This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

A Data-Driven Method for Online Monitoring Tube Wall Thinning Process in Dynamic Noisy Environment

Chen Zhang, Jun Long Lim, *Member, IEEE*, Ouyang Liu, Aayush Madan, *Member, IEEE*, Yongwei Zhu, *Member, IEEE*, Shili Xiang, *Member, IEEE*, Kai Wu, Rebecca Yen-Ni Wong, Jiliang Eugene Phua, *Member, IEEE*, Karan M. Sabnani, Keng Boon Siah, Wenyu Jiang, Yixin Wang, *Member, IEEE*, Emily Jianzhong Hao, *Member, IEEE*, and Steven C. H. Hoi, *Fellow, IEEE*

***Abstract*— Tube internal erosion, which corresponds to its wall thinning process, is one of the major safety concerns for tubes. Many sensing technologies have been developed to detect a tube wall thinning process. Among them, fiber Bragg grating (FBG) sensors are the most popular ones due to their precise measurement properties. Most of the current works focus on how to design different types of FBG sensors according to certain physical laws and only test their sensors in controlled laboratory conditions. However, in practice, an industrial system usually suffers from harsh and dynamic environmental conditions, and FBG signals are affected by many unpredictable factors. Consequently, the FBG signals have more fluctuations and are polluted by noises. Hence, the signals no longer directly follow the assumed physical laws and their proposed thinning detection mechanisms no longer work. Targeting at this, this article develops a data-driven model for FBG signal feature extraction and tube wall thickness monitoring using data analytic techniques. In particular, we develop a spatiotemporal model to describe dynamic FBG signals and extract features related to thickness. By taking physical law as guideline, we trace the relationship between the extracted features and the tube wall thickness, based on which we construct an online statistical monitoring scheme for tube wall thinning process. We use both laboratory test and field trial experiment to demonstrate the efficacy and efficiency of the proposed scheme.**

***Note to Practitioners*—This article is motivated by the real industrial needs of inner erosion detection of tubes in harsh environment. Most of the current research works focus on designing various sensing apparatuses based on fiber Bragg grating (FBG) sensors for nondestructive erosion detection. These apparatuses prove to be able to collect signals reflecting tube wall thickness in static and controllable laboratory environment qualitatively. However, in reality, the industrial environment, which is impacted by many changing factors, is dynamic and uncontrollable. Consequently, the signals collected by these FBG apparatuses would have larger variations that mask the signals related to thickness. Furthermore, current methods have neither mentioned how to process their collected data to capture the unnoticeably slow but accumulative erosion information efficiently nor constructed online monitoring algorithms to detect the tube wall thinning process based on the collected signals quantitatively. Built upon their apparatuses but targeting at their unsolved challenges, we propose a novel data-driven approach for FBG signal analysis that can remove the environmental influence and extract features only related to tube wall thickness, and using the extracted features, we construct a statistical process control scheme to monitor tube wall thickness and detect erosion in real time efficiently.**

***Index Terms*— Fiber Bragg grating (FBG) sensors, online monitoring, spatiotemporal model, statistical process control, tube erosion detection.**

Manuscript received February 28, 2020; revised June 30, 2020; accepted August 24, 2020. This article was recommended for publication by Associate Editor K. Liu and Editor F.-T. Cheng upon evaluation of the reviewers' comments. This work was supported in part by the National Natural Science Foundation of China under Grant 71932006 and Grant 71901131; in part by the Tsinghua University Intelligent Logistics and Supply Chain Research Center under Grant THUCSL20182911756-001; and in part by the National Research Foundation (NRF) Singapore, Energy Innovation Research Programme, through its Sembcorp-EMA Energy Technology Partnership (SEETP) (Singapore, LA) under Contract NRF2015EWT-EIRP001-010. *(Corresponding author: Chen Zhang.)*

Chen Zhang was with the School of Information Systems, Singapore Management University, Singapore 178902. She is now with the Department of Industrial Engineering, Tsinghua University, Beijing 100084, China (e-mail: zhangchen01@tsinghua.edu.cn).

Jun Long Lim, Ouyang Liu, Aayush Madan, Yongwei Zhu, Shili Xiang, Kai Wu, Rebecca Yen-Ni Wong, Jiliang Eugene Phua, Wenyu Jiang, Yixin Wang, and Emily Jianzhong Hao are with the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore 138632.

Karan M. Sabnani and Keng Boon Siah are with Sembcorp Industries Ltd., Singapore 179360.

Steven C. H. Hoi is with the School of Information Systems, Singapore Management University, Singapore 178902.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TASE.2020.3038708.

Digital Object Identifier 10.1109/TASE.2020.3038708

I. Introduction

TUBE (or pipe or pipeline) is an important structure for liquid transportation in many manufacturing systems or civil systems. However, due to the effect of inner fluids and harsh environmental conditions, the tube is particularly prone to be corroded. Once a tube is subjected to continuous erosion, its wall would become thinner and thinner and may probably cause leakage or rupture. Thus, it is of great significance to monitor the erosion level, i.e., thinning process of tube wall, to ensure its safe operation. In general, the thinning processes can be classified as corrosion of external walls and erosion of internal walls. External corrosion refers to the

1545-5955 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

(a) (b)

Fig. 1. Erosion of U-bend tubes in an HRSG boiler. (a) U-bend tube facilities. (b) Internal erosion.

deterioration processes occurring on metal surfaces of tubes that are exposed to outdoor environments and is largely driven by atmospheric and climatic conditions. External corrosion can be easily prevented by protecting the external surface of tube with insulation and cathodic materials. Internal erosion, as shown in Fig. 1(b), occurs on the surfaces inside the tubes due to the joint action of fluid flushing, corrosive medium, and microorganism and is more difficult to be prevented in general. As such, monitoring internal erosion, focusing on detecting the initiation of erosion and assessing the tube safety performance, has become one of the best options to offer early warning and to avoid high cost of later repairing and has become an important research topic [1].

There are many nondestructive methods to detect tube internal erosion. Some representative and mature techniques include electrochemical impedance spectroscopy [2], linear polarization resistance [3], [4], and electronic resistance technique [5]. However, for oil and gas tube monitoring, the use of these electrical sensors above will cause potential dangers to tube safety operation [6]. Moreover, when the erosion location is far away from the monitoring apparatus, the signals of electronic sensors would be really vulnerable to harsh environment and external electromagnetic disruptions. The optical fiber sensor is a promising technology, due to its advantages of geometric versatility and high measurement accuracy and reliability that are particularly attractive in harsh environment [7]–[9] with high temperatures and pressures, and strong electromagnetic fields. Considering their superior abilities, optical fiber sensing technologies have received increasing attention in the study of erosion monitoring in recent years [10]–[16].

One of the most direct phenomena caused by erosion (wall thinning process) is the change of tube circumferential strain. This is because the thinning process will cause the circumferential hoop strain to increase gradually according to the hoop strain theory, i.e.,

$$\epsilon = \frac{\Delta P R}{E \mu} \tag{1}$$

where ϵ is the circumferential strain of the tube; ΔP is the applied pressure; E is Young's modulus, R is the tube radius, and μ is the wall thickness of the tube. It is assumed that E and R are the constants. Therefore, the circumferential strain is inversely proportional to the wall thickness. Consequently, its change can be reflected by the circumferential strain directly. As such, monitoring the hoop strain variation of tube is a desirable way to detect the change of tube wall thickness [16]. Among different types of optical fiber sensors, fiber Bragg grating (FBG) sensors can accurately measure the circumferential strain of a tube with millimeter-scale resolution and microstrain measurement precision and hence can provide an effective method for tube erosion monitoring [15]. In particular, the FBG wavelength shift $\Delta\lambda$ is related to the circumferential strain in the following way:

$$\frac{\Delta\lambda}{\lambda^0} = (1 - P_e)\epsilon + [(1 - P_e)\alpha + \xi]\Delta T \tag{2}$$

where λ^0 is the initial wavelength when the FBG sensor is standalone at room temperature T_0, P_e stands for the photoelastic constant of fiber, α stands for the thermal expansion coefficient of fiber, ξ stands for the thermal-optics coefficient, and ΔT is the temperature change from T_0. Then, with the knowledge of ΔT and $\Delta\lambda$, we can infer ϵ and consequently monitor μ according to (1).

There exist some works for tube corrosion or erosion detection using FBG sensors based on the above physical law. In particular, Gao *et al.* [17] developed an FBG corrosion sensor to monitor the corrosion rate in reinforced concrete structures. Hu *et al.* [18] developed a Fe-C-coated FBG sensor for steel corrosion monitoring by measuring strains in the radius direction when wrapped on the steel bar. Ren *et al.* [19] designed an FBG hoop-strain sensor that could be used to measure tube uniform corrosion levels. Deng *et al.* [20] further discussed applying FBG sensors into corrosion detection of tubes with soft coatings.

All of these works focus on how to design FBG sensors and the measurement system, and they only test their designed apparatuses in laboratory environments by controlling all environmental variables to be consistent. However, in real applications, besides tube wall thickness, other environmental factors, such as tube temperature and pressure, i.e., ΔP and ΔT, would fluctuate and also affect FBG signals. Then, if we omit the variation of ΔP and ΔT and treat them as constant when inferring μ via (2), these variations would be transferred to the variation of the estimated μ and hence deteriorate the online monitoring accuracy.

For example, Fig. 2 shows the signals of 13 FBG sensors with 1-kHz sampling rate mount at 13 tubes of a heat recovery stream generator (HRSG) boiler as shown in Fig. 1(a) over a time period (19 days). The wavelength shift of each FBG sensor changes with respect to λ_i^0. Since different tubes have various wall thicknesses as presented in Table V (due to confidential reason, we hide the measurement unit here and mask the thickness magnitude by setting it as a reference thickness of μ_0 plus a deviation). Note that μ_0 is a fixed constant throughout this article), according to (2), their Bragg wavelength change is expected to have different shift magnitudes, i.e., FBG sensors on thinner wall tubes are presumed to have larger shifts in general. However, by plotting the FBG signals of tubes with smallest to largest wall thicknesses via lines from yellow to red in Fig. 2, we cannot observe this

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

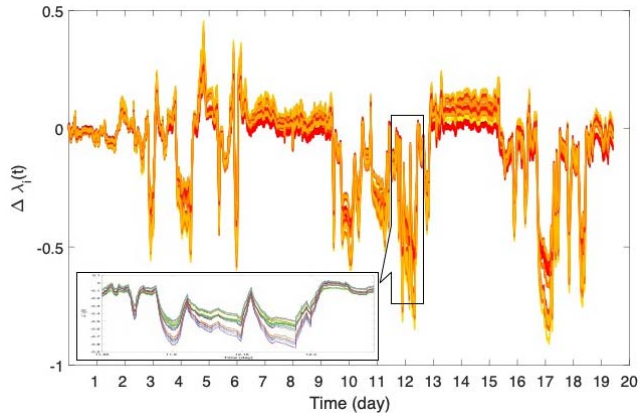


Fig. 2. Wavelength shifts of 13 FBG sensors on the HRSG boiler from their initial values over time. Each line represents one FBG sensor's signals over time. The color from red to yellow represents FBG sensors with the thinnest wall thickness to the thickest wall thickness.

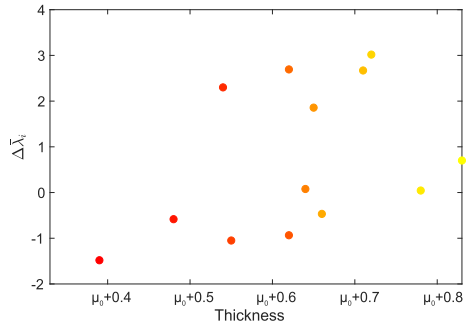


Fig. 3. Mean of wavelength shifts calculated from Fig. 2 from day 13 to day 15 (time window) versus their wall thicknesses (due to confidential reason, we hide the measurement unit here and mask the thickness magnitude by setting it as a reference thickness of μ_0 plus a deviation).

monotone property at all. Clearly, all the signals have large longitudinal dynamics and evolve in a similar pattern. This is caused by the change of environmental factors, such as temperature and pressure over time, driven by the manufacturing dynamics in practice. Furthermore, even if the global environmental factors are kept stable in some short time windows (such as the window from day 13 to day 15 in Fig. 2), each FBG sensor still has its own variation features due to local environmental fluctuations. If we calculate the average wavelength shift $\Delta\bar{\lambda}_i$ of each sensor in this flat time window and plot them in Fig. 3, the monotone relationship between $\Delta\bar{\lambda}_i$ and μ_i is still unobservable. This is because even at the same time point, the temperatures and pressures of different tubes are still different from each other, due to local environmental variations. Consequently, shifts of FBG signals caused by various wall thicknesses (or thinning process) would be hidden by these shifts caused by either global or local dynamic environmental changes.

In a real application environment, the behaviors of FBG sensors are much more random than expected. Some works propose to mount additional sensors to track the local temperatures and pressures as well and plug their real-time values in (2). However, it is hard to guarantee that their measured temperatures and pressures are exactly equal to the ones where the FBG sensors are located. Even if we can measure ΔP and ΔT accurately, direct plugging them into (2) for inference is still inaccurate since the other physical constants in (2), such as E, R, and P_e, may also change with ΔT, while their change magnitudes are hard to be quantitatively described. To remove the influence of all these factors, some other works suggest mounting unstained FBG sensors around each target FBG sensor on the tube as references. By subtracting these reference sensors' signals from the target sensors, it is expected that signals related to environmental changes can be compensated and signals only related to thickness can be exposed. However, this compensation method is not always applicable. In reality, reference sensors sometimes cannot be mounted due to environmental or financial constraints. Even if they can be mounted, they have limited functions in practice since the environmental conditions of these unstained FBG sensors may still be different from conditions of the target FBG sensors on the tubes. Consequently, this compensation loses its meaning. Furthermore, all the existing methods mentioned earlier only prove that the FBG signals collected from designed apparatuses can reflect the tube wall thickness qualitatively in the offline erosion detection test, while they still have a far way to go for constructing a quantitative algorithm for online tube wall thickness change detection.

