal posteriors or strongly nonlinear observation models, which are not represented here.
- All the filters use the process-noise covariance of the original implementation and standard settings. The process noise of the kinematic states, the inflation of the EnKF and the proposal distribution of the particle filters were not optimised, and other choices may change the ranking, in particular under broadband excitation.
- The least-action smoother estimates the stiffness at time t with the data up to $t + L$. In the detection analysis, the methods are compared at equal latency, i.e., the lag is added to the delays, but the tracking errors of the smoother are not directly comparable with those of a filter, which uses only the past data.
- The thermal law of the damage-factor filters is learned from healthy simulations with the same temperature profile and excitation as the test simulations, and the monitoring starts from the calibrated healthy state. In practice, the law would be learned from a healthy monitoring period, over a limited temperature range, and it would be less accurate. Thus, the margin of the damage constraint should be increased accordingly.
- The input force is assumed to be known exactly and, in the main campaign, the accelerations of all three masses are measured. Only one reduced sensor layout, i.e., masses 1 and 2, was tested.
- The harmonic input at 30 Hz lies far above the natural frequencies of the system, i.e., 0.37–0.88 Hz. The stiffness-dependent part of the acceleration response has an RMS of 0.07–0.09 $\mathrm{m/s^2}$, i.e., only 4–5 times the measurement-noise standard deviation. A 1% change of any stiffness alters the accelerations by about 3% of the noise standard deviation per sample, so that the stiffnesses are only weakly observable from individual samples. The noise study extends the standard deviation up to four times the nominal value, and a broadband excitation is considered as an informative alternative.
- With the harmonic excitation, the hysteretic variable remains small ($|r| \leq 1.4 \times 10^{-2}$ m), and the nonlinear terms of Equation (6) are at most 5×10^{-4} times the linear term. In practice, the Bouc–Wen element behaves linearly in the main case study, and it becomes strongly nonlinear only under the broadband excitation. Its parameters are weakly identifiable, and they are not damaged in any scenario. Their estimates are reported in the Supplementary Material (Figure S1).
- Damage affects a single spring with one severity per scenario, i.e., Spring 3 in the main campaign, and Spring 1 (the stiffness of the hysteretic element) or Spring 2 in the localisation study. Multiple simultaneous damages are not considered. The thermal effect on the stiffness is about 4% over the operational range (about 15% in the extended-temperature study), and within the healthy interval between the end of the baseline and the damage onset it amounts to only about 0.06 N/m. Temperature and damage evolve over seconds (tens of seconds in the slow-drift study), which is far faster than in real structures.
- The normalisation models are trained on the same temperature range in which they are tested, and extrapolation to unseen environmental conditions is not assessed.
- Dividing the likelihood by a path-dependent action is a heuristic re-weighting. Thus, the resulting weighted ensemble is not, in general, an approximation of the Bayesian posterior $p(\mathbf{z}_k|\mathbf{y}_{1:k})$, and the A-SIR uncertainty bands should not be interpreted as posterior credible intervals.
- All the results refer to a 3-DOF numerical model, and no experimental data or higher-dimensional structures have been considered.

5. Conclusions

This work compared Bayesian filters that provide the stiffness estimates of a temperature-compensated damage-detection chain for a hysteretic structure. The compared methods were: (i) sequential importance resampling particle filters with resampling at every step and with adaptive systematic resampling, (ii) the action-aided particle filter and two reformulations of the action-aided idea, (iii) tempered and Liu–West particle filters, (iv) a damage-factor particle filter, (v) the extended, unscented and ensemble Kalman filters, and (vi) a least-action fixed-lag smoother. The comparison was based on paired Monte Carlo simulations of a three-degree-of-freedom benchmark with a Bouc–Wen element, step and progressive degradation and varying temperatures. The normalisation models were trained on the estimates of each filter, the thresholds were calibrated at equal false-alarm rates, and the analysis was complemented by threshold-free ROC curves and robustness studies. The key findings of this study are as follows:

- Among the filters that use the same model, the Kalman-type filters were the most accurate in all scenarios and gave the best damage detection and localisation, at about 5% of the cost of a particle filter with 3×10^4 particles. The EKF was as accurate as the UKF, even under broadband excitation with strongly nonlinear hysteresis, while the EnKF with 200 members was slightly less accurate and, without inflation, over-confident when the data were informative.
- The resampling scheme had a larger effect on the particle filters than any of the weighting modifications considered. At every step, the weights were almost uniform, so that resampling them only added Monte Carlo noise. In contrast, resampling only when the effective sample size dropped below $N/2$, with systematic sampling, reduced the stiffness errors of the SIR PF by 35–53%, eliminated the lost runs and brought the particle filters to within 2–6% of the Kalman-type filters in the main case study. It also reduced the number of particles needed from 3×10^4 to about 2000, at which a particle filter costs about as much as a Kalman-type filter. Under broadband excitation, it outperformed the tempered likelihood, but it remained much less accurate than the Kalman-type filters.
- The action term of the A-SIR PF changes the normalised particle weights only marginally, with a total-variation distance of at most 10^{-5} per step, decreasing over time. This is because its coefficient cancels in the normalisation, and its dominant kinetic-like part is driven by measurement-noise increments common to all particles. Consistently, the A-SIR PF does not improve the stiffness tracking of the SIR PF, and its detection is worse (AUC of Spring 3 of 0.865/0.886 against 0.996/0.999). An incremental reformulation showed a modest, non-significant improvement in tracking the progressive degradation, while an action-consistent prior on the stiffness trajectories, derived from the path-integral form of the posterior, was significantly worse. The improvements reported for a single simulation in the original version of this work were within the variability between simulations, as confirmed by the control test.
- When formulated consistently with Bayesian inference, the least-action principle is an estimation principle for the trajectory, whose minimum-action path based on the measurements is the MAP path. The resulting least-action smoother, i.e., a fixed-lag RTS recursion on the EKF, reduced the stiffness errors of the EKF by 25–62% in the main case study and by 37–70% under broadband excitation. It also gave narrower and better-calibrated bands and a better localisation, at the price of a 1 s delay, which lengthens the detection delay by about 0.4–0.6 s. This is the role in which the action idea proved useful in this study.
- Prior knowledge of the structure that is not already contained in the model improves the particle filter when it enters the transition model. With a temperature–stiffness law learned from healthy data and the constraint that damage cannot increase the stiffness, the damage-factor particle filter was the most accurate filter of the study. Its errors on

the healthy springs were 18–31% lower than those of the SIR-ESS PF and of the EKF; in all simulations, its accuracy on the damaged spring was the same, and its detection delays were similar to those of the Kalman-type filters. The constraint required a small margin, because the healthy state lies on the boundary of the admissible set, and it was much less effective when imposed by projection in the EKF. Thus, hard physical constraints are where a particle representation pays off.

- Under an informative broadband excitation, the particle filters with the settings of the main case study lose track of the stiffnesses, because they cannot correct the kinematic states. Flattening or accumulating the weights mitigates the problem, whereas the Kalman-type filters remain accurate.
- The environmental normalisation is necessary when the thermal effect is large since, without it, the healthy simulations raise false alarms on all springs. With standardised inputs, the AE detected both the step and the progressive damage in the extended-temperature study, as did a conditional normalisation, i.e., the prediction of the stiffness from the temperature. Thus, both formulations are suitable, provided that the normalisation model is verified to reproduce the healthy relationship.

From these results, the following practical recommendations can be drawn for filter-based, temperature-compensated damage detection:

1. For joint state and stiffness estimation with mildly nonlinear models and Gaussian noise, start from an EKF or a UKF, which were the most accurate filters with the same model and among the cheapest in this study. The UKF does not require the derivation of Jacobians.
2. When a delay of the order of a second is acceptable, as for the slow degradation processes targeted by SHM, estimate the trajectory over a fixed lag (least-action smoothing) rather than filtering. With a 1 s lag, the least-action (fixed-lag RTS) smoother reduced the errors of the EKF by 25–62% and localised the damage better, while its alarms were raised about 0.4–0.6 s later.
3. When a particle filter is required, e.g., for non-Gaussian noise or multimodal posteriors, resample adaptively (for instance, systematic resampling when the ESS drops below $N/2$) rather than at every step, and verify the number of particles. With adaptive resampling, about 2000 particles were sufficient in this study, with no lost run already with 1000, whereas with resampling at every step even 3×10^4 particles did not prevent occasional divergences.
4. Encode the prior knowledge of the structure that is not already in the model, such as environmental laws calibrated on healthy data and the admissible range of the damage, in the transition model of a particle filter, with a small margin rather than a hard bound at the healthy state. In contrast, re-weighting the particles with quantities computed from their own states adds no information.
5. When the data are informative and the particles cannot correct the kinematic states, temper the likelihood or redesign the process noise or the proposal distribution. Re-weighting the particles according to the history of their residuals, as in the action-aided weighting, did not help.
6. For the environmental normalisation, standardise the inputs of the normalisation model, verify on held-out healthy data that it reproduces the relationship between the environmental variables and the damage-sensitive feature, e.g., through the explained variance of the reconstruction, and test the normalisation over environmental ranges large enough for the thermal effect to exceed the estimation noise.
7. Compare filters on repeated, paired simulations with common random numbers, at equal false-alarm rates and with threshold-free metrics. In addition, include simple reference filters and a control test, and report the variability between runs.

