Layout Optimization of Multi-Type Sensors and Human Inspection Tools With Probabilistic Detection of Localized Damages for Pipelines

AMIN ARIA, SHAPOUR AZARM, AND MOHAMMAD MODARRES
Department of Mechanical Engineering, Center for Risk and Reliability, University of Maryland, College Park, MD 20742, USA
Corresponding author: Amin Aria (aaria@umd.edu)
This work was supported by the Petroleum Institute, Khalifa University of Science and Technology, Abu Dhabi, UAE through the Pipeline System Integrity Management Project.

ABSTRACT In this paper, a new layout optimization approach for health monitoring of locally damaged and non-piggable segments of pipelines is presented. In contrast to the existing literature, the proposed optimization-based approach considers (i) sensors network and human inspection, as detection methods, together, (ii) severity level of damages, such as damage size and risk of failure, (iii) three probabilistic detection metrics (to detect, infer, and size different damages) and a health monitoring cost metric simultaneously, (iv) several key attributes of detection methods, such as data acquisition cost and frequency, in an optimization context, and (v) probabilistic sampling methods and probabilistic damage data. An optimization model for determining decision variables such as the type and location of sensors and human inspection areas and tools along a pipeline is formulated. Due to the unavailability of publicly available real-world data and benchmarks, applications of the proposed approach are demonstrated using two notional examples with synthetically generated damage data. In the first example, a step-by-step demonstration of the proposed approach is given. Considering the severity level of different localized damages, it is shown that the proposed layout optimization approach can obtain a better and more robust solution, especially when used during a pipeline design, compared to those that only consider a single detection method or rely on deterministic damage data. In the second example, a longer pipeline segment with a greater density of localized damages is considered to show applicability of the proposed approach to larger sized problems. Finally, based on the results obtained it is demonstrated that the proposed approach provides a practical solution for consideration of probabilistic detection metrics and probabilistic damage data corresponding to stochastic degradation processes in optimal health monitoring of pipelines.

INDEX TERMS Sensors placement, health monitoring, sensor layout optimization, probabilistic detection, human inspection, localized damages.

I. INTRODUCTION
Degradation processes that lead to localized damages, such as corrosion, while prominent, are uncertain and expensive to monitor in aging pipelines. This paper presents a new health monitoring approach with a focus on probabilistic detection of localized damages for (non-piggable) pipelines. Using this approach, an optimized layout of multi-type sensors and human inspection areas and tools can be determined for a pipeline. The proposed approach facilitates condition-based maintenance of pipelines where, in contrast to more expensive preventive maintenance schemes [1], the concept of the value of information (VoI), defined as the expected value of obtaining additional health monitoring information/data, is considered to reduce health monitoring costs while keeping high maintenance integrity [2]. A review of the literature on health monitoring for pipelines reveals that the proposed approach is different from and more general than those in the literature in several important ways:

The associate editor coordinating the review of this manuscript and approving it for publication was Ghufran Ahmed.
(i) Sensors network and human inspection are among the most common data gathering schemes (hereafter, also called “detection methods”), as reported in the literature for health monitoring of pipelines [3]. Each of these reported schemes has some strengths and weaknesses concerning pipeline applications [3], [4]. As an example, a network of sensors for a pipeline can provide frequent and inexpensive health monitoring measurement data. At the same time, it may include measurement biases and noises. On the other hand, human inspection-based data is more reliable but less frequent and more expensive to obtain. Nonetheless, these two schemes are often considered separately in optimization-based models [5]–[10]. Furthermore, for either of the two schemes (sensor network or human inspection), only a few of the key attributes such as data acquisition frequency, cost, location of sensors or inspection, type of sensors or inspection tool, and measurement error are considered in the reported works [4]–[10]. Meanwhile, considering the studies on the concept of VoI and affecting factors on the maintenance decision-making performance in single-scheme health monitoring systems [2], [11]–[14], an optimized multi-scheme health monitoring system can lead to more reliable maintenance decision-making from a cost vs. benefits point of view. As such, a combination of two schemes (sensors network and human inspection) is considered in the optimization-based approach of this paper to achieve a more cost-effective and reliable health monitoring of locally damaged pipelines. Moreover, using the proposed approach, a larger number of key attributes of both schemes, e.g., data acquisition frequency, data acquisition cost and measurement error, is considered in the formulated optimization problem considering the concept of VoI.