Considering the limitations of these hardware-based signal correction methods, another promising direction is to use data analytic techniques to extract FBG signals' features only related to wall thickness. Targeting at this, this article develops a novel data-driven approach for FBG sensor-based tube wall thinning process monitoring. Our approach is built upon physical law but adjusted with statistical models. In particular, we develop a spatiotemporal model to remove the environmental influence and extract features of FBG signals that relate to tube wall thickness. Specifically, for each FBG sensor, we treat its signals as response (dependent variable) and use the other sensors' signals as predictors (independent variables) for representation, with the basic idea that the representable part indicates features that coexist among all the sensors, and correspond to global environmental changes. Then, the left residual signals can be regarded as features that specially belong to the response sensor and correspond to features caused by its own properties, such as its wall thickness (wall thinning process) and local system vibrations. Consequentially, we can use the representation residuals for feature extraction. By taking the physical law as guideline, we describe the relationship between the extracted residual feature and the tube wall thickness. Finally, based on the relationship, we construct an online monitoring scheme for tube wall thinning process. The scheme monitors the residual feature in real time and triggers an alarm once it goes out of the prespecified control limit, which indicates that the tube wall thickness decreases to the minimum tolerable value and should be replaced.

The remainder of this article is organized as follows. Section II introduces our problem formulation to describe the FBG signals on tubes in dynamic noisy environments. Section III reviews related research works that have the potential to infer the formulated problem and their limitations.

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Section IV introduces our proposed spatiotemporal model for feature extraction in detail. Section V constructs the relationship between the extracted feature and the tube wall thickness and, built upon them, develops the online monitoring scheme. Section VI conducts some numerical studies using synthetic data to validate the proposed methods. Section VII applies the proposed methodology into tube wall thinning process detection with both lab test and real field trials. Finally, Section VIII concludes this article with remarks.

II. Problem Formulation

Suppose that we have p FBG sensors mounted at p different tubes. For each sensor i, we assume that its initial wavelength λ_i^0 is known and collect its signals $\lambda_i(t)$ in real time. We define the relative wavelength shift of sensor i at sensing time point t as

$$X_i(t) = \frac{\lambda_i(t) - \lambda_i^0}{\lambda_i^0}$$

which will change due to system dynamics in real time. In particular, we denote the real-time global temperature and pressure as $\Delta P(t)$ and $\Delta T(t)$, respectively. Furthermore, to address local environmental fluctuations of different tubes, for each sensor, we further add disturbance items $\tilde{P}_i(t)$ and $\tilde{T}_i(t)$ on $\Delta P(t)$ and $\Delta T(t)$. Consequently, the physical law that each sensor follows can be formulated as

$$X_i(t) = (1 - P_e)\frac{(\Delta P(t) + \tilde{P}_i(t))R}{E\mu_i} + [(1 - P_e)\alpha + \xi](\Delta T(t) + \tilde{T}_i(t)) \quad (3)$$

where μ_i denotes the tube wall thickness of sensor i's location. It is to be noted that in (3), we only consider t of a short time window, i.e., $W_0 = [0, L]$, $t \in W_0$, with the assumption that μ_i will not change in W_0. Furthermore, we assume that the temperature disturbance $\tilde{T}_i(t)$ follows a homogeneous distribution over time with mean as 0 and variance as σ_T^2 for all the sensors. Similarly, the pressure disturbance $\Delta\tilde{P}_i$ follows another homogeneous distribution over time with mean 0 and variance σ_P^2.

Though (3) provides a promising model to describe the FBG wavelength shifts in a real dynamic system, in practice, it is very challenging to infer the system based on (3), due to the unavailability of $\Delta T(t)$, $\Delta P(t)$, $\tilde{T}_i(t)$, and $\tilde{P}_i(t)$ as mentioned in Section I. Furthermore, for these FBG sensors, their physical constants, i.e., P_e, α, and ξ, will differ from their theoretical values on the manual book due to the influence of apparatus settings, such as the tube materials and sensor mounting techniques. Consequently, P_e, ξ, and α cannot be precisely inferred as well. All of these require us to estimate P_e, ξ, and α together with $\Delta T(t)$, $\Delta P(t)$, $\tilde{T}_i(t)$, and $\tilde{P}_i(t)$ from the system. However, these variables have intersections with each other, making it almost impossible to use traditional statistical parametric models for inference. Consequently, we have to bypass direct estimation of (3). Instead, we hope to extract intermediary features that relate to μ_i from (3) and use these features to estimate and monitor μ_i.

In particular, by reformulating (3), it can be separated into two parts, i.e., $X_i(t) = X_i^0(t) + \tilde{X}_i(t)$ where

$$X_i^0(t) = (1 - P_e)\frac{\Delta P(t)R}{E\mu_i} + [(1 - P_e)\alpha + \xi]\Delta T(t) \quad (4)$$

$$\tilde{X}_i(t) = (1 - P_e)\frac{\tilde{P}_i(t)R}{E\mu_i} + [(1 - P_e)\alpha + \xi]\tilde{T}_i(t) \quad (5)$$

where $X_i^0(t)$ is the FBG wavelength shift caused by global environmental dynamics. Its pattern is influenced by two sources, i.e., system temperature and pressure, and can be regarded as signals. $\tilde{X}_i(t)$ is the FBG wavelength shift caused by local dynamics and can be regarded as noise. According to our assumptions, it has $E(\tilde{X}_i(t)) = 0$ and $\mathrm{Var}(\tilde{X}_i(t)) = \theta_1/\mu_i^2 + \theta_2$, where $\theta_1 = (1 - P_e)^2R^2\sigma_P^2/E^2$ and $\theta_2 = [(1 - P_e)\alpha + \xi]^2\sigma_T^2$ are constants.

III. Literature Review

In this section, we review some research works that have potential to infer the system of (4) and (5).

As observed, both $X_i^0(t)$ and $\tilde{X}_i(t)$ include μ_i as parameters and can be potentially used for inference of μ_i. If we would like to focus on the signal part, we need to separate the influence of the two sources, i.e., $\Delta P(t)$ and $\Delta T(t)$, since μ_i only affects the magnitude of $\Delta P(t)$. This is related to the problem of blind source separation (BSS) [21]. BSS targets at separating the influence of different unknown blind sources only by using mixed signals of these sources. However, one crucial obstacle to applying BSS in our case is its well-known indeterminacy problem [22] that any permutation or scaling matrix will lead to equivalent separation but different source signals and coefficients. This indeterminacy will not bring any problem if we only do BSS for FBG signals once, for a single time window. However, for online monitoring, we have to do BSS for different time windows sequentially. If the BSS results of different windows have different latent permutations or scaling, it would be impossible to compare the coefficients of different windows, which are functions of μ_i, with the same rule. Consequently, we cannot set a decision criterion for thickness changes.

As an alternative, analyzing the noise part will simplify the problem. Recall that $\mathrm{Var}(\tilde{X}_i(t)) = \theta_1/\mu_i^2 + \theta_2$. This indicates that as long as we can estimate the noise part or simply just its variance, correctly, we can infer the information of μ_i based on this linear relationship between $\mathrm{Var}(\tilde{X}_i(t))$ and $1/\mu_i$. Signal–noise separation is a classical problem [23] in many applications, such as speech signal enhancement and recognition and EEG signal separation, and has been a major challenge for many researchers and engineers for more than four decades [24], [25]. Many approaches have been investigated, including power spectral subtraction [26], Wiener filtering [27], soft-decision estimation [28], and subspace methods [29]. Most of these pioneer works assume that only single-channel signal (i.e., only one FBG sensor in our case) is available and use a temporal prediction-related method for noise separation. Though for multichannel cases, these methods can still be used by denoising each channel separately, they miss the correlation information of different channels

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

and consequently may sacrifice the denoising performance. To use their correlation information for better noise removal, Huang *et al.* [30] and Sameni *et al.* [31] proposed spatiotemporal prediction approaches based on Wiener filtering to denoise multichannel signals together. Borowicz [32] also proposed a spatiotemporal prediction approach based on subspace methods. However, all these methods require knowing the variances of signals and noises, and their correlations in advance when conducting signal projections, while in our case, this information is unavailable. Another direction for separating the noise part from the signals is related to stationary subspace analysis (SSA) [33], [34]. SSA methods aim at extracting the stationary part of the original signals, by projecting the original data to a lower dimensional subspace such that the distribution of the projected data does not change over successive time windows. SSA may be applicable in our case since no matter how $X_i^0(t)$ changes dynamically, and $\tilde{X}_i(t)$ keeps random but stationary over time. However, the assumption that $X_i^0(t)$ should be always nonstationary over time might be too strong in many applications. Considering the fact that $X_i^0(t)$ might still be stationary in some time periods as well, but just have different patterns from $\tilde{X}_i(t)$, methods for SSA are still not reliable for solving our problem.

IV. Feature Transfer Algorithm for FBG Signal Modeling

With the limitations of the current signal–noise separation algorithms in mind, here, we present a novel signal–noise separation algorithm to identify $X_i^0(t)$ and $\tilde{X}_i(t)$. Our idea is to assume that $X_i^0(t)$ and $\tilde{X}_i(t)$ come from different subspaces. By using the spatiotemporal correlations of $X_i^0(t)$ of different sensors, a subspace learning algorithm for signal–noise separation can be learned, which is introduced as follows.

The core of our model is that the environmental change will cause all the FBG sensors to change similarly, i.e., the FBG wavelength shifts caused by environmental changes will coexist among all the sensors. Then, our key idea is, as $X_i^0(t)$, $i = 1, \ldots, p$, are linear functions of $\Delta T(t)$ and $\Delta P(t)$, by solving the linear equation system, and we can reversely reformulate $\Delta T(t)$ and $\Delta P(t)$ as linear functions of $X_i^0(t)$, i.e.,

$$\Delta P(t) = \sum_{j=1}^{p} a_j X_j^0(t), \quad \Delta T(t) = \sum_{j=1}^{p} c_j X_j^0(t) \tag{6}$$

where a_j and c_j are the linear coefficients. It is to be noted that strictly speaking, two channels of $X_i^0(t)$ would be enough to express $\Delta P(t)$ and $\Delta T(t)$ in (6). Yet, we incorporate all the channels for general expression. We temporarily put them aside for notation purpose. Then, by plugging (6) into (4), we can get the following self-expressive property:

$$X_i^0(t) = \sum_{j \neq i} b_{ij} X_j^0(t) \tag{7}$$

where $\forall i = 1, \ldots, p$. Here, b_{ij} are coefficients, which are linear functions of a_j and c_j, and other physical constants. Then, we have

$$X_i(t) = \sum_{j \neq i} b_{ij}(X_j(t) - \tilde{X}_j(t)) + \tilde{X}_i(t). \tag{8}$$

It is to be noted that (8) is similar to a regular linear regression, where b_{ij} represents regression coefficients of the predictors and $\tilde{X}_i(t)$ are the residuals. The only difference is that the predictors are not directly observable but are prediction outputs of other regressions. This motivates us an iterative inference algorithm to solve b_{ij} and $\tilde{X}_i(t)$. Specifically, for each iteration, the following conditions hold.

Step 1: For each i, we treat it as the response variable and construct the linear regression $X_i(t) = \sum_{j \neq i} b_{ij} X_j^0(t) + Z_i(t)$ with $X_j^0(t)$ estimated from the last iteration.

Step 2: For each i, we update $X_i^0(t) = X_i(t) - Z_i(t)$ where $Z_i(t)$ is its regression residual from Step 1.