However, the conclusions of this study are limited to a numerical benchmark with Gaussian noise. Thus, future work will address experimental data and higher-dimensional models with non-Gaussian noise and model-form uncertainty, i.e., the conditions in which particle filters are expected to be advantageous. In addition, particle filters for informative excitations may be improved with appropriately designed kinematic process noise, Kalman-type proposal distributions or Rao–Blackwellisation of the kinematic states. Other directions include jump-aware transition models or change-point detection for abrupt damage, e.g., damage-factor particle filters in which each spring remains exactly healthy until a damage onset, environmental laws learned online and over limited temperature ranges, and physics-based priors on the parameter trajectories in which the least-action idea is encoded consistently with the Bayesian formulation.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/buildings16224589/s1, Figure S1: Estimates of the damping coefficient and of the Bouc–Wen parameters in the main campaign. The left column refers to the progressive damage scenario and the right column to the healthy scenario. β stands for the proportional damping coefficient, while β_{BW}, γ and n are the Bouc–Wen parameters of Equation (6). Solid lines are the medians of the estimates of the SIR PF (blue) and of the A-SIR PF (red) over the 50 simulations, shaded areas are the 10–90% ranges, and black lines are the true values.

Author Contributions: Conceptualization, L.P. and L.L.; methodology, L.P. and L.L.; software, L.P.; validation, L.P. and L.L.; formal analysis, L.P.; investigation, L.P.; resources, L.L.; data curation, L.P. and L.L.; writing—original draft preparation, L.P.; writing—review and editing, L.P. and L.L.; visualization, L.P.; supervision, L.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AE	Autoencoders
AUC	Area Under the receiver-operating-characteristic Curve
PF	Particle Filter
SIR PF	Sequential Importance Resampling Particle Filter
SIR-ESS PF	SIR PF with adaptive (ESS-triggered) systematic resampling
A-SIR PF	Action-aided Sequential Importance Resampling Particle Filter
A-SIR-ESS PF	A-SIR PF with adaptive (ESS-triggered) systematic resampling
A-SIR-FK PF	A-SIR Particle Filter with Feynman–Kac (incremental) action weighting
AC-SIR PF	Action-consistent SIR Particle Filter
DF-PF	Damage-factor Particle Filter
DF-EKF	Damage-factor Extended Kalman Filter
EKF	Extended Kalman Filter
UKF	Unscented Kalman Filter
EnKF	Ensemble Kalman Filter
ESS	Effective Sample Size
MDOF	Multiple Degrees Of Freedom
MLP	Multilayer Perceptron
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
SE	Squared Error

SHM Structural Health Monitoring

Appendix A. Detailed Results

Tables A1–A3 report the per-spring tracking metrics of all the filters of the main campaign. Tables A4 and A5 report, for every additional study, the RMSE of the damaged spring obtained with the SIR PF and with the compared method on the same simulations.

Table A1. Stiffness-tracking accuracy in the Progressive Damage scenario. Mean ± standard deviation over the Monte Carlo simulations ($N = 3 \times 10^4$ particles, $N_e = 200$ members). RMSE and MAE are computed over [2.78, 16.67] s. The 95% interval width and coverage refer to the 2.5–97.5% (weighted) quantiles of the particles, and to the mean ±1.96 standard deviations for the Kalman-type filters. ΔRMSE is the mean paired difference with respect to the SIR PF (negative values mean lower error), with the Holm-corrected Wilcoxon signed-rank p-value in brackets.