(ii) Most of the previously reported approaches on pipeline health monitoring only consider a binary detection metric [4], [15]. As a result, health monitoring layouts produced by these approaches [16], [17] cannot find the severity level of damages with different sizes and types and different risk of missing potential failures. On the other hand, several probabilistic detection metrics have been reported in the literature, which can be used for reliable and optimal health monitoring of pipelines. Probability of Detection (POD) and probabilistic Measurement Error (ME) are two of these metrics [7], [18]–[20]. Also, Inference Probability (IP) is another detection metric, defined as the probability of true prediction of the pipeline’s state at a particular point given the spatial correlation of localized damages and corresponding health monitoring evidence at other locations. While some studies have considered the spatial variability of localized damages along a pipeline [7], [21], to the best of our knowledge, IP has not been considered in the corresponding literature. In this paper, all three metrics of POD, IP, and ME, along with their dependencies on damage type, damage size, and damage-to-sensor distance, are considered. Note that human inspection is also modeled here as a low-frequency and high-coverage sensor. In this way, using the proposed approach, the final multi-scheme layout will be consistent with the severity and risk of failure corresponding to different localized damages. Moreover, the cost-effectiveness of the proposed approach can be improved by inference of the pipeline state at all its points considering limited health monitoring evidence and analysis of partial coverage inspection data [22].

(iii) Only a handful of previous works on optimal sensor placement have considered cost-effectiveness while maximizing probabilistic detection of damages [10], [23]. Additionally, simplifying assumptions, such as having a fixed number of the same-type sensors and a one-dimensional sensor network, are used in the corresponding literature to minimize the computational cost of multi-type sensor optimization models [5], [6], [9], [17], [24], [25]. Moreover, the literature on human inspection is more focused on optimizing inspection time rather than deciding on inspection location and tool type [4], [7], [23]. As such, an optimization-based solution for placement and type of sensors and human inspection on the exterior surface of pipelines is provided in the approach of this paper, where health monitoring cost and average detection probability are optimized concurrently with only a few simplifying assumptions considered. As a result, the proposed approach can be used for cost-effective and reliable detection of localized damages with different sizes and types located on inner or outer surface of a pipeline.

(iv) Finally, most of the existing optimal sensor placement approaches are based on deterministic information about location, type, and size of damages [25]. Unfortunately, the stochastic nature of localized damages [7] can undermine the performance of such approaches [25]. The proposed approach considers the stochasticity of the damage configuration in terms of location, type, and size through damage simulation and using probabilistic damage data. Consequently, the proposed approach is well-suited for pipelines subject to stochastic degradation even if it is used in the early stages of pipeline design.

The remainder of this paper is organized as follows. In Section 2, an overall description of the problem statement is given. Three probabilistic detection metrics, their aggregation construct, and proposed mathematical modeling of detection methods are detailed in Section 3. The steps of the proposed approach are discussed in Section 4 while the formulation of optimization problem is presented in Section 5. Next, two examples and corresponding results are discussed in Section 6. Finally, the paper ends with some concluding remarks in Section 7.

II. PROBLEM STATEMENT

Consider a pipeline that is subjected to localized damages. The problem is to determine an optimized health monitoring layout for different segments of this pipeline. As shown in Figure 1, the layout determines the location and type of all sensors and human inspection areas and tools for obtaining data regarding the location and size of localized damages over the pipeline. Here, it is assumed that sensors are mounted, and inspection tools are used on the exterior surface of the pipeline even when the localized damages are on the inner
The objective function of the layout optimization problem is formulated to be a combination of an average probabilistic detection metric and a health monitoring cost metric.