Now, we talk about how to estimate the regression coefficients in Step 1. Though it can be formulated as a standard linear regression and its estimation can be achieved by minimizing the sum of square of the residuals, yet in the following, we design a more complex inference model by carefully considering more factors. On the one hand, recall that the linear system $\Delta P(t) = \sum_{j=1}^{p} a_j X_j^0(t)$ and $\Delta T(t) = \sum_{j=1}^{p} c_j X_j^0(t)$ may have infinite solutions for a_j and c_j. Consequently, the corresponding $b_{ij}, j = 1, \ldots, p$ are not unique and the regression may be underdetermined. In order to give preference to a particular solution with desirable property, a regularization term can be enforced on the norm of $\mathbf{b}_i = [b_{i1}, \ldots, b_{ip}]$. On the other hand, in large-scale systems, even $\Delta P(t)$ and ΔT may be different, due to spatial variations and fluid mechanics in different parts of the system. In other words, there are different subsystems with various environmental conditions. For example, fluid in tubes far away from each other may have different flow speeds, temperatures, and pressures according to requirements of the fluid system. Suppose that we have M subsystems, each of them has its own $\{\Delta P_m(t), \Delta T_m(t)\}_{m=1,\ldots,M}$. This means that $X_i^0(t)$ of different FBG sensors would also come from different M signal subspaces (subsystems) $\mathcal{S}_m, m = 1, \ldots, M$. $X_i^0(t)$ of sensors from the same subspace $\mathcal{S}_m$ shares the same $\{\Delta T_m(t), \Delta P_m(t)\}$, whereas $X_i^0(t)$ of sensors from different subspaces, $\mathcal{S}_{m_1}$ and $\mathcal{S}_{m_2}$, have different $\{\Delta T_{m_1}(t), \Delta P_{m_1}(t)\}$ and $\{\Delta T_{m_2}(t), \Delta P_{m_2}(t)\}$. This indicates that for a particular subspace m, not all the p sensors should necessarily have nonzero a_j and c_j. Instead, only sensors from this subspace has contributions, i.e., $\Delta P_m(t) = \sum_{j \in \mathcal{S}_m} a_j X_j^0(t)$ and $\Delta T_m(t) = \sum_{j \in \mathcal{S}_m} c_j X_j^0(t)$; hence, for $X_i(t) \in \mathcal{S}_m$, its $\mathbf{b}_i$ is naturally a sparse vector. Though in practice, we do not know the subspace structure in advance, this sparsity property of $\mathbf{b}_i$ suggests adding a regularization term on the norm of $\mathbf{b}_i$ to force sensors from other subspaces $\mathcal{S}_k, k \neq m$ to have zero coefficients. Considering these two points, we propose to add l_1 penalty on the norm of $\mathbf{b}_i$. Then, the regression in Step 1 can be estimated via the following optimization function:

$$\min_{\mathbf{b}_i} \sum_{t=1}^{L} Z_i(t)^2 + \lambda_1 \|\mathbf{b}_i\|_1$$
$$\text{s.t. } Z_i(t) = X_i(t) - \sum_{j=1}^{p} b_{ij} X_j^0(t), \quad b_{ii} = 0. \tag{9}$$

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

This is similar to the idea of sparse subspace learning (SSL) [35], [36], where they assume that $X_i(t), i = 1, \ldots, p$ come from different subspaces that are composed of different basis functions. Each $X_i(t), i = 1, \ldots, p$ is a linear combination of the basis functions of its own subspace. Therefore, sensors from the same subspace can be mutually self-represented. However, in practice, the subspace structure is unknown. Thus, a sparse regression can be achieved via adding the l_1 penalty on the self-expressive coefficients. According to SSL, it can guarantee that with high probability, the estimated $\mathbf{b}_i$ would only have nonzero values on components of variables from the same subspace as the output variable [35], [36]. Consequently, benefited from the similar formulation of (9) to SSL, we can also deal with signal–noise separation when multiple subsystems exist. However, note that it is not compulsory to apply (9) into cases with more than one subspace in the system since (9) is only used for estimation of $X_i^0(t)$. This is our largest difference from SSL, whose objective is to cluster different channels into different subspaces by treating b_{ij} as the affinity criterion between two channels. Furthermore, in (9), we do not construct the exact self-expressed regression as SSL by using $X_k(t)$ as the predictor. Instead, we use $X_k^0(t)$. This is because our aim is to estimate the signal part of each channel $X_i^0(t)$ as accurate as possible. Yet, for SSL, they would only like to calculate b_{ij} as the partial dependence between two channels. From this perspective, our method of (9) is not the exact SSL algorithm.

Furthermore, considering the inherent system dynamics, for different t, the spatial distribution of $\Delta P(t)$ and $\Delta T(t)$ may change due to external disturbance. Consequently, the subspace structures of $X_i^0(t), i = 1, \ldots, p$, would change. For example, as shown in Fig. 2, around the time period of day 11.5, all the FBG signals move together, following the global pattern. Later around day 11.8, after a sudden falling slope, the FBG signals begin to move in two different clusters. Each cluster has its local $\Delta P_m(t)$ and $\Delta T_m(t), m = 1, 2$. Then, after another arising and falling slope, the FBG signals begin to move in three subsystems, with three local $\Delta P_m(t)$ and $\Delta T_m(t), m = 1, 2, 3$. To capture this temporal subspace structure of $X_i^0(t)$, we further relax $b_{ij}, j = 1, \ldots, p$ by allowing them to change over time as $b_{ij}(t), t = 1, \ldots, L$. However, this naive relaxation would give too much flexibility on the change of $b_{ij}(t)$, which could lead to severe overfitting. Considering that, in reality, temporal distributions of environmental conditions usually change smoothly over time, we suggest penalizing the change of $b_{ij}(t)$ to encourage its smoothness. As such, the final estimation of Step 1 can be formulated as

$$\begin{aligned}\min_{\mathbf{b}_i(t), t=1,\ldots,L} \quad & \sum_{t=1}^{L} Z_i(t)^2 + \lambda_1 \sum_{t=1}^{L} \|\mathbf{b}_i(t)\|_1 \\ & + \lambda_2 \sum_{t=2}^{L} \|\mathbf{b}_i(t) - \mathbf{b}_i(t-1)\|_2^2 \\ \text{s.t. } & Z_i(t) = X_i(t) - \sum_{j=1}^{p} b_{ij}(t) X_j^0(t), \quad b_{ii}(t) = 0 \end{aligned} \tag{10}$$

where $\mathbf{b}_i(t) = [b_{i1}(t), \ldots, b_{ip}(t)]' \in \mathcal{R}^p$. This construction is similar to the formulation of [37], where a dynamic algorithm of SSL is proposed by allowing $\mathbf{b}_i(t)$ to change over time as well. However, besides the abovementioned difference between our method and SSL, we also differ from [37] by using a different way to regularize the system dynamics. In particular, we add the Laplacian penalty $\|\mathbf{b}_i(t) - \mathbf{b}_i(t-1)\|_2^2$ to enforce its smooth varying over time. This is because in our application, the fluid dynamics in tubes change over time gradually. Yet, Zhang *et al.* [37] added an elementwise l_1 penalty, i.e., the fused LASSO penalty, $\|\mathbf{b}_i(t) - \mathbf{b}_i(t-1)\|_1$ to encourage stepwise change of $\mathbf{b}_i(t)$. This is because in their application of manufacturing systems, they assume that some ON–OFF operations would lead to stepwise change of the dependence relationships of different channels.

Proposition 1: Define $\mathbf{Y} \in \mathcal{R}^{L \times L(p-1)}$ with the components $Y_{t,(t-1)(p-1)+1:t(p-1)} = [X_1^0(t), \ldots, X_{i-1}^0(t), X_{i+1}^0(t), \ldots, X_p^0(t)]$ and other components are 0. $\mathbf{D} \in \mathcal{R}^{(L-1)(p-1) \times L(p-1)}$ is the first-order difference matrix whose $D_{i,i} = -1$ and $D_{i,i+p-1} = 1$ for $i = 1, \ldots, (L-1) \times (p-1)$. $\mathbf{L}_i = (\mathbf{Y}'\mathbf{Y} + \lambda_2 \mathbf{D}'\mathbf{D})$. Then, the optimization problem of (10) is equivalent to a LASSO problem, i.e.,

$$\min_{\boldsymbol{\beta}_i} \ \|\mathbf{W}_i - \mathbf{R}_i \boldsymbol{\beta}_i\|_2^2 + \lambda_1 \|\boldsymbol{\beta}_i\|_1 \tag{11}$$

where $\mathbf{W}_i = \mathbf{L}_i^{-\frac{1}{2}} \mathbf{Y}' \mathbf{X}_i$ with $\mathbf{X}_i = [X_i(1), \ldots, X_i(L)]$; $\mathbf{R}_i = \mathbf{L}_i^{\frac{1}{2}}$. The solved $\boldsymbol{\beta}_i = [\mathbf{b}_{i(-i)}(1)', \ldots, \mathbf{b}_{i(-i)}(L)']' \in \mathcal{R}^{L(p-1)}$ with $\mathbf{b}_{i(-i)}(t) = [b_{i1}(t), \ldots, b_{i(i-1)}(t), b_{i(i+1)}(t), \ldots, b_{ip}(t)]' \in \mathcal{R}^{p-1}$.

Proposition 1 indicates that (10) can be effectively solved using standard LASSO solvers. Then, based on Step 1, Step 2 can be trivially derived.

Remark 1: It is to be noted for different sensors, since their $\mathrm{Var}(\tilde{X}_i(t))$ are different, the same set of $\{\lambda_1, \lambda_2\}$ may have different regularization results, which in turn makes the estimation results of $\mathrm{Var}(\tilde{X}_i(t))$ incomparable. To avoid this, we propose to normalize the first term in (10) by replacing $Z_i(t)^2$ to $Z_i(t)^2 / V_i$, where $V_i = \mathrm{Var}(\tilde{X}_i(t))$ is the variance of the estimated $\tilde{X}_i(t)$ in the last iteration, i.e.,

$$\begin{aligned}\min_{\mathbf{b}_i(t), t=1,\ldots,L} \quad & \sum_{t=1}^{L} \frac{Z_i(t)^2}{V_i} + \lambda_1 \sum_{t=1}^{L} \|\mathbf{b}_i(t)\|_1 \\ & + \lambda_2 \sum_{t=2}^{L} \|\mathbf{b}_i(t) - \mathbf{b}_i(t-1)\|_2^2 \\ \text{s.t. } & Z_i(t) = X_i(t) - \sum_{j=1}^{p} b_{ij}(t) X_j^0(t), \quad b_{ii}(t) = 0 \end{aligned} \tag{12}$$

which can be solved with the corresponding LASSO solution in (11) by replacing $\mathbf{L}_i = (\mathbf{Y}'\mathbf{Y}/V_i + \lambda_2 \mathbf{D}'\mathbf{D})$ and $\mathbf{W}_i = \mathbf{L}_i^{-(1/2)} \mathbf{Y}' \mathbf{X}_i / V_i$. In this way, we may guarantee the fair variable regularization and get comparable $\tilde{X}_i(t)$ for different sensors for the next iteration. The detailed procedure of our proposed signal–noise separation method is shown in Algorithm 1.

Remark 2: In general, the selection of optimal tuning parameters for a given model can be a difficult task, which

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Algorithm 1 Iterative Signal–Noise Separation of $X_i(t)$ for a Short Window Based on the Spatiotemporal Model

Input: Data $X_i(t), t = 1, \ldots, L, i = 1, \ldots, p$, pre-specified λ_1 and λ_2.
In Iteration $k = 1$,
for *each* i **do**
  Solve $\boldsymbol{\beta}_i^1$ and get the corresponding $\mathbf{b}_i^1(t), t = 1, \ldots, L$ according to (12) by setting $X_j^0(t) = X_j(t)$ for $j \neq i$ and $V_i = \mathrm{Var}(X_i(t))$.
  Get the residual $Z_i(t) = X_i(t) - \sum_{j \neq i} b_{ij}^1(t) X_j^0(t)$.
  Calculate $V_i = \mathrm{Var}(Z_i(t))$.
for *Iteration* $k = 2, \ldots$ **do**
  for *each* i **do**
    Solve $\boldsymbol{\beta}_i^k$ and get the corresponding $\mathbf{b}_i^k(t), t = 1, \ldots, L$ according to (12) by setting $X_j^0(t) = X_j(t) - Z_j(t)$ for $j \neq i$.
    Get the residual $Z_i(t) = X_i(t) - \sum_{j \neq i} b_{ij}^k(t) X_j^0(t)$.
    Update $V_i = \mathrm{Var}(Z_i(t))$.
  if $\sum_{i=1}^p \|\boldsymbol{\beta}_i^k - \boldsymbol{\beta}_i^{k-1}\|_2 < \epsilon$ **then**
    Break, and set the final estimated $\tilde{X}_i(t) = Z_i(t)$, and $V_i = \mathrm{Var}(\tilde{X}_i(t))$.

becomes further complicated as the number of tuning parameters increases. Here, we propose to follow the tuning procedure for the fused LASSO in [38] to reduce the computation. Specifically, to simplify the search for the optimal tuning parameters, given the sample size L, we reparameterize λ_1 and λ_2 in terms of λ_0 and $\rho \in (0, 1)$ such that $\lambda_1 = \rho\lambda_0$ and $\lambda_2 = (1 - \rho)\lambda_0$. We can think of λ_0 as an overall tuning parameter with ρ determining how much emphasis is placed on sparsity versus smoothness. By fixing the possible values that ρ can take, we effectively reduce the search over λ_1 and λ_2, to a search over one parameter λ_0. In particular, we initially fix the possible values of ρ (e.g., $\{0.1, 0.3, 0.5, 0.7, 0.9\}$). For each value of ρ, we find the value of λ_0 that results in each estimated variable to be 0 and denote this value by $\lambda_{0,\rho}^{\max}$. Then, we chose a fixed number of candidate values for λ_0 from the interval $(0, \lambda_{0,\rho}^{\max})$. For a pair of parameter $\{\lambda_0, \rho\}$ (or $\{\lambda_1, \lambda_2\}$), we input them into Algorithm 1 and get the converged $Z_i(t)$ and V_i. Then, we evaluate this pair of parameters based on the following criterion:

$$L \sum_{i=1}^{p} \log\left(\frac{\sum_{t=1}^{L} Z_i(t)^2}{V_i L}\right) + \log(L) \sum_{i=1}^{p} k_{\rho,\lambda_0}(i) \tag{13}$$

and select the pair of parameters that has the smallest value of (13). In (13), $k_{\rho,\lambda_0}(i)$ is the number of nonzero elements in $\mathbf{b}_i(t), t = 1, \ldots, L$, given the current ρ and λ_0. The term of $\sum_{i=1}^p k_{\rho,\lambda_0}(i)$ represents the complexity of the model, with bigger values indicating bigger complexity. This criterion is similar to the Bayesian information criterion. Its rational is that by minimizing (13), we attempt to find an appropriate model without overfitting the data. The first term will tend to be smaller for complex models, whereas the second term will tend to be smaller for simple models. For computational reasons, we prefer this approach for selecting the optimal tuning parameters.