Method	Spring	RMSE [N/m]	MAE [N/m]	95% Width [N/m]	Coverage [-]	ΔRMSE (p)
SIR PF ($n = 50$)	Spring 1	0.155 ± 0.064	0.123 ± 0.054	0.450 ± 0.034	0.84 ± 0.14	–
SIR PF	Spring 2	0.206 ± 0.267	0.146 ± 0.172	0.420 ± 0.027	0.84 ± 0.14	–
SIR PF	Spring 3	0.385 ± 0.221	0.247 ± 0.126	0.430 ± 0.033	0.67 ± 0.09	–
SIR PF	Spring 4	0.203 ± 0.202	0.151 ± 0.134	0.449 ± 0.031	0.82 ± 0.14	–
A-SIR PF ($n = 50$)	Spring 1	0.207 ± 0.284	0.162 ± 0.210	0.442 ± 0.037	0.80 ± 0.14	+0.052 (0.296)
A-SIR PF	Spring 2	0.282 ± 0.747	0.204 ± 0.532	0.420 ± 0.035	0.82 ± 0.14	+0.076 (1.000)
A-SIR PF	Spring 3	0.386 ± 0.140	0.253 ± 0.100	0.426 ± 0.034	0.65 ± 0.10	+0.002 (1.000)
A-SIR PF	Spring 4	0.193 ± 0.122	0.149 ± 0.087	0.443 ± 0.036	0.80 ± 0.14	−0.010 (1.000)
A-SIR-FK PF ($n = 50$)	Spring 1	0.162 ± 0.048	0.128 ± 0.039	0.440 ± 0.026	0.82 ± 0.12	+0.007 (1.000)
A-SIR-FK PF	Spring 2	0.171 ± 0.083	0.131 ± 0.059	0.416 ± 0.020	0.82 ± 0.14	−0.034 (1.000)
A-SIR-FK PF	Spring 3	0.323 ± 0.124	0.219 ± 0.071	0.426 ± 0.024	0.66 ± 0.09	−0.062 (0.296)
A-SIR-FK PF	Spring 4	0.183 ± 0.061	0.143 ± 0.043	0.443 ± 0.023	0.79 ± 0.12	−0.020 (1.000)
AC-SIR PF ($n = 50$)	Spring 1	0.266 ± 0.325	0.202 ± 0.223	0.351 ± 0.055	0.67 ± 0.22	+0.111 (0.049)
AC-SIR PF	Spring 2	0.435 ± 0.841	0.304 ± 0.606	0.331 ± 0.050	0.67 ± 0.21	+0.229 (0.031)
AC-SIR PF	Spring 3	0.553 ± 0.516	0.369 ± 0.362	0.332 ± 0.052	0.53 ± 0.14	+0.168 (0.022)
AC-SIR PF	Spring 4	0.279 ± 0.278	0.210 ± 0.204	0.349 ± 0.055	0.65 ± 0.20	+0.076 (0.094)
SIR-ESS PF ($n = 50$)	Spring 1	0.091 ± 0.015	0.072 ± 0.012	0.562 ± 0.005	0.99 ± 0.02	−0.064 (<0.001)
SIR-ESS PF	Spring 2	0.081 ± 0.012	0.064 ± 0.009	0.518 ± 0.006	1.00 ± 0.01	−0.125 (<0.001)
SIR-ESS PF	Spring 3	0.249 ± 0.027	0.164 ± 0.016	0.534 ± 0.005	0.77 ± 0.05	−0.136 (<0.001)
SIR-ESS PF	Spring 4	0.093 ± 0.015	0.073 ± 0.013	0.563 ± 0.006	0.99 ± 0.01	−0.110 (<0.001)
A-SIR-ESS PF ($n = 50$)	Spring 1	0.091 ± 0.014	0.072 ± 0.010	0.562 ± 0.007	0.99 ± 0.02	−0.064 (<0.001)
A-SIR-ESS PF	Spring 2	0.081 ± 0.014	0.064 ± 0.010	0.519 ± 0.005	1.00 ± 0.01	−0.125 (<0.001)
A-SIR-ESS PF	Spring 3	0.245 ± 0.027	0.162 ± 0.017	0.536 ± 0.006	0.77 ± 0.05	−0.140 (<0.001)
A-SIR-ESS PF	Spring 4	0.093 ± 0.016	0.073 ± 0.013	0.562 ± 0.005	0.99 ± 0.01	−0.110 (<0.001)
DF-PF ($n = 50$)	Spring 1	0.072 ± 0.014	0.054 ± 0.010	0.433 ± 0.016	1.00 ± 0.01	−0.083 (<0.001)
DF-PF	Spring 2	0.060 ± 0.012	0.047 ± 0.009	0.403 ± 0.014	1.00 ± 0.00	−0.146 (<0.001)
DF-PF	Spring 3	0.248 ± 0.024	0.156 ± 0.015	0.460 ± 0.012	0.77 ± 0.04	−0.137 (<0.001)
DF-PF	Spring 4	0.076 ± 0.018	0.057 ± 0.014	0.437 ± 0.017	0.99 ± 0.01	−0.127 (<0.001)
EKF ($n = 50$)	Spring 1	0.088 ± 0.012	0.070 ± 0.010	0.570 ± 0.000	1.00 ± 0.01	−0.068 (<0.001)
EKF	Spring 2	0.078 ± 0.012	0.062 ± 0.009	0.524 ± 0.000	1.00 ± 0.00	−0.128 (<0.001)
EKF	Spring 3	0.243 ± 0.022	0.159 ± 0.015	0.544 ± 0.001	0.78 ± 0.04	−0.142 (<0.001)
EKF	Spring 4	0.090 ± 0.015	0.071 ± 0.013	0.571 ± 0.001	1.00 ± 0.01	−0.113 (<0.001)
UKF ($n = 50$)	Spring 1	0.087 ± 0.012	0.069 ± 0.010	0.570 ± 0.000	1.00 ± 0.01	−0.068 (<0.001)
UKF	Spring 2	0.077 ± 0.012	0.062 ± 0.009	0.524 ± 0.000	1.00 ± 0.00	−0.128 (<0.001)
UKF	Spring 3	0.243 ± 0.022	0.159 ± 0.014	0.544 ± 0.001	0.78 ± 0.04	−0.142 (<0.001)
UKF	Spring 4	0.089 ± 0.015	0.071 ± 0.013	0.571 ± 0.001	1.00 ± 0.01	−0.114 (<0.001)
EnKF ($n = 50$)	Spring 1	0.091 ± 0.013	0.073 ± 0.011	0.556 ± 0.006	1.00 ± 0.01	−0.064 (<0.001)
EnKF	Spring 2	0.081 ± 0.012	0.064 ± 0.010	0.513 ± 0.006	1.00 ± 0.01	−0.125 (<0.001)
EnKF	Spring 3	0.244 ± 0.022	0.161 ± 0.015	0.531 ± 0.006	0.77 ± 0.05	−0.140 (<0.001)
EnKF	Spring 4	0.093 ± 0.015	0.073 ± 0.012	0.557 ± 0.006	0.99 ± 0.01	−0.110 (<0.001)
DF-EKF ($n = 50$)	Spring 1	0.059 ± 0.009	0.046 ± 0.008	0.358 ± 0.000	0.99 ± 0.01	−0.096 (<0.001)
DF-EKF	Spring 2	0.049 ± 0.008	0.039 ± 0.006	0.320 ± 0.000	1.00 ± 0.01	−0.156 (<0.001)
DF-EKF	Spring 3	0.322 ± 0.015	0.198 ± 0.010	0.328 ± 0.000	0.63 ± 0.01	−0.063 (0.896)
DF-EKF	Spring 4	0.069 ± 0.013	0.053 ± 0.011	0.355 ± 0.000	0.97 ± 0.04	−0.134 (<0.001)
LAS ($L = 1$ s) ($n = 50$)	Spring 1	0.063 ± 0.013	0.050 ± 0.011	0.356 ± 0.000	0.99 ± 0.02	−0.092 (<0.001)
LAS ($L = 1$ s)	Spring 2	0.052 ± 0.010	0.041 ± 0.008	0.311 ± 0.000	1.00 ± 0.01	−0.153 (<0.001)
LAS ($L = 1$ s)	Spring 3	0.091 ± 0.017	0.059 ± 0.011	0.334 ± 0.001	0.96 ± 0.03	−0.294 (<0.001)
LAS ($L = 1$ s)	Spring 4	0.062 ± 0.013	0.050 ± 0.011	0.357 ± 0.000	0.99 ± 0.01	−0.141 (<0.001)

Table A2. Stiffness-tracking accuracy in the Step Damage scenario. Mean ± standard deviation over the Monte Carlo simulations ($N = 3 \times 10^4$ particles, $N_e = 200$ members). RMSE and MAE are computed over [2.78, 16.67] s. The 95% interval width and coverage refer to the 2.5–97.5% (weighted) quantiles of the particles, and to the mean ±1.96 standard deviations for the Kalman-type filters. ΔRMSE is the mean paired difference with respect to the SIR PF (negative values mean lower error), with the Holm-corrected Wilcoxon signed-rank p-value in brackets.