The key elements of the framework for this problem are illustrated in Figure 2. As shown in Figure 2, it is assumed that there is an existing damage model and pipeline vulnerability model available. The damage model refers to a damage size distribution for each damage type based on historical data. The damage type refers to an underlying failure mechanism that results in localized damage. It is also assumed that an existing pipeline vulnerability model is known based on historical data from pipeline health monitoring and/or inspection. The vulnerability model consists of distributions that reveal a pointwise probability of having a damage of specific type and size at an arbitrary location on the inner or outer surface of the pipeline.

III. METRICS FOR PROBABILISTIC DETECTION AND MODELS FOR DETECTION METHODS

Three probabilistic detection metrics, i.e., POD, IP, and ME are formulated and employed in the proposed approach. These metrics and their application with respect to different detection methods (i.e., sensors and human inspection) are described in Sections 3.1 to 3.3.

Each detection method has a specific coverage range, cost, precision (measurement error), data acquisition frequency, and inference capability. As an example, human inspection is modeled as a high-coverage and low-frequency data generating detection method (sensor). In this modeling, it is assumed that the inspection area is independent of inspection tool type and the inspection tool is taken to the exact location of all localized damages in the inspection area. As such, no damage-to-sensor detection dependency is considered for human inspection.

In the proposed approach, a graph is defined for each particular realization of the simulated pipeline with localized damages. Corresponding graph is formed by placing a node, as candidate location for detecting localized damage, near each damage of the pipeline realization of interest (see Figure 4). Upon finding a solution to the optimization problem, it is determined which detection method (if any) should be used at each of the graph nodes to maximize an average detection probability. (Note that in the case of the assignment of human inspection to a node, the inspection area will be centered on the corresponding node.)

Details about the models for detection methods and their relationship with the probabilistic detection of damages are provided in the two following sections.

A. PROBABILITY OF DETECTION (POD)

For damage of a particular type and size to be detected by a detection device (e.g., acoustic emission (AE) sensor), POD depends on factors such as size and type of damage and the
distance between the damage and the corresponding detection device. In the literature, these dependencies are typically not investigated altogether [7], [26]. In this paper, a new tabularized, and aggregate model of POD as a function of damage size and damage-to-sensor distance is developed (see Section 4, Step 6).

The use of this new aggregate POD model is facilitated through a binary (hit or miss) detection model, namely $P_{\text{exist}}$. A damage is considered detectable in the proposed $P_{\text{exist}}$ model, regardless of its type and size, if it is located in the coverage range (footprint) of a detection device. The coverage range of a detection device is defined as the distance beyond which the marginal POD of a damage by that particular detection device is zero. For the case of sensors, coverage range of the detection method is equal to that of the sensor. However, for human inspection coverage range is equal to half of the inspection length along the pipeline since inspection tool coverage range is not of importance as long as it is assumed that the tool is taken to the exact location of all damages in the inspection area.

As shown in Figure 3 for an unrolled pipeline segment, the longitudinal and (unrolled) circumferential coordinates are used in the $P_{\text{exist}}$ model (and throughout this paper) as coordinates of the location of damages on the surface of the pipeline. To account for the circularity of the circumferential coordinate (y-axis in Figure 3), a replica of each damage is considered with the same longitudinal component as that of the original damage. The replica’s circumferential component, as defined, is $(y-2\pi R)$ where $y$ denotes the circumferential component of the original damage location and $R$ is the radius of the pipeline. The origin of the x-y coordinate system is located at the leftmost point of the bottom line of the pipeline (Figure 3). The location vector of both the damage and its replica are inserted into the $P_{\text{exist}}$ model (Eq. 1) to determine whether damage ($d_i$) is in the coverage range ($cr_j$) of the detection method at a node $j$ ($dm_j$). To do this, consider a damage $k$ denoted by $d_k$ (located at: $a_k = (x_k, y_k)$) and damage $i$, $d_i$ (located at: $a_i = (x_i, y_i)$), as shown in Figure 3. The replica of $d_i$ (located at: $a'_i = (x_i, y_i-2\pi R)$) and $d_k$ itself are in the coverage range of the detection method at node $j$ (located at: $b_j = (x_j, y_j)$). Hence, both damages are detectable by the detection method at node $j$, $dm_j$.