Remark 3: Algorithm 1 is a standard block coordinate descent algorithm, and its convergence rate depends on the convergence criterion, ϵ, and convergence rate of the solver of each block, i.e., the solver of (12). In general, ϵ is a small positive variable. Here, we set it to be 10^{-3} in all the numerical studies of this article. For the solver of (12), we use the iterative shrinkage thresholding algorithm (ISFA), which is generally used for large-scale LASSO problems [39]. Its convergence rate is $O(1/k)$, where k is the iteration index. To accelerate the convergence rate, other algorithms, such as fast ISFA (FISFA) [40] that can guarantee $O(1/k^2)$ convergence rate, can also be used. For ISTA or FISTA, since its computation only involves matrix and vector multiplication, the computation complexity of each iteration in Algorithm 1 is in the order of $O(L^2 p^3)$, where L is the length of the time window and p is the number of sensors. It is such a small deal for current high-performance computation resources. To illustrate one case with $p = 20$ and $L = 100$ (in our following numerical studies), it only takes less than 2 min to compute one iteration on a single-core personal laptop. The convergence can be achieved very fast as well within 30–40 iterations. As such, the total computation time of Algorithm 1 is not computing intensive. Furthermore, in practice, the wall thickness thinning process is very slow, in units of weeks or even months. What it means is that the time window of online monitoring can also be set in units of weeks or even longer period of time. Therefore, the computation time of modeling inference of (12) based on Algorithm 1 for each time window is almost negligible compared with the length of each time window.

V. Online Monitoring Scheme for Real-Time Single Sensor Anomaly Detection

A. Modeling the Relationship Between the Extracted Feature and the Thickness

Once we extract $\tilde{X}_i(t)$ and V_i, we can infer its relationship with μ_i according to the physical law

$$V_i = \frac{\theta_1}{\mu_i^2} + \theta_2. \tag{14}$$

Here, $\boldsymbol{\theta} = [\theta_1, \theta_2]$ can be easily estimated according to the minimum least square algorithm

$$\min_{\boldsymbol{\theta}} \sum_{i=1}^{p} (V_i - h(\mu_i, \boldsymbol{\theta}))^2, \quad \text{where } h(\mu_i, \boldsymbol{\theta}) = \frac{\theta_1}{\mu_i^2} + \theta_2 \tag{15}$$

with the solved

$$\hat{\boldsymbol{\theta}} = (\mathbf{U}'\mathbf{U})^{-1}(\mathbf{U}'\mathbf{V}) \tag{16}$$

where $\mathbf{U} = [[1/\mu_1^2, \ldots, 1/\mu_p^2]', \mathbf{1}]$ and $\mathbf{1} \in \mathcal{R}^p$ is the vector with all components equal to 1, and $\mathbf{V} = [V_1, \ldots, V_p]'$. Based on the estimated $\hat{\boldsymbol{\theta}}$, for any given tube with wall thickness $\mu_\#$, we can predict its residual variance as $h(\mu_\#, \hat{\boldsymbol{\theta}})$. Furthermore, we can derive its $1 - \alpha$ upper prediction interval

$$\left[0, h(\mu_\#, \hat{\boldsymbol{\theta}}) + q_{1-\alpha}^{t_{p-2}} \sqrt{\sigma_\#^2 + g(\mu_\#)}\right] \tag{17}$$

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

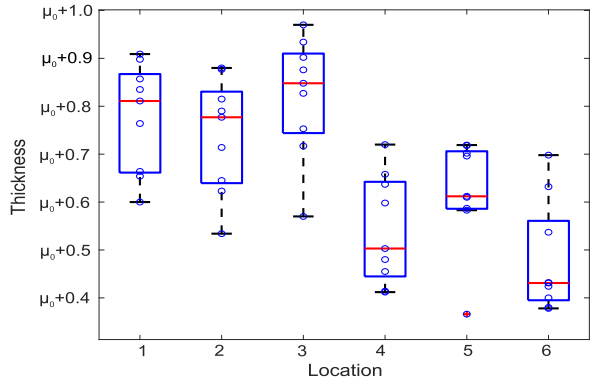


Fig. 4. Multiple thickness measurement for six randomly selected locations of FBG sensors in Fig. 2.

where $g(\mu_\#) = \sigma_\#^2(\mathbf{a}_\#'(\mathbf{A}'\mathbf{A})^{-1}\mathbf{a}_\#)^{1/2}$. Here $\mathbf{a}_\# = ((\partial h(\mu_\#, \boldsymbol{\theta}))/\partial\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$, $\mathbf{A} \in \mathcal{R}^{p\times 2}$ with $A_{i,j} = ((\partial h(\mu_i, \boldsymbol{\theta}))/\partial\theta_j)|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$, $j = 1, 2$, and $\sigma_\#^2 = \sum_{i=1}^{p}(V_i - h(\mu_i, \hat{\boldsymbol{\theta}}))^2/(p-2)$.

Remark 4: It is worth mentioning that in practice, the measured wall thickness with limited measurement precision may introduce additional errors. For example, Fig. 4 shows the repeated measurements (eight or nine times) of tube wall thicknesses of some selected FBG sensors' locations in Fig. 2, whose standard deviations are almost as large as 0.2. In this case, the model of (14) should consider the error in variable to avoid introducing additional noise:

$$\begin{aligned} \mu_{ik} &= s_i + e_{ik} \\ V_i &= h(s_i, \boldsymbol{\theta}) = \frac{\theta_1}{s_i^2} + \theta_2 \end{aligned} \tag{18}$$

where e_{ik} is the unobservable measurement noise for the kth measurement of true wall thickness s_i of sensor i's location. Due to it, the observed thickness μ_{ik} is different from the true one that affects V_i. To solve (18), and we can adopt the approximation estimation method in [41]. In particular, we estimate $\boldsymbol{\theta}$ according to

$$\min_{\boldsymbol{\theta}} \sum_{i=1}^{p}(V_i - \tilde{h}(\bar{\mu}_i, \boldsymbol{\theta}))^2 \tag{19}$$

with

$$\tilde{h}(\bar{\mu}_i, \boldsymbol{\theta}) = h(\bar{\mu}_i, \boldsymbol{\theta}) + \frac{1}{2}\frac{\partial^2 h(\bar{\mu}_i, \boldsymbol{\theta})}{\partial \bar{\mu}_i^2}\sigma_s^2 = \frac{\theta_1}{\bar{\mu}_i^2} + \theta_2 + \frac{3\theta_1}{\bar{\mu}_i^4}\sigma_s^2 \tag{20}$$

where $\bar{\mu}_i$ is the average of repeated wall thickness measurements and σ_s^2 is the variance of e_{ik} and determined by the precision of measurement instrument. Then, this leads to the estimation with error adjustment as

$$\tilde{\boldsymbol{\theta}} = (\tilde{\mathbf{U}}'\tilde{\mathbf{U}})^{-1}(\tilde{\mathbf{U}}'\mathbf{V}) \tag{21}$$

with $\tilde{\mathbf{U}} = [[1/\bar{\mu}_1^2 + 3\sigma_s^2/\bar{\mu}_1^4, \ldots, 1/\bar{\mu}_p^2 + 3\sigma_s^2/\bar{\mu}_p^4]', \mathbf{1}]$. Based on the estimated $\tilde{\boldsymbol{\theta}}$ in (21), for any given tube with wall thickness $\mu_\#$, we can predict its residual variance and the corresponding prediction interval following the similar procedure as in (17). Furthermore, we can derive its $1-\alpha$ upper prediction interval

$$\left[0, h(\mu_\#, \tilde{\boldsymbol{\theta}}) + q_{1-\alpha}^{t_{p-2}}\sqrt{\sigma_\#^2 + \tilde{g}(\mu_\#)}\right] \tag{22}$$

where $\tilde{g}(\mu_\#) = \sigma_\#^2(\tilde{\mathbf{a}}_\#'(\tilde{\mathbf{A}}'\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{a}}_\#)^{1/2}$. Here $\tilde{\mathbf{a}}_\# = ((\partial\tilde{h}(\mu_\#, \boldsymbol{\theta}))/\partial\boldsymbol{\theta})|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}$, and $\tilde{\mathbf{A}} \in \mathcal{R}^{p\times 2}$ with $\tilde{A}_{i,j} = ((\partial\tilde{h}(\mu_i, \boldsymbol{\theta}))/\partial\theta_j)|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}}$, $j = 1, 2$, and $\sigma_\#^2 = \sum_{i=1}^{p}(V_i - \tilde{h}(\mu_i, \tilde{\boldsymbol{\theta}}))^2/(p-2)$.

B. Thinning Process Detection Scheme

Based on the developed model in Section V-A, we can design an online monitoring scheme for the wall thinning process by calculating V_i based on online samples. Remember that since the wall thinning process is very slow compared with the sampling frequency, in Section II, we assume that in the short window with sampling time length L, the tube wall thickness will not change. As such, for online monitoring, we can also divide the sampled FBG signals into sequential short windows of length L, i.e., the nth window $W_n = [L(n-1)+1, nL]$ includes online samples from $t = (n-1)L+1$ to $t = nL$. Then, for each short window, we can fit the data using our spatiotemporal-based signal–noise separation algorithm developed in Section IV and get the residuals' variance for each sensor, denoted as $V_i(n)$. Keeping in mind that as $V_i(n)$ increases, the thickness decreases, and $V_i(n)$ can be used to construct the online monitoring statistic. Suppose that the minimum tolerable tube wall thickness is $\mu^\star$, and then, based on the prediction interval of residual variance in (17) [or (22)], we can get

$$V^\star = h(\mu^\star, \hat{\boldsymbol{\theta}}) + q_{1-\alpha}^{t_{p-2}}\sqrt{\sigma_0^2 + g(\mu^\star)}. \tag{23}$$

Then, if $V_i(n) > V^\star$, it indicates that with enough confidence, the current tube wall thickness of location i is smaller than the minimum tolerable $\mu^\star$. Note that here we construct p univariate monitoring schemes for each sensor (or tube) individually since we assume the erosion of different tubes would not affect each other. However, if erosion of close tubes is driven by some common environmental factors, we recommend to cluster tubes and construct multivariate monitoring schemes for each cluster of tubes jointly. This can be easily achieved by combining $V_i(n)$ of different values of i together into a Hotelling T^2 test statistic. Furthermore, consider that the thickness decrease magnitude is cumulatively large over time, and we further incorporate the exponential weighted moving average (EWMA) technique into the monitoring scheme, to improve the detection efficiency, i.e.,

$$\Gamma_i(n) = (1-\gamma)\Gamma_i(n-1) + \gamma V_i(n) \tag{24}$$

where γ is the tuning parameter. For the EWMA monitoring statistic [42], the corresponding control limit would approximately be $V^\star = h(\mu^\star, \hat{\boldsymbol{\theta}}) + q_{1-\alpha}^{t_{p-2}}((\gamma/(2-\gamma))(\sigma_0^2 + g(\mu^\star)))^{1/2}$. Finally, we define that if $\Gamma_i(n) > V^\star$, the monitoring scheme will trigger an alarm because the current wall thickness of sensor i has dropped below $\mu^\star$ and should require maintenance. The schematic of the monitoring procedure is shown in Fig. 5.

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

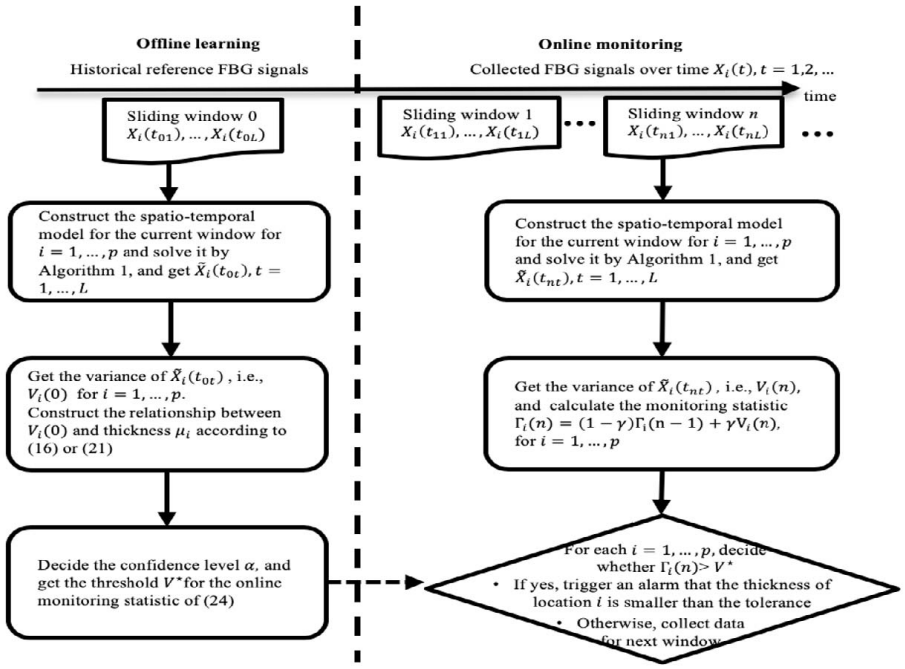


Fig. 5. Schematic of the proposed online monitoring scheme for tube wall thinning process.

VI. Numerical Studies

To evaluate the efficacy of the proposed signal–noise separation algorithm and online monitoring scheme, we generate some synthetic data for analysis.