Method	Spring	RMSE [N/m]	MAE [N/m]	95% Width [N/m]	Coverage [-]	ΔRMSE (p)
SIR PF ($n = 50$)	Spring 1	0.163 ± 0.090	0.130 ± 0.077	0.448 ± 0.034	0.84 ± 0.15	–
SIR PF	Spring 2	0.190 ± 0.297	0.136 ± 0.180	0.420 ± 0.028	0.85 ± 0.14	–
SIR PF	Spring 3	0.214 ± 0.192	0.153 ± 0.126	0.434 ± 0.032	0.81 ± 0.13	–
SIR PF	Spring 4	0.194 ± 0.156	0.150 ± 0.108	0.450 ± 0.030	0.80 ± 0.15	–
A-SIR PF ($n = 50$)	Spring 1	0.205 ± 0.286	0.159 ± 0.210	0.439 ± 0.034	0.81 ± 0.15	+0.043 (0.320)
A-SIR PF	Spring 2	0.250 ± 0.731	0.188 ± 0.524	0.419 ± 0.033	0.85 ± 0.15	+0.060 (1.000)
A-SIR PF	Spring 3	0.207 ± 0.110	0.148 ± 0.087	0.429 ± 0.033	0.80 ± 0.14	−0.007 (1.000)
A-SIR PF	Spring 4	0.196 ± 0.177	0.152 ± 0.126	0.446 ± 0.037	0.79 ± 0.14	+0.002 (1.000)
A-SIR-FK PF ($n = 50$)	Spring 1	0.167 ± 0.054	0.131 ± 0.040	0.434 ± 0.024	0.81 ± 0.12	+0.004 (1.000)
A-SIR-FK PF	Spring 2	0.180 ± 0.080	0.135 ± 0.055	0.414 ± 0.019	0.81 ± 0.14	−0.009 (0.320)
A-SIR-FK PF	Spring 3	0.207 ± 0.089	0.147 ± 0.059	0.426 ± 0.026	0.79 ± 0.12	−0.007 (1.000)
A-SIR-FK PF	Spring 4	0.193 ± 0.088	0.149 ± 0.055	0.442 ± 0.024	0.77 ± 0.14	−0.001 (1.000)
AC-SIR PF ($n = 50$)	Spring 1	0.287 ± 0.336	0.214 ± 0.233	0.348 ± 0.056	0.66 ± 0.21	+0.125 (0.010)
AC-SIR PF	Spring 2	0.403 ± 0.809	0.288 ± 0.584	0.328 ± 0.051	0.68 ± 0.22	+0.213 (0.103)
AC-SIR PF	Spring 3	0.359 ± 0.581	0.257 ± 0.389	0.337 ± 0.055	0.66 ± 0.19	+0.145 (0.072)
AC-SIR PF	Spring 4	0.275 ± 0.286	0.212 ± 0.203	0.351 ± 0.057	0.64 ± 0.20	+0.081 (0.248)
SIR-ESS PF ($n = 50$)	Spring 1	0.102 ± 0.020	0.078 ± 0.013	0.546 ± 0.009	0.98 ± 0.02	−0.061 (<0.001)
SIR-ESS PF	Spring 2	0.083 ± 0.014	0.065 ± 0.010	0.509 ± 0.006	0.99 ± 0.01	−0.107 (<0.001)
SIR-ESS PF	Spring 3	0.136 ± 0.015	0.082 ± 0.010	0.535 ± 0.005	0.95 ± 0.02	−0.078 (<0.001)
SIR-ESS PF	Spring 4	0.096 ± 0.016	0.074 ± 0.013	0.561 ± 0.006	0.98 ± 0.02	−0.098 (<0.001)
A-SIR-ESS PF ($n = 50$)	Spring 1	0.097 ± 0.018	0.076 ± 0.013	0.547 ± 0.007	0.99 ± 0.02	−0.065 (<0.001)
A-SIR-ESS PF	Spring 2	0.084 ± 0.016	0.066 ± 0.011	0.509 ± 0.005	0.99 ± 0.01	−0.106 (<0.001)
A-SIR-ESS PF	Spring 3	0.134 ± 0.012	0.081 ± 0.009	0.536 ± 0.005	0.95 ± 0.01	−0.080 (<0.001)
A-SIR-ESS PF	Spring 4	0.093 ± 0.015	0.072 ± 0.012	0.559 ± 0.006	0.98 ± 0.02	−0.101 (<0.001)
DF-PF ($n = 50$)	Spring 1	0.062 ± 0.012	0.047 ± 0.008	0.407 ± 0.018	1.00 ± 0.00	−0.100 (<0.001)
DF-PF	Spring 2	0.062 ± 0.013	0.048 ± 0.009	0.399 ± 0.015	1.00 ± 0.00	−0.127 (<0.001)
DF-PF	Spring 3	0.126 ± 0.012	0.071 ± 0.009	0.460 ± 0.012	0.95 ± 0.01	−0.088 (<0.001)
DF-PF	Spring 4	0.070 ± 0.017	0.053 ± 0.013	0.424 ± 0.018	0.99 ± 0.02	−0.124 (<0.001)
EKF ($n = 50$)	Spring 1	0.087 ± 0.011	0.070 ± 0.009	0.569 ± 0.000	1.00 ± 0.00	−0.076 (<0.001)
EKF	Spring 2	0.079 ± 0.012	0.063 ± 0.010	0.521 ± 0.000	1.00 ± 0.00	−0.111 (<0.001)
EKF	Spring 3	0.127 ± 0.010	0.076 ± 0.008	0.544 ± 0.001	0.96 ± 0.01	−0.087 (<0.001)
EKF	Spring 4	0.088 ± 0.012	0.067 ± 0.010	0.571 ± 0.001	0.99 ± 0.01	−0.106 (<0.001)
UKF ($n = 50$)	Spring 1	0.087 ± 0.011	0.070 ± 0.009	0.569 ± 0.000	1.00 ± 0.00	−0.076 (<0.001)
UKF	Spring 2	0.079 ± 0.012	0.062 ± 0.010	0.521 ± 0.000	1.00 ± 0.00	−0.111 (<0.001)
UKF	Spring 3	0.128 ± 0.010	0.076 ± 0.008	0.544 ± 0.001	0.96 ± 0.01	−0.086 (<0.001)
UKF	Spring 4	0.088 ± 0.012	0.067 ± 0.010	0.571 ± 0.001	0.99 ± 0.01	−0.106 (<0.001)
EnKF ($n = 50$)	Spring 1	0.091 ± 0.015	0.073 ± 0.012	0.555 ± 0.006	1.00 ± 0.01	−0.071 (<0.001)
EnKF	Spring 2	0.082 ± 0.013	0.065 ± 0.010	0.509 ± 0.006	1.00 ± 0.01	−0.108 (<0.001)
EnKF	Spring 3	0.131 ± 0.012	0.078 ± 0.009	0.530 ± 0.006	0.96 ± 0.01	−0.083 (<0.001)
EnKF	Spring 4	0.091 ± 0.013	0.070 ± 0.011	0.558 ± 0.006	0.99 ± 0.01	−0.103 (<0.001)
DF-EKF ($n = 50$)	Spring 1	0.058 ± 0.008	0.047 ± 0.007	0.358 ± 0.000	1.00 ± 0.00	−0.104 (<0.001)
DF-EKF	Spring 2	0.054 ± 0.008	0.043 ± 0.007	0.318 ± 0.000	1.00 ± 0.01	−0.135 (<0.001)
DF-EKF	Spring 3	0.148 ± 0.008	0.067 ± 0.005	0.327 ± 0.000	0.92 ± 0.01	−0.066 (<0.001)
DF-EKF	Spring 4	0.053 ± 0.010	0.042 ± 0.008	0.355 ± 0.000	1.00 ± 0.01	−0.141 (<0.001)
LAS ($L = 1$ s) ($n = 50$)	Spring 1	0.068 ± 0.013	0.054 ± 0.011	0.354 ± 0.000	0.99 ± 0.02	−0.094 (<0.001)
LAS ($L = 1$ s)	Spring 2	0.054 ± 0.009	0.042 ± 0.007	0.307 ± 0.000	0.98 ± 0.02	−0.135 (<0.001)
LAS ($L = 1$ s)	Spring 3	0.086 ± 0.008	0.055 ± 0.007	0.336 ± 0.001	0.94 ± 0.01	−0.128 (<0.001)
LAS ($L = 1$ s)	Spring 4	0.069 ± 0.012	0.052 ± 0.010	0.357 ± 0.000	0.98 ± 0.02	−0.125 (<0.001)

Table A3. Stiffness-tracking accuracy in the healthy structure scenario. Mean ± standard deviation over the Monte Carlo simulations ($N = 3 \times 10^4$ particles, $N_e = 200$ members). RMSE and MAE are computed over [2.78, 16.67] s. The 95% interval width and coverage refer to the 2.5–97.5% (weighted) quantiles of the particles, and to the mean ±1.96 standard deviations for the Kalman-type filters. ΔRMSE is the mean paired difference with respect to the SIR PF (negative values mean lower error), with the Holm-corrected Wilcoxon signed-rank p-value in brackets.

Method	Spring	RMSE [N/m]	MAE [N/m]	95% Width [N/m]	Coverage [-]	ΔRMSE (p)
SIR PF ($n = 50$)	Spring 1	0.149 ± 0.061	0.121 ± 0.050	0.462 ± 0.023	0.87 ± 0.14	–
SIR PF	Spring 2	0.134 ± 0.059	0.105 ± 0.045	0.429 ± 0.021	0.89 ± 0.12	–
SIR PF	Spring 3	0.139 ± 0.056	0.106 ± 0.039	0.446 ± 0.025	0.89 ± 0.12	–
SIR PF	Spring 4	0.146 ± 0.054	0.118 ± 0.045	0.469 ± 0.025	0.88 ± 0.13	–
A-SIR PF ($n = 50$)	Spring 1	0.209 ± 0.244	0.165 ± 0.195	0.449 ± 0.064	0.84 ± 0.23	+0.060 (1.000)
A-SIR PF	Spring 2	0.264 ± 0.484	0.196 ± 0.350	0.420 ± 0.056	0.84 ± 0.24	+0.131 (1.000)
A-SIR PF	Spring 3	0.299 ± 0.722	0.217 ± 0.461	0.434 ± 0.061	0.85 ± 0.23	+0.161 (1.000)
A-SIR PF	Spring 4	0.206 ± 0.231	0.160 ± 0.166	0.450 ± 0.061	0.82 ± 0.22	+0.060 (1.000)
A-SIR-FK PF ($n = 50$)	Spring 1	0.209 ± 0.417	0.160 ± 0.293	0.447 ± 0.041	0.85 ± 0.16	+0.060 (1.000)
A-SIR-FK PF	Spring 2	0.275 ± 0.938	0.209 ± 0.702	0.425 ± 0.041	0.87 ± 0.18	+0.142 (1.000)
A-SIR-FK PF	Spring 3	0.189 ± 0.393	0.148 ± 0.302	0.432 ± 0.039	0.87 ± 0.15	+0.050 (1.000)
A-SIR-FK PF	Spring 4	0.172 ± 0.107	0.136 ± 0.084	0.450 ± 0.042	0.83 ± 0.18	+0.026 (1.000)
AC-SIR PF ($n = 50$)	Spring 1	0.210 ± 0.194	0.168 ± 0.153	0.367 ± 0.067	0.72 ± 0.25	+0.061 (1.000)
AC-SIR PF	Spring 2	0.293 ± 0.554	0.225 ± 0.445	0.341 ± 0.057	0.72 ± 0.25	+0.159 (0.283)
AC-SIR PF	Spring 3	0.373 ± 0.759	0.262 ± 0.488	0.349 ± 0.062	0.75 ± 0.27	+0.234 (1.000)
AC-SIR PF	Spring 4	0.256 ± 0.318	0.196 ± 0.224	0.369 ± 0.066	0.72 ± 0.24	+0.110 (1.000)
SIR-ESS PF ($n = 50$)	Spring 1	0.084 ± 0.009	0.067 ± 0.008	0.563 ± 0.004	1.00 ± 0.00	−0.065 (<0.001)
SIR-ESS PF	Spring 2	0.074 ± 0.011	0.058 ± 0.008	0.518 ± 0.004	1.00 ± 0.00	−0.060 (<0.001)
SIR-ESS PF	Spring 3	0.074 ± 0.012	0.058 ± 0.009	0.540 ± 0.005	1.00 ± 0.00	−0.065 (<0.001)
SIR-ESS PF	Spring 4	0.085 ± 0.017	0.068 ± 0.014	0.572 ± 0.004	1.00 ± 0.01	−0.061 (<0.001)
A-SIR-ESS PF ($n = 50$)	Spring 1	0.084 ± 0.010	0.067 ± 0.008	0.563 ± 0.006	1.00 ± 0.00	−0.065 (<0.001)
A-SIR-ESS PF	Spring 2	0.074 ± 0.010	0.058 ± 0.007	0.517 ± 0.003	1.00 ± 0.00	−0.060 (<0.001)
A-SIR-ESS PF	Spring 3	0.074 ± 0.012	0.059 ± 0.009	0.540 ± 0.005	1.00 ± 0.00	−0.065 (<0.001)
A-SIR-ESS PF	Spring 4	0.085 ± 0.016	0.068 ± 0.013	0.571 ± 0.005	1.00 ± 0.01	−0.061 (<0.001)
DF-PF ($n = 50$)	Spring 1	0.064 ± 0.012	0.048 ± 0.009	0.426 ± 0.017	1.00 ± 0.00	−0.085 (<0.001)
DF-PF	Spring 2	0.055 ± 0.012	0.042 ± 0.008	0.402 ± 0.015	1.00 ± 0.00	−0.078 (<0.001)
DF-PF	Spring 3	0.057 ± 0.011	0.044 ± 0.008	0.412 ± 0.014	1.00 ± 0.00	−0.082 (<0.001)
DF-PF	Spring 4	0.064 ± 0.016	0.049 ± 0.011	0.431 ± 0.019	1.00 ± 0.00	−0.082 (<0.001)
EKF ($n = 50$)	Spring 1	0.082 ± 0.010	0.065 ± 0.008	0.566 ± 0.000	1.00 ± 0.00	−0.067 (<0.001)
EKF	Spring 2	0.072 ± 0.010	0.057 ± 0.008	0.520 ± 0.000	1.00 ± 0.00	−0.061 (<0.001)
EKF	Spring 3	0.072 ± 0.011	0.057 ± 0.008	0.542 ± 0.001	1.00 ± 0.00	−0.067 (<0.001)
EKF	Spring 4	0.083 ± 0.015	0.066 ± 0.012	0.574 ± 0.001	1.00 ± 0.01	−0.063 (<0.001)
UKF ($n = 50$)	Spring 1	0.082 ± 0.010	0.065 ± 0.008	0.566 ± 0.000	1.00 ± 0.00	−0.067 (<0.001)
UKF	Spring 2	0.072 ± 0.010	0.057 ± 0.008	0.520 ± 0.000	1.00 ± 0.00	−0.062 (<0.001)
UKF	Spring 3	0.072 ± 0.010	0.057 ± 0.008	0.542 ± 0.001	1.00 ± 0.00	−0.067 (<0.001)
UKF	Spring 4	0.083 ± 0.015	0.066 ± 0.012	0.574 ± 0.001	1.00 ± 0.01	−0.063 (<0.001)
EnKF ($n = 50$)	Spring 1	0.087 ± 0.011	0.069 ± 0.009	0.551 ± 0.006	1.00 ± 0.01	−0.062 (<0.001)
EnKF	Spring 2	0.074 ± 0.010	0.059 ± 0.008	0.509 ± 0.005	1.00 ± 0.00	−0.059 (<0.001)
EnKF	Spring 3	0.075 ± 0.012	0.060 ± 0.010	0.530 ± 0.006	1.00 ± 0.00	−0.064 (<0.001)
EnKF	Spring 4	0.086 ± 0.018	0.069 ± 0.013	0.561 ± 0.007	1.00 ± 0.01	−0.060 (<0.001)
DF-EKF ($n = 50$)	Spring 1	0.050 ± 0.007	0.040 ± 0.006	0.356 ± 0.000	1.00 ± 0.00	−0.099 (<0.001)
DF-EKF	Spring 2	0.042 ± 0.006	0.033 ± 0.005	0.317 ± 0.000	1.00 ± 0.00	−0.092 (<0.001)
DF-EKF	Spring 3	0.041 ± 0.006	0.032 ± 0.005	0.326 ± 0.000	1.00 ± 0.00	−0.098 (<0.001)
DF-EKF	Spring 4	0.050 ± 0.010	0.040 ± 0.008	0.357 ± 0.000	1.00 ± 0.00	−0.096 (<0.001)
LAS ($L = 1$ s) ($n = 50$)	Spring 1	0.057 ± 0.010	0.046 ± 0.008	0.351 ± 0.000	1.00 ± 0.01	−0.092 (<0.001)
LAS ($L = 1$ s)	Spring 2	0.047 ± 0.010	0.036 ± 0.007	0.308 ± 0.000	1.00 ± 0.00	−0.087 (<0.001)
LAS ($L = 1$ s)	Spring 3	0.048 ± 0.010	0.038 ± 0.007	0.330 ± 0.001	1.00 ± 0.00	−0.091 (<0.001)
LAS ($L = 1$ s)	Spring 4	0.057 ± 0.013	0.045 ± 0.010	0.362 ± 0.000	1.00 ± 0.01	−0.089 (<0.001)

Table A4. Summary of the main campaign, of the action-based variants and of the comparators (main case study and broadband excitation). For each study, the RMSE of the damaged spring is reported for the SIR PF and for the compared method on the same simulations (Spring 3 unless otherwise stated, and Spring 3 also for the healthy scenario). The mean paired difference (method − SIR PF) is reported with its 95% confidence interval, together with the Wilcoxon p-value, Holm-corrected over all the comparisons of Tables A4 and A5.

Study	Scenario	Method	n	SIR [N/m]	Method [N/m]	Difference [N/m]	p
Test control (other filter seed)	Progressive	SIR PF	30	0.376	0.541	+0.164 [−0.117, +0.446]	1.000
	Step	SIR PF	30	0.188	0.394	+0.206 [−0.084, +0.495]	1.000
	Healthy	SIR PF	30	0.143	0.140	−0.003 [−0.049, +0.043]	1.000
Ablation: T only	Progressive	A-SIR PF, T only	30	0.376	0.313	−0.064 [−0.126, −0.002]	1.000
	Step	A-SIR PF, T only	30	0.188	0.191	+0.003 [−0.025, +0.031]	1.000
Ablation: V only	Progressive	A-SIR PF, V only	10	0.414	2.695	+2.281 [+0.870, +3.691]	0.137
	Step	A-SIR PF, V only	10	0.203	2.611	+2.407 [+0.903, +3.912]	0.137
Tempered SIR (0.5)	Progressive	Tempered SIR PF	10	0.414	0.362	−0.053 [−0.110, +0.004]	1.000
	Step	Tempered SIR PF	10	0.203	0.177	−0.026 [−0.100, +0.047]	1.000
Liu–West SIR	Progressive	Liu–West SIR PF	10	0.414	0.380	−0.034 [−0.138, +0.070]	1.000
	Step	Liu–West SIR PF	10	0.203	0.243	+0.040 [−0.047, +0.127]	1.000
Control: FK V only (0.3)	Progressive	A-SIR-FK PF (T+V , $\lambda = 0.3$)	10	0.414	0.292	−0.122 [−0.253, +0.009]	1.000
	Step	A-SIR-FK PF (T+V, $\lambda = 0.3$)	10	0.203	0.191	−0.012 [−0.084, +0.060]	1.000
	Healthy	A-SIR-FK PF (T+V, $\lambda = 0.3$)	10	0.177	0.165	−0.011 [−0.047, +0.024]	1.000

Table A5. Summary of the robustness and sensitivity studies (SIR PF vs. A-SIR PF unless otherwise stated). For each study, the RMSE of the damaged spring is reported for the SIR PF and for the compared method on the same simulations (Spring 3 unless otherwise stated, and Spring 3 also for the healthy scenario). The mean paired difference (method − SIR PF) is reported with its 95% confidence interval, together with the Wilcoxon p-value, Holm-corrected over all the comparisons of Tables A4 and A5.