A. Scenario I

First, we consider a simple scenario with data of only one window, to test the performance of the signal–noise separation model in Algorithm 1. In particular, we assume that for each sensor i, we generate $X_i^0(t)$ as

$$X_i^0(t) = \boldsymbol{\phi}(t)' \mathbf{b}_i^0(t)$$

where $\boldsymbol{\phi}(t) = [\phi_1(t), \ldots, \phi_d(t)]' \in \mathcal{R}^d$ are d basis functions and $\mathbf{b}_i^0(t) = [b_{i1}^0(t), \ldots, b_{id}^0(t)]' \in \mathcal{R}^d$ are the coefficients. In particular, we set $\phi_k(t), k = 1, \ldots, d$, as Fourier bases $\phi_k(t) = \cos(5\pi kt/L), t = 1, \ldots, L$, and we set $\mathbf{b}_i^0(t)$ to change over time with $b_{ik}^0(t) = z_{ik}^0 \cos(\pi t/L), t = 1, \ldots, L$, and $\boldsymbol{\zeta}_k^0 = [z_{1k}^0, \ldots, z_{pk}^0] \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\Sigma}_0$ is the covariance matrix with $\sigma_{ij} = 0.5^{|i-j|}$. The evolving of $\boldsymbol{\phi}(t)$ and $\mathbf{b}_i^0(t)$ over time demonstrates the temporal dynamics, and the correlation of $b_{ik}^0(t), i = 1, \ldots, p$ demonstrates the spatial structure of multiple sensors. We further generate $\tilde{X}_i(t) \sim N(0, \sigma^2 s_i^2)$ where s_i^2 is different for sensors. In particular, we sample s_i from uniform distribution $U(1, 3)$. Then, we set $X_i(t) = X_i^0(t) + \tilde{X}_i(t)$ and use Algorithm 1 for inference.

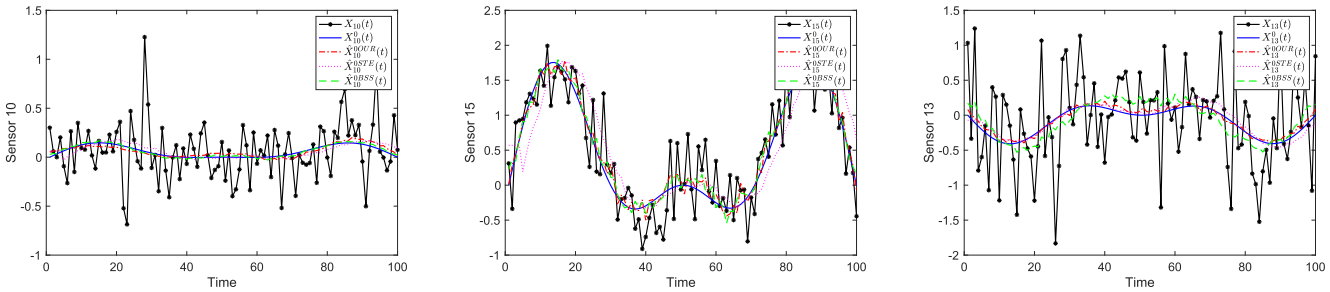
Besides our proposed signal–noise separation Algorithm 1 (in the following, we abbreviate it as OUR), for intensive study, we further consider some baselines in the field of signal source separation and enhancement for performance comparison.

1) *BSS Mentioned in [43]:* The algorithm is based on independent vector analysis (IVA) and nonnegative matrix factorization (NMF). After source separation, we reconstruct the signal part $X_i^0(t)$ and use the left part as noise $\tilde{X}_i(t)$ for inferring V_i.
2) *Spatiotemporal Model-Based Signal Enhancement (STE) Mentioned in [44]:* Since the original algorithm requires knowing the correlation of noise $\tilde{X}_i(t)$, we assume that this information is known in advance. Though this assumption may not be valid in reality, we can still use here and treat the results as the upper limit of the algorithm's performance.
3) *Naive Estimation (NE) Method:* For each window, we further divide the signal into M bins (segments) $l_m, m = 1, \ldots, M$. For each bin, we average the data $X_i(t)$ of each bin as the $X_i^0(t), t \in l_m$ and get the corresponding $\tilde{X}_i(t), t \in l_m$, and then, we use $\tilde{X}_i(t)$ of all the bins together to infer V_i.

Fig. 6 shows the $X_i(t)$, the true $X_i^0(t)$, and the estimated $\hat{X}_i^0(t)$ based on OUR, BSS, and STE (here, we do not show the results of NE since it is a straightforward piecewise estimation) of three selected sensors for the setting with $p = 20, L = 100, d = 3$, and $\sigma = 0.2$. We can see that $\hat{X}_i^{0\ \mathrm{OUR}}(t)$ can be very close to the true $X_i^0(t)$, demonstrating the efficiency of the proposed signal–noise separation algorithm, while $\hat{X}_i^{0\mathrm{BSS}}(t)$ and $\hat{X}_i^{0\mathrm{STE}}(t)$ lose accurate tracking. Then, based on $\hat{\tilde{X}}_i(t) = X_i(t) - \hat{X}_i^0(t)$, we further calculate $\hat{V}_i$ and observe its difference from V_i in Fig. 7. It is clear that among these four methods, $\hat{V}_i^{\mathrm{OUR}}$ has the smallest difference from the true V_i. This validates our idea that the proposed signal–noise separation algorithm can track V_i accurately and

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



(a) (b) (c)

Fig. 6. Synthetic signals of three selected sensors for Scenario I. (a) Sensor 10. (b) Sensor 15. (c) Sensor 13.

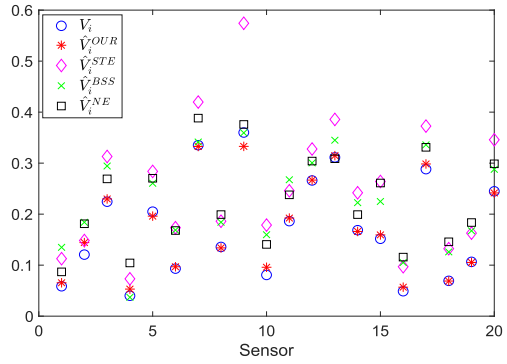


Fig. 7. Estimated residual variance $\hat{V}_i$ and the true values V_i for Scenario I.

TABLE I

ESTIMATION BIAS ($\times 10^{-2}$ AND VARIANCE IN THE PARENTHESIS) OF DIFFERENT METHODS FOR CASE I UNDER DIFFERENT VALUES OF σ

σ	OUR	STE	BSS	NE
0.05	-0.14(0.00)	4.69(0.43)	0.86(0.06)	4.44(0.30)
0.10	0.28(0.00)	3.91(0.30)	1.42(0.34)	4.02(0.28)
0.15	0.59(0.04)	4.26(0.43)	1.11(1.16)	3.11(0.31)
0.20	0.71(0.14)	4.54(0.55)	0.78(1.11)	2.82(0.44)

hence is potential to detect the change of V_i, which will be further evaluated later.

We further evaluate the performance of our signal–noise separation under different values of $\sigma = 0.05, 0.10, 0.15, 0.20$. In each experiment for a particular σ, we random generate $s_i \sim U(1,3), i = 1,\ldots,20$ and conduct experiments following the same procedure as above and then record the estimation results $\hat{V}_i$ of different methods. We repeat the experiment for in total 100 times and use the results to calculate the bias and variance of the estimators $\hat{V}_i$ of different methods. The results are shown in Table I. Clearly, OUR performs dominantly better than the other algorithms under different magnitudes of σ, showing the efficacy of our proposed algorithm.

B. Scenario II

Based on these basic results, we further conduct experiments to imitate the scenario of FBG signals in real apparatus of tube wall thinning process detection. We still temporarily focus on data of single window. In particular, we mimic the temperature and pressure changes in the tube by generating $\Delta P(t) = \sum_{k=1}^{d} a_{1k}(t)\phi(5\pi kt/L)$ and $\Delta T(t) = \sum_{k=1}^{d} a_{2k}(t)\phi(5\pi kt/L)$ for $t = 1,\ldots,L$. We consider $a_{lk}(t) = z_{lk}\cos(\pi t/L), t = 1,\ldots,L$, and $\boldsymbol{\zeta}_k = [z_{1k}, z_{2k}] \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\alpha)$ with $\sigma_{a,ij} = 0.5^{|i-j|}$ for $k = 1,\ldots,d$.

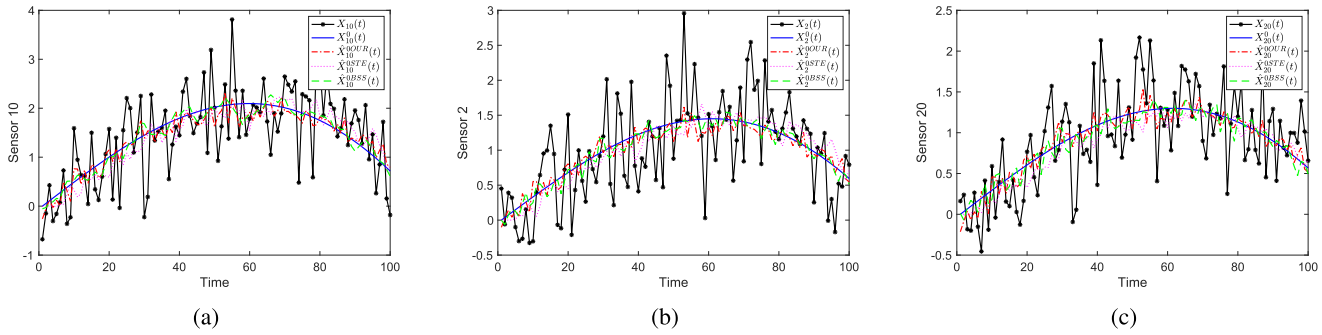
For each sensor, we further mimic its local temperature and pressure variations by generating its own $\tilde{P}_i(t) \sim N(0,\sigma_P^2)$ and $\tilde{T}_i(t) \sim N(0,\sigma_T^2)$ for $t = 1,\ldots,L$. Then, for each sensor, we generate its wall thickness μ_i from uniform distribution $U(1,5)$ and set $X_i^0(t) = \beta_1/\mu_i \Delta P(t) + \beta_2 \Delta T(t)$, $\tilde{X}_i(t) = \beta_1/\mu_i \tilde{P}_i(t) + \beta_2 \tilde{T}_i(t)$, where β_1 and β_2 are fixed constants. Then, we combine $X_i(t) = X_i^0(t) + \tilde{X}_i(t)$ and use OUR together with the three baselines for inference.

Fig. 8 shows the $X_i(t)$, the true $X_i^0(t)$, and the estimated ones in our simulation replication for the case with the setting $p = 20, L = 100, q = 3, \sigma_T = \sigma_P = 0.5, \beta_1 = 4$, and $\beta_2 = 1$. In this case, all the methods seem to be able to track $X_i^0(t)$ to some degree. Yet, they still have larger fluctuations, and our method $\hat{X}_i^{0\mathrm{OUR}}(t)$ has the smallest distance from the true $X_i^0(t)$. This can be better shown in Fig. 9, where based on $\hat{\tilde{X}}_i(t) = X_i(t) - \hat{X}_i^0(t)$, we calculate $\hat{V}_i, i = 1,\ldots,p$ and use them to estimate θ_1 and θ_2 in (15) according to (16). In particular, $\hat{V}_i^{\mathrm{OUR}}$ is closest to V_i. Yet $\hat{V}_i^{\mathrm{BSS}}$ and $\hat{V}_i^{\mathrm{NE}}$ have larger fluctuations around V_i, and $\hat{V}_i^{0\ \mathrm{STE}}$ is consistently larger than V_i^0. Furthermore, it is clear that the estimated regression line of OUR is quite close to the line of true relationship, demonstrating that the proposed method can recover the relationship between μ_i and V_i. Yet, for the other methods, their lines deviate from the true line, especially for thinner wall thicknesses, which is the most crucial part required to be accurately estimated and monitored. To further verify the accuracy of the proposed method, we repeat the experiment for 100 replicates. Each replicate includes data generation, signal–noise separation, and regression estimation. Based on the estimated $\hat{V}_i, \hat{\theta}_1$, and $\hat{\theta}_2$, we calculate that the estimation bias and variance of $\hat{V}_i$, $\hat{\theta}_1$ and $\hat{\theta}_2$ of different methods and report them in Table II. Clearly, OUR has much smaller bias than the other methods for V_i and consequently much smaller bias for θ_1 and θ_2. This builds the foundation of applying OUR algorithm for thinning process monitoring.

Finally, we also evaluate the performance of our different algorithms under different values of (σ_T, σ_P). In each

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



(a) (b) (c)

Fig. 8. Synthetic signals of three selected sensors for Scenario II. (a) Sensor 10. (b) Sensor 2. (c) Sensor 20.

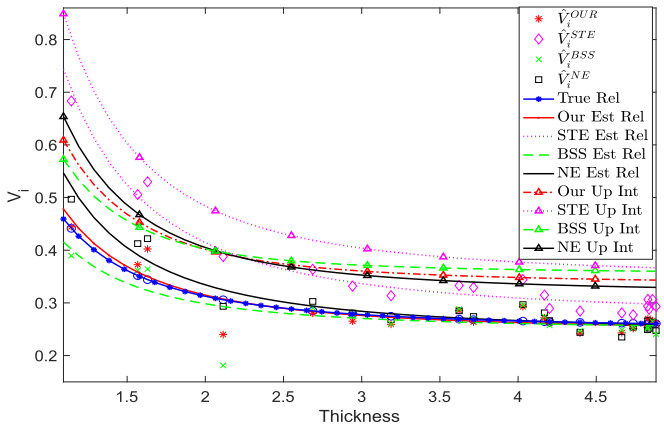


Fig. 9. Scenario II: synthetic thickness μ_i versus the estimated residual variance $\hat{V}_i$. The lines of different colors are the fitted relationships between V_i and μ_i using $\hat{V}_i$ based on different numerical methods. The lines with triangle markers are their corresponding 0.99 upper prediction intervals based on different numerical methods; The blue line with "*" marker is the true relationship between V_i and μ_i.