Study	Scenario	Method	n	SIR [N/m]	Method [N/m]	Difference [N/m]	p
Noise sigma x2	Progressive	A-SIR PF	10	0.393	0.357	−0.036 [−0.124, +0.052]	1.000
	Step	A-SIR PF	10	0.233	0.194	−0.038 [−0.053, −0.024]	0.137
Noise sigma x4	Progressive	A-SIR PF	10	0.512	0.489	−0.023 [−0.091, +0.045]	1.000
	Step	A-SIR PF	10	0.219	0.223	+0.003 [−0.028, +0.034]	1.000
	Healthy	A-SIR PF	10	0.159	0.145	−0.014 [−0.053, +0.024]	1.000
Noise sigma x4, FK	Progressive	A-SIR-FK PF	10	0.512	0.532	+0.021 [−0.069, +0.110]	1.000
	Step	A-SIR-FK PF	10	0.219	0.268	+0.048 [−0.005, +0.101]	1.000
	Healthy	A-SIR-FK PF	10	0.159	0.168	+0.009 [−0.038, +0.057]	1.000
Model mismatch	Progressive	A-SIR PF	10	0.360	0.366	+0.006 [−0.068, +0.080]	1.000
	Step	A-SIR PF	10	0.196	0.179	−0.017 [−0.066, +0.031]	1.000
	Healthy	A-SIR PF	10	0.533	0.129	−0.404 [−1.293, +0.486]	1.000
Model mismatch, FK	Progressive	A-SIR-FK PF	10	0.360	0.389	+0.029 [−0.043, +0.101]	1.000
	Step	A-SIR-FK PF	10	0.196	0.217	+0.022 [−0.031, +0.075]	1.000
	Healthy	A-SIR-FK PF	10	0.533	0.260	−0.274 [−1.257, +0.710]	1.000
Sensors 1-2 only	Progressive	A-SIR PF	10	1.184	1.578	+0.394 [−0.839, +1.628]	1.000
	Step	A-SIR PF	10	1.203	1.605	+0.403 [−0.923, +1.728]	1.000
	Healthy	A-SIR PF	10	0.292	0.964	+0.672 [−0.650, +1.994]	1.000
Sensors 1-2 only, FK	Progressive	A-SIR-FK PF	10	1.184	0.816	−0.368 [−1.347, +0.610]	1.000
	Step	A-SIR-FK PF	10	1.203	0.798	−0.405 [−1.498, +0.688]	1.000
	Healthy	A-SIR-FK PF	10	0.292	0.952	+0.659 [−0.402, +1.721]	1.000

Table A5. *Cont.*

Study	Scenario	Method	*n*	SIR [N/m]	Method [N/m]	Difference [N/m]	*p*
Damage on Spring 1	Progressive	A-SIR PF	10	0.300	0.266	−0.033 [−0.120, +0.053]	1.000
	Step	A-SIR PF	10	0.314	0.276	−0.038 [−0.110, +0.035]	1.000
Damage on Spring 2	Progressive	A-SIR PF	10	0.231	0.263	+0.032 [−0.036, +0.100]	1.000
	Step	A-SIR PF	10	0.311	0.286	−0.025 [−0.109, +0.058]	1.000
Extended temperature	Progressive	A-SIR PF	10	0.334	0.617	+0.284 [−0.135, +0.702]	1.000
	Step	A-SIR PF	10	0.195	0.396	+0.201 [−0.102, +0.505]	1.000
	Healthy	A-SIR PF	10	0.150	0.158	+0.008 [−0.027, +0.044]	1.000
Slow drift (3x)	Progressive	A-SIR PF	8	0.390	0.722	+0.332 [−0.764, +1.429]	1.000
	Step	A-SIR PF	8	0.278	0.633	+0.355 [−0.706, +1.416]	1.000
	Healthy	A-SIR PF	8	1.140	1.205	+0.065 [−0.036, +0.166]	1.000
Coefficient m = 0.015	Progressive	A-SIR PF	10	0.414	0.524	+0.109 [−0.240, +0.459]	1.000
Coefficient m = 15	Progressive	A-SIR PF	10	0.414	0.380	−0.035 [−0.199, +0.130]	1.000
N = 1000	Progressive	A-SIR PF	10	3.318	3.422	+0.104 [−0.977, +1.184]	1.000
	Progressive	SIR-ESS PF	10	3.318	0.286	−3.033 [−4.334, −1.732]	0.137
	Step	A-SIR PF	10	3.218	3.315	+0.097 [−1.004, +1.198]	1.000
	Step	SIR-ESS PF	10	3.218	0.156	−3.062 [−4.556, −1.568]	0.137
	Healthy	A-SIR PF	10	3.149	3.575	+0.425 [−2.230, +3.080]	1.000
	Healthy	SIR-ESS PF	10	3.149	0.095	−3.054 [−4.569, −1.539]	0.137
N = 2000	Progressive	A-SIR PF	10	3.710	3.118	−0.592 [−2.551, +1.368]	1.000
	Progressive	SIR-ESS PF	10	3.710	0.250	−3.460 [−4.988, −1.931]	0.137
	Step	A-SIR PF	10	3.974	3.188	−0.786 [−2.722, +1.150]	1.000
	Step	SIR-ESS PF	10	3.974	0.148	−3.826 [−5.392, −2.260]	0.137
	Healthy	A-SIR PF	10	1.675	2.885	+1.211 [+0.359, +2.063]	0.537
	Healthy	SIR-ESS PF	10	1.675	0.083	−1.592 [−2.484, −0.699]	0.137
N = 5000	Progressive	A-SIR PF	30	1.576	1.943	+0.367 [−0.598, +1.333]	1.000
	Progressive	SIR-ESS PF	10	1.530	0.244	−1.286 [−2.253, −0.318]	0.219
	Step	A-SIR PF	30	1.478	1.957	+0.478 [−0.562, +1.518]	1.000
	Step	SIR-ESS PF	10	1.507	0.138	−1.369 [−2.387, −0.351]	0.137
	Healthy	A-SIR PF	30	1.773	1.849	+0.076 [−0.923, +1.075]	1.000
	Healthy	SIR-ESS PF	10	1.583	0.087	−1.497 [−2.694, −0.300]	0.137
N = 10,000	Progressive	A-SIR PF	10	0.642	0.945	+0.304 [−0.159, +0.766]	1.000
	Progressive	SIR-ESS PF	10	0.642	0.242	−0.399 [−0.841, +0.043]	0.137
	Step	A-SIR PF	10	0.512	0.819	+0.307 [−0.218, +0.832]	1.000
	Step	SIR-ESS PF	10	0.512	0.134	−0.379 [−0.900, +0.143]	0.137
	Healthy	A-SIR PF	10	0.304	0.804	+0.501 [−0.123, +1.124]	1.000
	Healthy	SIR-ESS PF	10	0.304	0.081	−0.223 [−0.458, +0.012]	0.137

References

1. Malekloo, A.; Ozer, E.; AlHamaydeh, M.; Girolami, M. Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights. *Struct. Health Monit.* **2022**, *21*, 1906–1955. https://doi.org/10.1177/14759217211036880.
2. Kamariotis, A.; Chatzi, E.; Straub, D.; Dervilis, N.; Goebel, K.; Hughes, A.J.; Lombaert, G.; Papadimitriou, C.; Papakonstantinou, K.G.; Pozzi, M.; et al. Monitoring-supported value generation for managing structures and infrastructure systems. *Data-Centric Eng.* **2024**, *5*, e27. https://doi.org/10.1017/dce.2024.24.
3. Aghelizadeh, P.; Tang, Y.; Sychterz, A.; Sneed, H.L. Review on vibration-based damage detection of vehicular bridges considering traffic effects. *J. Civ. Struct. Health Monit.* **2026**, *16*, 53. https://doi.org/10.1007/s13349-026-01102-4.
4. Zhang, C.; Mousavi, A.A.; Masri, S.F.; Gholipour, G.; Yan, K.; Li, X. Vibration feature extraction using signal processing techniques for structural health monitoring: A review. *Mech. Syst. Signal Process.* **2022**, *177*, 109175. https://doi.org/10.1016/j.ymssp.2022.109175.