TABLE II

ESTIMATION BIAS ($\times 10^{-2}$ AND VARIANCE IN THE PARENTHESIS) OF DIFFERENT METHODS FOR CASE II UNDER DIFFERENT VALUES OF σ_T AND σ_P

(σ_T, σ_P)	Para	OUR	STE	BSS	NE
$(0.25, 0.25)$	$\hat{V}_i$	0.34(0.00)	6.67(0.65)	1.11(0.06)	2.63(0.11)
	$\hat{\theta}_1$	2.01(0.03)	22.6(5.49)	6.27(0.23)	10.9(1.10)
	$\hat{\theta}_2$	0.35(0.00)	1.30(0.06)	2.37(0.02)	0.70(0.02)
$(0.50, 0.50)$	$\hat{V}_i$	1.20(0.00)	7.43(1.78)	2.82(0.32)	2.76(1.03)
	$\hat{\theta}_1$	3.42(0.03)	25.4(7.68)	24.1(2.37)	14.3(2.14)
	$\hat{\theta}_2$	1.95(0.00)	1.45(0.07)	7.87(0.15)	-0.30(0.02)
$(0.75, 0.75)$	$\hat{V}_i$	1.38(0.33)	10.9(4.05)	7.89(1.19)	3.02(3.99)
	$\hat{\theta}_1$	-5.77(2.81)	28.3(11.0)	58.5(7.66)	11.3(2.23)
	$\hat{\theta}_2$	2.64(0.03)	1.36(0.11)	17.5(0.71)	-1.56(0.03)

experiment for a particular pair of (σ_T, σ_P), we randomly generate $\mu_i \sim U(1,5), i = 1, \ldots, 20$ and conduct experiments following the same procedure as above, we record the estimation results $\hat{V}_i$ of different methods and further use $\hat{V}_i$ to get the regression coefficients $\hat{\theta}_1$ and $\hat{\theta}_2$. We repeat the experiment for in total 100 times and use the results to calculate the bias and variance of the estimators $\hat{V}_i$ of different methods, $\hat{\theta}_1$ and $\hat{\theta}_2$. The results are also tabulated in Table II. Similar to Scenario I, OUR performs dominantly better than the other algorithms under different magnitudes of (σ_T, σ_P).

C. Scenario III

Finally, we set up an online thinning scenario for the experiment. It is to be noted that from the first two scenarios, we have demonstrated that the baseline algorithms cannot estimate V_i and the relationship between V_i and μ_i accurately. Consequently, their online detection schemes have too poor performance to be presented here. As such, we only show the performance results of our detection scheme in the following. In particular, in each simulation replicate example, for the first window with $n = 0$, we randomly get the wall thicknesses of the 20 sensors by sampling from the uniform distribution $U(1,5)$ and generate the data in the same way as Scenario II. Then, for later windows $n = 1, \ldots, 100$, we mimic the wall thinning process by setting $\mu_i(n) = \max\{\mu_i - 0.03n, 0.1\}$, which means that the wall thicknesses of all the sensors decrease with speed 0.03 per time window and stop when they decrease to 0.1. Then, we use the proposed online monitoring scheme to detect the change of $\mu_i(n)$. In particular, we evaluate its performance in terms of five different criteria: 1) the average detection delay is ADD $= E[s_i - \tau_i | s_i \geq \tau_i]$, where $\tau_i = \inf\{n > 0, \mu_i(n) < \mu^\star\}$ and $s_i = \inf\{\Gamma_i(n) > V^\star, n > \tau_i\}$; 2) the average false positive number is AFP $= E[\sum_{n=1}^{s_i-1} I(\Gamma_i(n) > V^\star)]$, and here, $I(\cdot) = 1$ if the condition is satisfied and $I(\cdot) = 0$ otherwise; 3) the average false negative number is AFN $= E[\sum_{n=s_i}^{N} I(\Gamma_i(n) < V^\star)]$; 4) the Recall $= E[(\sum_{n=s_i}^{N} I(\Gamma_i(n) > V^\star))/(N - s_i + 1)]$; and 5) the Precision $= E[(\sum_{n=s_i}^{N} I(\Gamma_i(n) > V^\star))/(\sum_{n=1}^{N} I(\Gamma_i(n) > V^\star))]$. The first three criteria focus on evaluating the sequential decision performance, whereas the last two criteria focus on evaluating the overall performance. In our numerical studies, we consider different parameter settings of $\mu^\star$, α (the corresponding $V^\star$), and γ. For each artificial setting $r = 1, \ldots, 5$, we conduct the experiments for 100 replicates and calculate their five criteria, namely ADD, AFP, AFN, Recall, and Precision, as shown in Table III.

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE III

DETECTION PERFORMANCE WITH $\gamma = 0.1$ AND $\gamma = 0.5$ FOR DIFFERENT CHANGE ARTIFICIAL SETTINGS (NUMBERS IN THE PARENTHESES ARE THE CORRESPONDING STANDARD DEVIATIONS)

		$\gamma = 0.1$					$\gamma = 0.5$				
$\mu^\star$	α	ADD	AFP	AFN	Recall	Precision	ADD	AFP	AFN	Recall	Precision
2.5	0.50	1.80(0.98)	17.5(7.51)	2.64(1.70)	0.96(0.02)	0.78(0.08)	1.45(0.26)	16.7(6.49)	3.39(1.52)	0.95(0.02)	0.78(0.08)
	0.60	2.94(2.32)	11.8(7.29)	4.47(2.97)	0.94(0.04)	0.84(0.09)	1.82(0.49)	13.0(6.12)	5.26(2.20)	0.92(0.03)	0.82(0.08)
	0.70	5.13(3.50)	5.55(5.55)	7.89(4.15)	0.89(0.06)	0.91(0.08)	2.49(0.98)	9.24(5.30)	7.71(2.83)	0.89(0.04)	0.86(0.07)
	0.80	9.27(4.33)	1.66(2.70)	13.0(4.61)	0.81(0.07)	0.96(0.04)	3.65(1.65)	5.65(4.07)	11.1(3.33)	0.84(0.05)	0.91(0.06)
	0.90	16.0(4.07)	0.22(0.64)	20.1(4.06)	0.71(0.06)	0.99(0.01)	6.04(2.71)	2.33(2.27)	15.9(3.43)	0.77(0.05)	0.95(0.04)
	0.95	21.1(3.66)	0.03(0.10)	24.7(3.17)	0.64(0.06)	1.00(0.00)	9.19(3.46)	1.05(1.33)	19.8(3.22)	0.71(0.05)	0.98(0.03)
	0.99	28.4(3.08)	0.00(0.00)	30.8(2.51)	0.55(0.05)	1.00(0.00)	16.7(4.31)	0.19(0.32)	25.7(2.66)	0.63(0.05)	1.00(0.01)
2	0.50	1.38(0.71)	19.8(10.6)	1.42(1.04)	0.98(0.02)	0.73(0.12)	1.19(0.19)	20.9(8.72)	1.80(0.90)	0.97(0.01)	0.71(0.11)
	0.60	2.15(1.56)	12.2(9.36)	2.53(1.95)	0.96(0.03)	0.81(0.12)	1.44(0.27)	16.2(8.23)	2.73(1.34)	0.95(0.02)	0.76(0.11)
	0.70	3.37(2.19)	6.22(6.96)	4.19(2.71)	0.93(0.04)	0.89(0.10)	1.80(0.57)	11.6(7.18)	3.97(1.83)	0.93(0.03)	0.81(0.10)
	0.80	5.79(2.94)	2.37(4.05)	7.23(3.48)	0.88(0.06)	0.95(0.07)	2.43(0.91)	7.29(5.54)	5.78(2.37)	0.90(0.04)	0.87(0.09)
	0.90	10.1(3.33)	0.52(1.49)	11.9(3.60)	0.79(0.06)	0.98(0.03)	3.94(1.70)	3.43(3.41)	8.77(2.81)	0.85(0.04)	0.93(0.07)
	0.95	13.4(3.37)	0.13(0.67)	15.3(3.40)	0.73(0.05)	0.99(0.02)	5.58(2.37)	1.73(2.19)	11.2(2.93)	0.80(0.04)	0.96(0.05)
	0.99	18.6(3.15)	0.00(0.00)	20.3(3.08)	0.64(0.05)	1.00(0.00)	9.56(3.24)	0.42(0.76)	15.7(2.98)	0.72(0.04)	0.99(0.02)
1.5	0.50	1.14(0.43)	14.1(11.0)	0.84(0.58)	0.98(0.01)	0.75(0.15)	1.00(0.15)	19.5(10.3)	0.82(0.46)	0.98(0.01)	0.69(0.14)
	0.60	1.48(0.80)	9.15(8.92)	1.32(1.01)	0.97(0.02)	0.82(0.14)	1.10(0.20)	15.1(9.29)	1.11(0.66)	0.98(0.01)	0.74(0.14)
	0.70	2.10(1.32)	5.42(6.59)	2.06(1.56)	0.96(0.03)	0.88(0.12)	1.25(0.31)	11.3(7.92)	1.52(0.94)	0.97(0.02)	0.79(0.13)
	0.80	3.18(1.97)	2.84(4.60)	3.32(2.19)	0.93(0.04)	0.93(0.09)	1.53(0.48)	7.68(6.31)	2.22(1.27)	0.95(0.02)	0.84(0.12)
	0.90	5.16(2.64)	0.99(2.48)	5.55(2.79)	0.88(0.05)	0.97(0.06)	0.94(0.37)	4.66(3.80)	0.43(0.49)	0.99(0.01)	0.86(0.11)
	0.95	7.04(2.80)	0.34(1.30)	7.53(2.99)	0.83(0.05)	0.99(0.03)	2.73(1.23)	2.64(3.06)	4.60(2.15)	0.90(0.04)	0.93(0.07)
	0.99	10.4(3.06)	0.06(0.32)	10.9(3.20)	0.76(0.05)	1.00(0.01)	4.58(2.23)	0.93(1.45)	7.05(2.60)	0.84(0.04)	0.97(0.04)
1	0.50	0.92(0.35)	6.12(6.17)	0.24(0.32)	0.99(0.01)	0.83(0.14)	0.72(0.14)	11.3(7.52)	0.05(0.09)	1.00(0.00)	0.74(0.14)
	0.60	1.11(0.59)	4.71(5.15)	0.43(0.55)	0.99(0.02)	0.86(0.12)	0.75(0.17)	9.55(6.54)	0.10(0.15)	1.00(0.00)	0.77(0.14)
	0.70	1.37(0.85)	3.49(4.30)	0.70(0.83)	0.98(0.02)	0.89(0.11)	0.78(0.20)	7.92(5.71)	0.15(0.21)	1.00(0.01)	0.80(0.13)
	0.80	1.72(1.17)	2.43(3.37)	1.06(1.14)	0.97(0.03)	0.92(0.10)	0.82(0.25)	6.37(4.77)	0.24(0.31)	0.99(0.01)	0.83(0.12)
	0.90	2.35(1.60)	1.40(2.38)	1.69(1.57)	0.95(0.05)	0.95(0.08)	0.97(0.42)	4.10(3.60)	0.41(0.50)	0.99(0.01)	0.88(0.10)
	0.95	2.99(1.86)	0.86(1.77)	2.35(1.82)	0.93(0.05)	0.97(0.06)	1.04(0.45)	3.54(3.13)	0.61(0.62)	0.98(0.02)	0.89(0.09)
	0.99	4.17(2.18)	0.31(0.93)	3.55(2.15)	0.89(0.06)	0.99(0.03)	1.46(0.80)	2.11(2.14)	1.16(1.02)	0.97(0.03)	0.93(0.07)
0.5	0.50	1.13(0.76)	1.07(1.48)	0.55(0.71)	0.98(0.03)	0.94(0.08)	0.72(0.14)	11.3(7.52)	0.05(0.09)	1.00(0.00)	0.74(0.14)
	0.60	1.17(0.80)	0.96(1.37)	0.60(0.75)	0.97(0.03)	0.95(0.08)	0.60(0.15)	3.68(2.19)	0.01(0.03)	1.00(0.00)	0.85(0.10)
	0.70	1.22(0.84)	0.88(1.29)	0.65(0.79)	0.97(0.03)	0.95(0.08)	0.60(0.15)	3.51(2.10)	0.01(0.04)	1.00(0.00)	0.85(0.10)
	0.80	1.29(0.87)	0.78(1.22)	0.71(0.82)	0.97(0.03)	0.96(0.07)	0.60(0.15)	3.30(2.01)	0.02(0.05)	1.00(0.00)	0.86(0.10)
	0.90	1.39(0.92)	0.66(1.07)	0.82(0.87)	0.96(0.04)	0.96(0.07)	0.61(0.16)	3.04(1.86)	0.03(0.05)	1.00(0.00)	0.87(0.09)
	0.95	1.48(0.96)	0.57(0.97)	0.91(0.92)	0.96(0.04)	0.97(0.06)	0.62(0.16)	2.85(1.77)	0.03(0.07)	1.00(0.00)	0.87(0.09)
	0.99	1.61(1.01)	0.42(0.80)	1.05(0.96)	0.95(0.04)	0.97(0.05)	0.63(0.17)	2.49(1.66)	0.05(0.09)	1.00(0.00)	0.89(0.09)

First, for all the artificial settings of γ and $\mu^\star$, as α increases with larger $V^\star$, ADD and AFN increase, while AFP decreases. This is easily interpretable since larger $V^\star$ indicates smaller $P\{\Gamma_i(n) > V^\star\}$ and consequently larger ADD and AFN and smaller AFP. Similarly, as α increases, we can get lower Recall and higher Precision. Another thing to be noted is that for all the settings of γ, if restricting AFP to be fixed, as $\mu^\star$ decreases, the ADD and AFN decrease. This desirable phenomenon is not very intuitive. It is a merit of the quadratic property of $V = h(\mu, \boldsymbol{\theta})$: as $\mu^\star$ decreases and $V^\star$ increases, the decreasing speed of $\int_{\mu<\mu^\star} P(h(\mu, \boldsymbol{\theta}) > V^\star) d\mu$ is smaller than the decreasing speed of $\int_{\mu>\mu^\star} P(h(\mu, \boldsymbol{\theta}) > V^\star) d\mu$. Finally, as $\mu^\star$ decreases, given a fixed AFP, ADD and AFN decrease moderately. Similarly, given a fixed Precision value, Recall increases.

The influence of γ is a bit complex. First, for a certain $\mu^\star$, given the same α level, larger γ leads to smaller ADD and AFN, but larger AFP. This is reasonable since larger γ puts more stress on the current single window and hence is more aggressive and unstable with larger variation. Therefore, $\Gamma_i(n)$ is more probable to go out of $V^\star$ and leads to either earlier true positive detection or more false-positive detection. It is also interesting to observe that for the monitoring schemes with different values of γ, even if we restrict their AFP to be

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

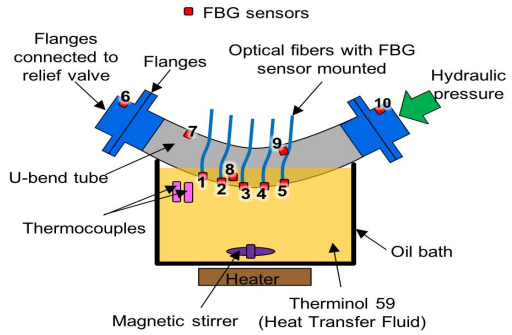


Fig. 10. Schematic of the oil bath test apparatus.

the same by setting their corresponding α to be different, their ADD and AFN are still slightly different in the sense that the scheme with larger γ would have smaller ADD and yet larger AFN. In practice, γ can be selected based on the preference of either aggressive detection or conservative detection.

VII. Case Studies

In this section, we apply the proposed methodology to tube wall thinning detection with both laboratory experiments and real field trials.

A. Lab Experiment via Oil Bath Test

Before applying the designed modeling and monitoring scheme in the real plant environment, we first conduct a thorough laboratory test for performance evaluation. To imitate the real plant environment, we design our laboratory test apparatus, as shown in Fig. 10.

In particular, the apparatus includes a section of the tube submerged in a heating pot. The material of the tube is SA213, which is a type of American standard seamless alloy steel pipe, typically employed for boiler tubes used in oil and refinery industries. We mount in total $p = 10$ FBG sensors at ten different locations of the tube. In the experiment, Sensors 1–5 are fixed with adhesive at the section of the tube where there is noticeably localized inner wall thinning. The other sensors are fixed elsewhere with no changes in the wall thickness for the whole experiment.

In our experiment, we artificially accelerate the tube wall thinning process via etching the inner wall of the tube for several runs. For each run, we inject aqua regia into the tube and let the aqua regia react with the bottom part (i.e., the bending part) of the tube until the reaction stops. The bottom wall will become thinner and we measure the wall thicknesses at the locations of Sensors 1–5 with repeated measurements. The measurement results after each of in total five etching runs are shown in Table IV. After each run of etching, we place the tube on a heating pot, which is filled with Therminol 59 (termed as oil for the rest of this article). The bending section of the tube is submerged in the oil bath, which can create a relatively stable and controlled high-temperature environment surrounding the tube. To mimic the operation environment of the tube in the real plant, we heat the oil up to 150 °C in the meanwhile pressurize the tube with water

TABLE IV

Averaged Thickness, $\bar{\mu}_i = \mu_0 + \mu_i^e$, for Sensors $i = 1, 2, 3, 4, 5$ From Repeated Measurements After Each of the Five Runs (Numbers in Parentheses Are Standard Deviations of Repeated Measurements)

	μ_i^e				
Etching Run	$i=1$	2	3	4	5
$e=1$	0.495 (0.086)	0.396 (0.114)	0.394 (0.109)	0.394 (0.083)	0.484 (0.076)
2	0.115 (0.085)	-0.005 (0.068)	0.015 (0.084)	-0.040 (0.110)	0.110 (0.090)
3	-0.168 (0.057)	-0.433 (0.100)	-0.421 (0.099)	-0.474 (0.093)	-0.225 (0.055)
4	-0.611 (0.131)	-0.877 (0.133)	-0.889 (0.103)	-0.945 (0.099)	-0.665 (0.077)
5	-0.713 (0.027)	-1.180 (0.116)	-1.192 (0.110)	-1.229 (0.093)	-0.864 (0.045)

by a hydraulic pressure pump. Then, we collect the FBG signals for a certain time period (about 1–2 h) for analysis. To evaluate the performance of the proposed online monitoring scheme, we treat the signals of the five runs as data coming sequentially, with stepwise thinning process. In particular, we set the window size $L = 50$. In reality, we recommend to set $L = 2880$, i.e., one day. This sample size can provide sufficient samples for accurate model estimation from the statistical perspective. However, from a practical perspective, if the tube thinning process is very slow and real-time data transmission is limited, a longer window size (say one week) is also feasible. Then, for each run, we have different numbers of windows, i.e., $N_1 = 3, N_2 = 4, N_3 = 3, N_4 = 4$, and $N_5 = 3$ (note that we discard the signals at the end of each run if they are not enough to form a window). Then, we cascade signals of all these five runs together sequentially and get in total $N = \sum_{i=1}^{5} N_i = 17$ time windows. For each window n, we implement the spatiotemporal model for the $p = 10$ sensors and get their $V_i(n), i = 1, \ldots, p$. However, since we only measured the thicknesses of Sensors 1–5 where the thinning process occurred, we use these sensors' $V_i(n)$ and μ_i for further analysis. The relationship between $V_i(n)$ and $\mu_i(n)$ for all the $5 \times 17 = 85$ points is shown in Fig. 11. Clearly, there is a general pattern that as $\mu_i(n)$ decreases, $V_i(n)$ increases. Suppose that $\mu^\star = \mu_0 - 1$ is the minimum tolerable thickness, and then, the thicknesses of all the sensors in first four runs are tolerable, whereas the thicknesses of Sensors 2–4 in the last run are out of tolerance and to be detected. We use the data of the first four runs to infer the relationship between $\mu_i(n)$ and $V_i(n)$ and use the data of the last run for online detection. The fitted regression using $V_i(n), i = 1, \ldots, 5$, of $n = 1, \ldots, 14$, together with its 0.99 upper prediction interval is shown in Fig. 11. $R^2 = 0.691$. Based on the upper prediction interval for $\mu^\star$, we can get the detection threshold $V^\star = 1.8726 \times 10^{-4}$. Then, we implement the monitoring scheme of (24) for these five locations.

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

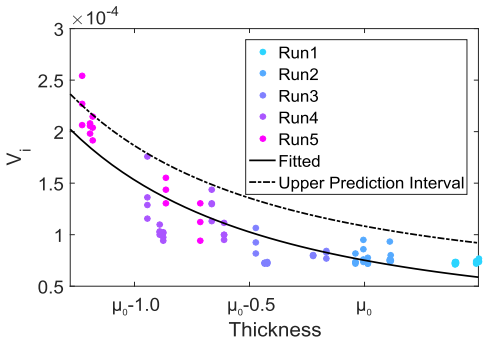


Fig. 11. $\mu_i(n)$ versus estimated $V_i(n)$ based on our proposed signal–noise separation model for the oil bath test data. The black line is the fitted relationship based on (21). Its $R^2 = 0.691$. The dashed line is the corresponding 0.99 upper prediction interval calculated by (22).

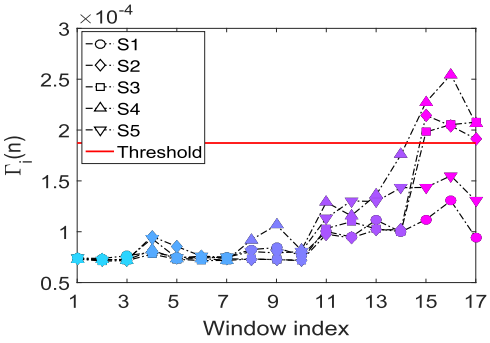


Fig. 12. Monitoring statistic (24) of Sensors 1, 2, 3, 4, and 5 over $n = 1, \ldots, 17$. The red line is the threshold $V^\star$.

The monitoring statistic $\Gamma_i(n)$ for $i = 1, \ldots, 5$ and $n = 1, \ldots, 17$ is shown in Fig. 12. For the first four runs with the first $n = 1, \ldots, 14$, the monitoring statistics of all the five sensors keep consistently below the threshold, whereas for the last $n = 15, 16$, and 17, the monitoring statistics of Sensors 2–4 go out of the threshold, triggering an alarm that their tube wall thicknesses are below $\mu^\star$, whereas the monitoring statistics of Sensors 1 and 5 are still below the threshold. This is consistent with the real situation that their thicknesses of Sensors 1 and 5 are still above $\mu^\star$, whereas the thicknesses of Sensor 2–4 have dropped below $\mu^\star$, demonstrating the efficiency of the proposed monitoring scheme.

B. Field Trial in HRSG

We have investigated the performance of the proposed methodology in lab experiments. In this section, we carry out field trial to further evaluate the implementation applicability of the proposed methodology for tube wall thinning detection in the HRSG plant, where the wall thinning of tubes is one of the common causes of failures in practice.

In our field trial, we choose a representative HRSG site, and $p = 13$ tubes connected to the same compartment as Fig. 1(a) are mounted with FBG sensors and monitored. Due to confidentiality and security reasons, we cannot present details about the plant, the tube structure, or the mounting process. The mounting procedure for the FBG sensors onto the tube is with high resemblance to those we used for laboratory environment test in Section VII-A. The wall thicknesses of tubes (location), before the FBG sensors are mounted, are measured at the beginning of the experiment and are presented in Table V (due to confidential reason, we hide the measurement unit here and mask the thickness magnitude by setting it as a reference thickness of μ_0 plus a deviation).

TABLE V

THICKNESSES, $\mu_i = \mu_0 + \mu_i^e$, OF $p = 13$ TUBES IN THE FIELD TRIAL (NUMBERS IN PARENTHESES ARE STANDARD DEVIATIONS OF REPEATED MEASUREMENTS)

Location	1	2	3	4	5	6	7
μ_i^e	0.83 (0.15)	0.78 (0.13)	0.62 (0.11)	0.72 (0.13)	0.39 (0.14)	0.55 (0.11)	0.64 (0.12)
Location	8	9	10	11	12	13	
μ_i^e	0.62 (0.14)	0.54 (0.13)	0.65 (0.11)	0.71 (0.15)	0.48 (0.13)	0.66 (0.13)	

The FBG sensors have a sampling rate of 1 KHz, with in total data of about 19 days for the whole test duration. We first preprocess the data by downsampling the data to $1/30$ Hz by averaging samples of each half minute for system noise reduction. For the downsampled data, we treat the window size $L = 720$ (6 h), with in total 76 windows available for the 19 days. Denote their indices as $n = 0, 1, \ldots, 75$.

Then, we use window $n = 0$ for offline learning to estimate the relationship between $V_i(0), i = 1, \ldots, p$ and $\mu_i \equiv \mu_i(0)$ (because there is no thinning process in these 19 days) based on our signal–noise separation algorithm. From Fig. 13, we can see that the spatiotemporal model can fit the original FBG signals of window $n = 0$ very well. Then, we calculate the regression residuals and their variance $V_i(0), i = 1, \ldots, p$, and draw them together with their corresponding thicknesses in Fig. 14. Obviously, the variance of FBG sensor decreases as its thickness increases. This is consistent with our expectations. The regression result with consideration of error in a variable is shown in Fig. 14 with black line, and its corresponding 0.99 upper prediction interval is shown in the dashed line. $R^2 = 0.753$.

Suppose that the minimum tolerable thickness of the HRSG is $\mu^* = \mu_0 + 0.4$, and we can get the monitoring threshold $V^\star$ based on a specific α according to (23). We treat the FBG signals with $n = 1, \ldots, 75$ as online testing windows and implement the online monitoring scheme in the following way. In particular, we calculate $\Gamma_i(n), n = 1, \ldots, 75$, for each sensor $i = 1, \ldots, p$, and see whether $\Gamma_i(n)$ goes out of the threshold. Ideally, for signals of FBG sensor $i = 5$ with $\mu_5 = \mu_0 + 0.39$ (the dot in the top-left corner of Fig. 14), its $\Gamma_5(n)$ is expected to be higher than $V^\star$ for most windows, whereas for the other sensors with $\mu_i > \mu_0 + 0.4, i \neq 5$, their $\Gamma_i(n)$ are expected to be smaller than $V^\star$. As such, we evaluate the monitoring performance in terms of the true detection rate (TDR) of $i = 5$ and the false detection rate (FDR) of $i \neq 5$. The results under $V^\star$ with different values of α are shown in Fig. 15. It is clear that the TDR of $i = 5$ is almost equal to 1 all the way, whereas FDRs of all the other sensors

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

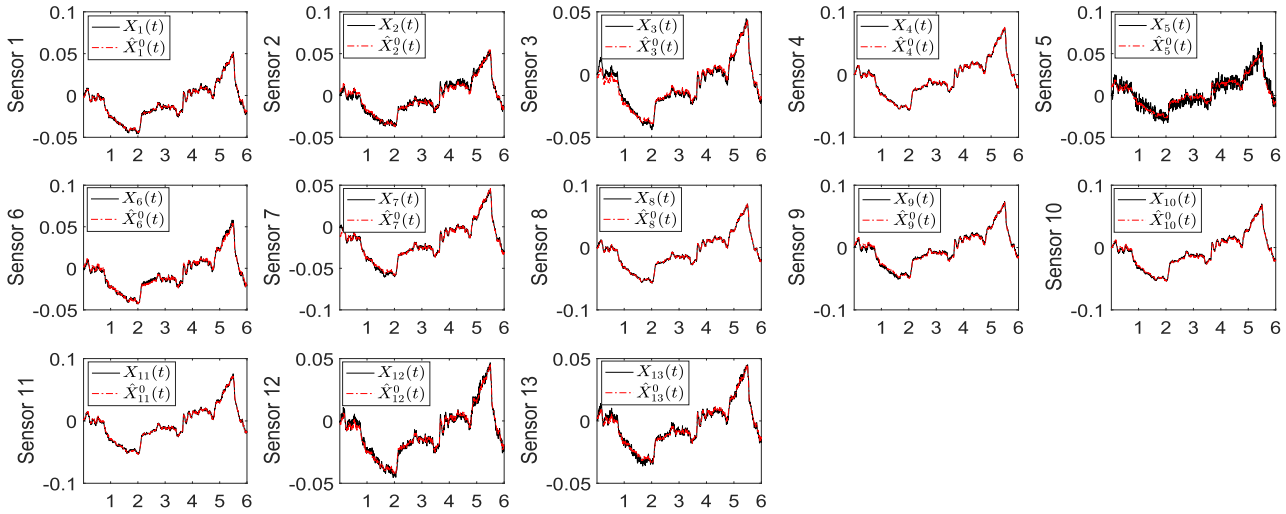


Fig. 13. Prediction of the FBG wavelength drift based on our proposed signal–noise separation model for window $n = 0$ for the field trial data.