5. Farrar, C.R.; Baker, W.E.; Bell, T.M.; Cone, K.M.; Darling, T.W.; Duffey, T.A.; Eklund, A.; Migliori, A. *Dynamic Characterization and Damage Detection in the I-40 Bridge over the Rio Grande*; Technical Report; Los Alamos National Laboratory: Los Alamos, NM, USA, 1994. https://doi.org/10.2172/10158042.
6. Farrar, C.R.; Doebling, S.W.; Cornwell, P.J.; Straser, E.G. Variability of modal parameters measured on the Alamosa Canyon Bridge. In Proceedings of the International Modal Analysis Conference, Orlando, FL, USA, 3–6 February, 1997.
7. Biondini, F.; Frangopol, D.M. Life-cycle performance of deteriorating structural systems under uncertainty: Review. *J. Struct. Eng.* **2016**, *142*, F4016001. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001544.
8. Figueiredo, E.; Peres, N.; Moldovan, I.; Nasr, A. Impact of climate change on long-term damage detection for structural health monitoring of bridges. *Struct. Health Monit.* **2025**, *24*, 2252–2270. https://doi.org/10.1177/14759217231224254.
9. Yang, K.; Kim, S.; Yue, R.; Yue, H.; Harley, J.B. Long-term guided wave structural health monitoring in an uncontrolled environment through long short-term principal component analysis. *Struct. Health Monit.* **2022**, *21*, 1501–1517. https://doi.org/10.1177/14759217211035532.
10. Svendsen, B.T.; Øiseth, O.; Frøseth, G.T.; Rønnquist, A. A hybrid structural health monitoring approach for damage detection in steel bridges under simulated environmental conditions using numerical and experimental data. *Struct. Health Monit.* **2023**, *22*, 540–561. https://doi.org/10.1177/14759217221098998.
11. Neumann, L.; Wittenberg, P.; Mendler, A.; Gertheiss, J. Signal and feature reconstruction in structural health monitoring under varying environmental conditions. *Eng. Struct.* **2026**, *361*, 122865. https://doi.org/10.1016/j.engstruct.2026.122865.
12. Yan, A.M.; Kerschen. G.; De Boe, P.; Golinval, J.C. Structural damage diagnosis under varying environmental conditions-Part II: Local PCA for non-linear cases. *Mech. Syst. Signal Process.* **2005**, *19*, 865–880.
13. Turrisi, S.; Cigada, A.; Zappa, E. A cointegration-based approach for automatic anomalies detection in large-scale structures. *Mech. Syst. Signal Process.* **2022**, *166*, 108483. https://doi.org/10.1016/j.ymssp.2021.108483.
14. Cross, E.J.; Worden, K.; Chen, Q. Cointegration: A novel approach for the removal of environmental trends in structural health monitoring data. *Proc. R. Soc. A* **2011**, *467*, 2712–2732.
15. Shi, H.; Worden, K.; Cross, E.J. A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring. *Mech. Syst. Signal Process.* **2018**, *103*, 381–397.
16. Römgens, N.; Abbassi, A.; Dierksen, N.; Grießmann, T.; Hofmeister, B.; Hübler, C.; Rolfes, R. Temperature-compensated autoencoders with sequences of raw time-domain signals for damage detection and localization in active and passive SHM. *Struct. Health Monit.* 2025, *advance online publication*. https://doi.org/10.1177/14759217251386293.
17. Li, L.; Morgantini, M.; Betti, R. Structural damage assessment through a new generalized autoencoder with features in the quefrency domain. *Mech. Syst. Signal Process.* **2023**, *184*, 109713. https://doi.org/10.1016/j.ymssp.2022.109713.
18. Li, X.; Bolandi, H.; Masmoudi, M.; Salem, T.; Jha, A.; Lajnef, N.; Boddeti, V.N. Mechanics-informed autoencoder enables automated detection and localization of unforeseen structural damage. *Nat. Commun.* **2024**, *15*, 9229. https://doi.org/10.1038/s41467-024-52501-4.
19. Wan, C.; Huang, R.; Huang, L.; Wen, B.; Sato, T. Damage identification using particle filters. *Procedia Eng.* **2017**, *188*, 41–47.
20. Doucet, A.; Godsill, S.l Andrieu, C. On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat. Comput.* **2000**, *10*, 197–208. https://doi.org/10.1023/A:1008935410038.
21. Li, T.; Sbarufatti, C.; Cadini, F.; Chen, J.; Yuan, S. Particle filter-based hybrid damage prognosis considering measurement bias. *Struct. Control Health Monit.* **2022**, *29*, e2914. https://doi.org/10.1002/stc.2914.
22. Corbetta, M.; Sbarufatti, C.; Giglio, M.; Todd, M.D. Optimization of nonlinear, non-Gaussian Bayesian filtering for diagnosis and prognosis of monotonic degradation processes. *Mech. Syst. Signal Process.* **2018**, *104*, 305–322. https://doi.org/10.1016/j.ymssp.2017.11.012.
23. Sbarufatti, C.; Corbetta, M.; Giglio, M.; Cadini, F. Adaptive prognosis of lithium-ion batteries based on the combination of particle filters and radial basis function neural networks. *J. Power Sources* **2017**, *344*, 128–140. https://doi.org/10.1016/j.jpowsour.2017.01.105.
24. Cadini, F.; Sbarufatti, C.; Cancelliere, F.; Giglio, M. State-of-life prognosis and diagnosis of lithium-ion batteries by data-driven particle filters. *Appl. Energy* **2019**, *235*, 661–672. https://doi.org/10.1016/j.apenergy.2018.10.095.
25. Chen, J.; Yuan, S.; Sbarufatti, C.; Jin, X. Dual crack growth prognosis by using a mixture proposal particle filter and on-line crack monitoring. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107758. https://doi.org/10.1016/j.ress.2021.107758.
26. Liu, X.; Lunde, E.; Diehl, F.; Zhang, A.; Vatn, J.; Yin, S. Degradation modeling under time-varying operating conditions: Inference and prognosis with particle filter. *Reliab. Eng. Syst. Saf.* **2024**, *245*, 109965. https://doi.org/10.1016/j.ress.2024.109965.
27. Chatzi, E.N.; Smyth, A.W. Particle filter scheme with mutation for the estimation of time-invariant parameters in structural health monitoring applications. *Struct. Control Health Monit.* **2013**, *20*, 1081–1095. https://doi.org/10.1002/stc.1520.
28. Ma, Y.; Cui, C.; Tang, K.; Xia, Y.; Zhang, Q.; Lao, W. A generalized probabilistic method for fatigue life prediction of butt-welded joints with multiple cracks based on equivalent initial flaw size distribution. *Eng. Struct.* **2026**, *349*, 121943. https://doi.org/10.1016/j.engstruct.2025.121943.