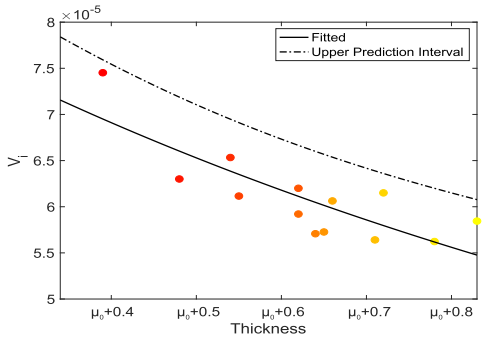


Fig. 14. μ_i versus the estimated $V_i(0), i = 1, \ldots, p$ for the field trial data. The black line is the fitted regression based on (19). Its $R^2 = 0.753$. The dashed line is the corresponding 0.99 upper prediction interval calculated by (22).

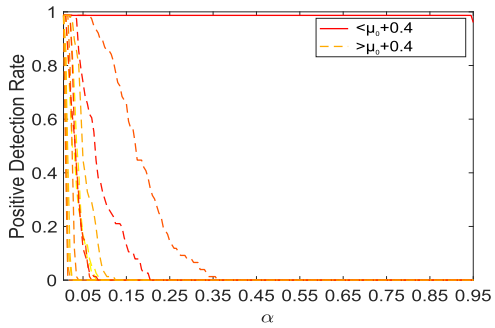


Fig. 15. Detection rate of the monitoring statistic (24) with different values of $V^\star$ (under different values of α) for the field trial data.

almost go to zero quickly as α increases, demonstrating the efficiency of the proposed monitoring scheme. In practice, we may specify α based on domain or prior knowledge.

VIII. Conclusion

We propose a data-driven framework for online monitoring of tube wall thinning process using FBG sensors in a dynamic and noisy environment. In particular, we develop a spatiotemporal model for FBG signal feature extraction. It can remove the environmental influence and extract the FBG signal feature that relates to tube wall thickness. Then, we construct the relationship between the selected feature and the tube wall thickness by taking physical law as a guideline. Finally, an online monitoring scheme is constructed to monitor the extracted feature in real time to detect thickness change. Both the laboratory test and the field trial experiment are conducted to demonstrate the efficacy and efficiency of the proposed methodology.

There are many potential extensions of the current method. First, in some cases, when the tube diameter is large, it is more reasonable to mount more than one FBG sensor on the same tube to jointly detect the erosion of different parts of the same tube. As FBG sensors on the same tube usually share the same environmental conditions and have similar wall thicknesses (thinning process), it would be better to analyze FBG sensors on the same tube together and extend the current methodology to a multivariate sensor-based online monitoring scheme, which considers the correlations of FBG sensors on the same tube. Second, in some cases, the global temperature $\Delta P(t)$ and pressure $\Delta T(t)$ have time delays on different tubes as $\Delta P(t - s_{iP})$ and $\Delta T(t - s_{iT})$. Here, s_{iP} and s_{iT} are the local time delays of $\Delta P(t)$ and $\Delta T(t)$ on Tube i, respectively. Then, the self-expressive model of (7) should consider signals of other time points of other sensors as predictors, and a more elegant model should be considered. In functional regression, there are also some works to address this problem by taking the integral of the predictor over the past [45]. How to incorporate this time delay into the modeling framework also deserves further exploration.

References

[1] Y. Huang and D. Ji, "Experimental study on seawater-pipeline internal corrosion monitoring system," *Sens. Actuators B, Chem.*, vol. 135, no. 1, pp. 375–380, Dec. 2008.

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[2] X. Li and H. Castaneda, "Coating studies of buried pipe in soil by novel approach of electrochemical impedance spectroscopy at wide frequency domain," *Corrosion Eng., Sci. Technol.*, vol. 50, no. 3, pp. 218–225, May 2015.

[3] G. Rocchini, "Some considerations on the polarization resistance method," *Corrosion Sci.*, vol. 41, no. 12, pp. 2353–2367, Dec. 1999.

[4] T. Jiang, L. Ren, Z.-G. Jia, D.-S. Li, and H.-N. Li, "Pipeline internal corrosion monitoring based on distributed strain measurement technique," *Struct. Control Health Monitor.*, vol. 24, no. 11, Nov. 2017, Art. no. e2016.

[5] S. Li, Y.-G. Kim, S. Jung, H.-S. Song, and S.-M. Lee, "Application of steel thin film electrical resistance sensor for *in situ* corrosion monitoring," *Sens. Actuators B, Chem.*, vol. 120, no. 2, pp. 368–377, Jan. 2007.

[6] D. Sonyok, B. Zhang, and J. Zhang, "Applications of non-destructive evaluation (NDE) in pipeline inspection," in *Proc. ASCE Int. Pipelines Conf., Pipeline Asset Manage., Maximizing Perform. Pipeline Infrastruct.*, Atlanta, GA, USA, Jul. 2008, pp. 1–10.

[7] U. D. Black, *Data Communications and Distributed Networks*. Upper Saddle River, NJ, USA: Prentice-Hall, 1987.

[8] J. M. Senior and M. Y. Jamro, *Optical Fiber Communications: Principles and Practice*. London, U.K.: Pearson Education, 2009.

[9] H.-N. Li, D.-S. Li, and G.-B. Song, "Recent applications of fiber optic sensors to health monitoring in civil engineering," *Eng. Struct.*, vol. 26, no. 11, pp. 1647–1657, Sep. 2004.

[10] K. D. Bennett and L. McLaughlin, "Monitoring of corrosion in steel structures using optical fiber sensors," *Proc. SPIE*, vol. 2446, pp. 48–60, Apr. 1995.

[11] L. Zou *et al.*, "Pipeline corrosion monitoring by fiber optic distributed strain and temperature sensors," in *Proc. NACE Int. Corrosion Conf. Expo*, New Orleans, LA, USA, Mar. 2008, pp. 16–20.

[12] J. F. Martins-Filho, E. Fontana, J. Guimaraes, D. F. Pizzato, and I. J. S. Coelho, "Fiber-optic-based corrosion sensor using OTDR," in *Proc. IEEE Sensors*, Oct. 2007, pp. 1172–1174.

[13] C. McCague *et al.*, "Novel sensor design using photonic crystal fibres for monitoring the onset of corrosion in reinforced concrete structures," *J. Lightw. Technol.*, vol. 32, no. 5, pp. 891–896, Mar. 2014.

[14] M. Bravo *et al.*, "Reinforced concrete structural corrosion monitoring using Hi-Bi photonic crystal fibres in a fiber loop structure," in *Proc. 23rd Int. Conf. Opt. Fibre Sensors*, vol. 9157, 2014, Art. no. 91579V.

[15] T. Jiang, L. Ren, Z. Jia, D. Li, and H. Li, "Application of FBG based sensor in pipeline safety monitoring," *Appl. Sci.*, vol. 7, no. 6, p. 540, May 2017.

[16] L. Ren, T. Jiang, D.-S. Li, P. Zhang, H.-N. Li, and G.-B. Song, "A method of pipeline corrosion detection based on hoop-strain monitoring technology," *Struct. Control Health Monitor.*, vol. 24, no. 6, Jun. 2017, Art. no. e1931.

[17] J. Gao, J. Wu, J. Li, and X. Zhao, "Monitoring of corrosion in reinforced concrete structure using Bragg grating sensing," *NDT&E Int.*, vol. 44, no. 2, pp. 202–205, Mar. 2011.

[18] W. Hu, H. Cai, M. Yang, X. Tong, C. Zhou, and W. Chen, "Fe–C-coated fibre Bragg grating sensor for steel corrosion monitoring," *Corrosion Sci.*, vol. 53, no. 5, pp. 1933–1938, May 2011.

[19] L. Ren, Z.-G. Jia, H.-N. Li, and G. Song, "Design and experimental study on FBG hoop-strain sensor in pipeline monitoring," *Opt. Fiber Technol.*, vol. 20, no. 1, pp. 15–23, Jan. 2014.

[20] F. Deng, Y. Huang, and F. Azarmi, "Corrosion detection for steel with soft coating using in-line fiber Bragg grating sensor," *Proc. SPIE*, vol. 10168, Apr. 2017, Art. no. 101681R.

[21] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*. New York, NY, USA: Academic, 2010.

[22] L. Tong, R.-W. Liu, V. C. Soon, and Y.-F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits Syst.*, vol. 38, no. 5, pp. 499–509, May 1991.

[23] P. S. K. Hansen, P. C. Hansen, S. D. Hansen, and J. A. Sorensen, "Experimental comparison of signal subspace based noise reduction methods," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Mar. 1999, pp. 101–104.

[24] J. Benesty, S. Makino, and J. Chen, *Speech Enhancement*. Berlin, Germany: Springer, 2005.

[25] J. Benesty *et al.*, *Noise Reduction in Speech Processing*, vol. 2. Berlin, Germany: Springer, 2009.

[26] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 27, no. 2, pp. 113–120, Apr. 1979.

[27] J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proc. IEEE*, vol. 67, no. 12, pp. 1586–1604, Dec. 1979.

[28] R. McAulay and M. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 28, no. 2, pp. 137–145, Apr. 1980.

[29] Y. Hu and P. C. Loizou, "A generalized subspace approach for enhancing speech corrupted by colored noise," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 4, pp. 334–341, Jul. 2003.

[30] Y. Huang, J. Benesty, and J. Chen, "Analysis and comparison of multichannel noise reduction methods in a common framework," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 16, no. 5, pp. 957–968, Jun. 2008.

[31] R. Sameni, C. Jutten, and M. B. Shamsollahi, "A deflation procedure for subspace decomposition," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2363–2374, Apr. 2010.

[32] A. Borowicz, "A signal subspace approach to spatio-temporal prediction for multichannel speech enhancement," *EURASIP J. Audio, Speech, Music Process.*, vol. 2015, no. 1, p. 5, Dec. 2015.

[33] P. von Bünau, F. C. Meinecke, F. C. Király, and K.-R. Müller, "Finding stationary subspaces in multivariate time series," *Phys. Rev. Lett.*, vol. 103, no. 21, Nov. 2009, Art. no. 214101.

[34] M. Baktashmotlagh, M. Harandi, A. Bigdeli, B. Lovell, and M. Salzmann, "Non-linear stationary subspace analysis with application to video classification," in *Proc. Int. Conf. Mach. Learn.*, Feb. 2013, pp. 450–458.

[35] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.

[36] Y.-X. Wang and H. Xu, "Noisy sparse subspace clustering," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 320–360, 2016.

[37] C. Zhang, H. Yan, S. Lee, and J. Shi, "Dynamic multivariate functional data modeling via sparse subspace learning," *Technometrics*, pp. 1–14, 2020.

[38] G. Nowak, T. Hastie, J. R. Pollack, and R. Tibshirani, "A fused lasso latent feature model for analyzing multi-sample aCGH data," *Biostatistics*, vol. 12, no. 4, pp. 776–791, Oct. 2011.

[39] S. Tao, D. Boley, and S. Zhang, "Local linear convergence of ISTA and FISTA on the LASSO problem," *SIAM J. Optim.*, vol. 26, no. 1, pp. 313–336, Jan. 2016.

[40] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Jan. 2009.

[41] C. Hsiao, L. Wang, and Q. Wang, "Estimation of nonlinear errors-in-variables models: An approximate solution," *Stat. Papers*, vol. 38, no. 1, pp. 1–25, Mar. 1997.

[42] D. C. Montgomery, *Introduction to Statistical Quality Control*, 8th ed. Hoboken, NJ, USA: Wiley, 2012.

[43] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 9, pp. 1622–1637, Sep. 2016.

[44] J. Chen, J. Benesty, and Y. Huang, "A minimum distortion noise reduction algorithm with multiple microphones," *IEEE Trans. Audio, Speech, Language Process.*, vol. 16, no. 3, pp. 481–493, Mar. 2008.

[45] J. S. Morris, "Functional regression," *Annu. Rev. Statist. Appl.*, vol. 2, pp. 321–359, Apr. 2015.

Authorized licensed use limited to: Tsinghua University. Downloaded on February 28,2021 at 05:52:56 UTC from IEEE Xplore. Restrictions apply.