29. Chatzi, E.; Smyth, A. The unscented Kalman filter and particle filter methods for nonlinear structural system identification with non-collocated heterogeneous sensing. *Struct. Control Health Monit.* **2009**, *16*, 99–123.
30. Cadini, F.; Lomazzi, L.; Ferrater Roca, M.; Sbarufatti, C.; Giglio, M. Neutralization of temperature effects in damage diagnosis of MDOF systems by combinations of autoencoders and particle filters. *Mech. Syst. Signal Process.* **2022**, *162*, 108048. https://doi.org/10.1016/j.ymssp.2021.108048.
31. Lomazzi, L.; Cadini, F.; Giglio, M. Particle filters and auto-encoders combination for damage diagnosis on hysteretic non-linear structures subject to changing environmental conditions. In *Dynamics of Civil Structures, Volume 2: Conference Proceedings of the Society for Experimental Mechanics*; Springer: Cham, Switzerland, 2021. https://doi.org/10.1007/978-3-030-77143-0_1.
32. Loh, C.H.; Mao, C.H.; Chao, S.H.; Weng, J.H. Feature extraction and system identification of reinforced concrete structures considering degrading hysteresis. *Struct. Control Health Monit.* **2010**, *17*, 712–729.
33. Baber, T.T.; Wen, Y.K. Random vibration of hysteretic, degrading systems. *J. Eng. Mech.* **1981**, *107*, 1069–1087.
34. Heredia-Pérez, M.; Alvarez, D.A.; Bedoya-Ruiz, D. A state-of-the-art review of the Bouc-Wen class model of hysteresis: Origin, evolution and current state. *Arch. Comput. Methods Eng.* **2026**, *33*, 1–51. https://doi.org/10.1007/s11831-025-10301-z.
35. Nguyen, X.B.; Komatsuzaki, T.; Truong, H.T. Adaptive parameter identification of Bouc-Wen hysteresis model for a vibration system using magnetorheological elastomer. *Int. J. Mech. Sci.* **2021**, *211*, 106848. https://doi.org/10.1016/j.ijmecsci.2021.106848.
36. Miguel, L.P.; Teloli, R.O.; da Silva, S. Bayesian model identification through harmonic balance method for hysteresis prediction in bolted joints. *Nonlinear Dyn.* **2022**, *107*, 77–98. https://doi.org/10.1007/s11071-021-06967-2.
37. Ding, Z.; Yu, Y.; Xia, Y. Nonlinear hysteretic parameter identification using an attention-based long short-term memory network and principal component analysis. *Nonlinear Dyn.* **2023**, *111*, 4559–4576. https://doi.org/10.1007/s11071-022-08095-x.
38. Zhai, W.; Bao, Y.; Tao, D. State space model-based Runge-Kutta gated recurrent unit networks for structural response prediction. *Nonlinear Dyn.* **2024**, *112*, 21901–21921. https://doi.org/10.1007/s11071-024-10229-2.
39. Julier, S.J.; Uhlmann, J.K. Unscented filtering and nonlinear estimation. *Proc. IEEE* **2004**, *92*, 401–422. https://doi.org/10.1109/JPROC.2003.823141.
40. Burgers, G.; van Leeuwen, P.J.; Evensen, G. Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.* **1998**, *126*, 1719–1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.
41. Evensen, G. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dyn* **2003**, *53*, 343–367. https://doi.org/10.1007/s10236-003-0036-9.
42. Wu, M.; Smyth, A.W. Application of the unscented Kalman filter for real-time nonlinear structural system identification. *Struct. Control Health Monit.* **2007**, *14*, 971–990. https://doi.org/10.1002/stc.186.
43. Ching, J.; Beck, J.L.; Porter, K.A. Bayesian state and parameter estimation of uncertain dynamical systems. *Probabilistic Eng. Mech.* **2006**, *21*, 81–96. https://doi.org/10.1016/j.probengmech.2005.08.003.
44. Kong, A.; Liu, J.S.; Wong, W.H. Sequential imputations and Bayesian missing data problems. *J. Am. Stat. Assoc.* **1994**, *89*, 278–288. https://doi.org/10.1080/01621459.1994.10476469.
45. Liu, J.S.; Chen, R. Sequential Monte Carlo methods for dynamic systems. *J. Am. Stat. Assoc.* **1998**, *93*, 1032–1044. https://doi.org/10.1080/01621459.1998.10473765.
46. Douc, R.; Cappé, O.; Moulines, E. Comparison of resampling schemes for particle filtering. In Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis (ISPA 2005), Zagreb, Croatia, 15–17 September, 2005; pp. 64–69. https://doi.org/10.1109/ISPA.2005.195385.
47. Liu, J.; West, M. Combined parameter and state estimation in simulation-based filtering. In *Sequential Monte Carlo Methods in Practice*; Doucet, A., de Freitas, N., Gordon, N., Eds.; Springer: New York, NY, USA, 2001; pp. 197–223. https://doi.org/10.1007/978-1-4757-3437-9_10.
48. Kantas, N.; Doucet, A.; Singh, S.S.; Maciejowski, J.; Chopin, N. On particle methods for parameter estimation in state-space models. *Stat. Sci.* **2015**, *30*, 328–351. https://doi.org/10.1214/14-STS511.
49. Herbst, E.; Schorfheide, F. Tempered particle filtering. *J. Econom.* **2019**, *210*, 26–44. https://doi.org/10.1016/j.jeconom.2018.11.003.
50. Pinello, L.; Lomazzi, L. Action-aided Particle Filter for Damage Diagnosis on Hysteretic Non-Linear Structures Under Varying Environmental Conditions. In Proceedings of the 12th European Workshop on Structural Health Monitoring, Toulouse, France, 7–10 July, 2026. Available online: https://www.ndt.net/search/docs.php?id=33676 (accessed on 5 August 2026).
51. Vanik, M.W.; Beck, J.L.; Au, S.K. Bayesian probabilistic approach to structural health monitoring. *J. Eng. Mech.* **2000**, *126*, 738–745. https://doi.org/10.1061/(ASCE)0733-9399(2000)126:7(738).
52. Teughels, A.; De Roeck, G. Structural damage identification of the highway bridge Z24 by FE model updating. *J. Sound Vib.* **2004**, *278*, 589–610. https://doi.org/10.1016/j.jsv.2003.10.041.
53. Peeters, B.; De Roeck, G. One-year monitoring of the Z24-Bridge: Environmental effects versus damage events. *Earthq. Eng. Struct. Dyn.* **2001**, *30*, 149–171. https://doi.org/10.1002/1096-9845(200102)30:2<149::AID-EQE1>3.0.CO;2-Z.
54. Moaveni, B.; Behmanesh, I. Effects of changing ambient temperature on finite element model updating of the Dowling Hall Footbridge. *Eng. Struct.* **2012**, *43*, 58–68. https://doi.org/10.1016/j.engstruct.2012.05.009.

55. Behmanesh, I.; Moaveni, B.; Lombaer, G.; Papadimitriou, C. Hierarchical Bayesian model updating for structural identification. *Mech. Syst. Signal Process.* **2015**, *64–65*, 360–376. https://doi.org/10.1016/j.ymssp.2015.03.026.
56. Simon, D. Kalman filtering with state constraints: A survey of linear and nonlinear algorithms. *IET Control Theory Appl.* **2010**, *4*, 1303–1318. https://doi.org/10.1049/iet-cta.2009.0032.
57. Lang, L.; Chen, W.; Bakshi, B.R.; Goel, P.K.; Ungarala, S. Bayesian estimation via sequential Monte Carlo sampling—Constrained dynamic systems. *Automatica* **2007**, *43*, 1615–1622. https://doi.org/10.1016/j.automatica.2007.02.012.
58. Shao, X.; Huang, B.; Lee, J.M. Constrained Bayesian state estimation—A comparative study and a new particle filter based approach. *J. Process Control* **2010**, *20*, 143–157. https://doi.org/10.1016/j.jprocont.2009.11.002.
59. Cheng, L.; Cigada, A. An analytical perspective about structural damage identification based on transmissibility funciton. *Struct. Health Monit.* **2020**, *19*, 142–155. https://doi.org/10.1177/1475921719838079.
60. Rayleigh, J.W.S.B. *The Theory of Sound*; Macmillan: New York, NY, USA, 1896.
61. Soo Loon Wah, W.; Chen, Y.-T. A new approach toward damage localization and quantification of structures under changing temperature condition. *J. Low Freq. Noise Vib. Act. Control* **2018**, *99*, 306–325. https://doi.org/10.1177/1461348418793079.
62. Kitagawa, G. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *J. Comput. Graph. Stat.* **1996**, *5*, 1–25. https://doi.org/10.1080/10618600.1996.10474692.
63. Del Moral, P. *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*; Springer: New York, NY, USA, 2004. https://doi.org/10.1007/978-1-4684-9393-1.
64. Simon, D. *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*; Wiley: Hoboken, NJ, USA, 2006. https://doi.org/10.1002/0470045345.
65. Rauch, H.E.; Tung, F.; Striebel, C.T. Maximum likelihood estimates of linear dynamic systems. *AIAA J.* **1965**, *3*, 1445–1450. https://doi.org/10.2514/3.3166.
66. Särkkä, S. *Bayesian Filtering and Smoothing*; Cambridge University Press: Cambridge, UK, 2013. https://doi.org/10.1017/CBO9781139344203.
67. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
68. Wilcoxon, F. Individual comparisons by ranking methods. *Biom. Bull.* **1945**, *1*, 80–83. https://doi.org/10.2307/3001968.
69. McNemar, Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* **1947**, *12*, 153–157. https://doi.org/10.1007/BF02295996.
70. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
71. Doucet, A.; Johansen, A.M. A tutorial on particle filtering and smoothing: Fifteen years later. In *The Oxford Handbook of Nonlinear Filtering*; Crisan. D., Rozovskii, B., Eds.; Oxford University Press: Oxford, UK, 2011; pp. 656–704.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.