B. INFERENCE PROBABILITY (IP)

IP is a metric used for detectability of damage of a specific size and type at a particular point of the pipeline, given the health monitoring data at some other point of the pipeline. To better illustrate IP, the differences between IP and POD in the proposed approach are illustrated in Figure 4. As shown, a pipeline segment with two damage types (shown with hollow diamond and hollow multiplication symbols with different sizes) and detection methods (solid triangle and solid plus symbols) are considered. A hollow square node is located near each of the damages and detection methods are assigned to these nodes. For example, a detection method (e.g., a sensor) of the type triangle is assigned to node 3 while a detection method type zero is assigned to node 4. (The assignment of no detection method to a node is considered as the detection method type zero.)

In Figure 4, the coverage boundaries of detection methods are shown with dashed circles, while inference boundaries are shown with dashed parabolas. Note that the other (left side) inference boundary of the detection method at node 1 ($dm_1$) and at node 3 ($dm_3$) is not shown in Figure 4 to avoid clutter.

For each node $j$, inference boundaries are the locus of zero IP, and an inference distance limit is defined as the minimum distance of zero IP locus from the corresponding node. IP is a function of the damage-to-node distance, detection method type, damage size and type, and similarity of the degradation behavior at the damage and node locations. Hence, $IP(d_i, dm_j)$ is zero at the inference boundaries of node $j$ while $IP(d_i, dm_j)$ is unity if $d_i$ is located at node $j$. 

FIGURE 3. Geometry of damages on the surface of a pipeline segment.
Sites of damage will be missed if POD and IP values corresponding to all neighboring detection methods are zero at the damage location. To minimize the probability of missing a true damage, a particular detection method should be chosen in a way that reduces the Probability of Not Detection (POND), accordingly:

\[ \Omega_i = (1 - \Theta_i) \times (1 - \Phi_i) = \prod_{j=1}^{N} (\theta_{i,j}) \times (1 - \Phi_i) \]  

(2)

where \( \Omega_i \) is POND of damage \( i \), \( \Theta_i \) is the aggregate IP of damage \( i \) (i.e., IP of all detection methods assigned to the nodes near damage \( i \)), and \( \Phi_i \) is the aggregate POD of damage \( i \) (i.e., resultant POD of all detection methods with damage \( i \) in their coverage range). \( N \) is the number of nodes. \( \theta_{i,j} \), an element of \( \Theta \) matrix, represents the probability of missing \( d_i \) through inference based on health monitoring data gathered by \( d_m \). In Eq. 2 (and elsewhere in the proposed approach), it is assumed that the probability of a false positive detection is zero.

To simplify the model, a log-linear version of POND, or LPOND, is used (Eq. 3) where \( \Omega'_i \) is LPOND of damage \( i \).

\[ \Omega'_i = \sum_{j=1}^{N} \log(\theta_{i,j}) + \log(1 - \Phi_i) \]  

(3)

Application of a tabularized LPOND based on Eq. 3 is discussed in Section 4 (Step 6).

C. MEASUREMENT ERROR (ME)

ME is an uncertainty metric defined as the difference between reported and true values of a measurement quantity. In this paper, the probabilistic ME, simply referred to as ME, is defined as the probability of the reported size value of a damage of known type not being within the interval \([(True size - \epsilon_l), (True size + \epsilon_u)]\). The quantities \( \epsilon_l \) and \( \epsilon_u \) are respectively for pre-specified acceptable lower and upper error margins, specified by an expert, and “True size” refers to the actual size of the damage. Note that only tool measurement errors, but not the human error, are considered in the calculation of ME values for human inspection. Nonetheless, human error can be easily included, as reported elsewhere [27].

The LPOND model, as an aggregate of POD and IP models, and probabilistic ME are considered in a step-by-step approach for pipeline health monitoring, as discussed next.

IV. PROPOSED APPROACH

The proposed approach simulates and randomly places localized damages on the pipeline surface and repeats the process for many pipeline realizations, each with the consideration of degradation and maintenance history of similar pipelines. The proposed approach is general, and it is envisioned to be used for any pipeline regardless of detection methods used. However, minor modeling modifications may be needed considering pipeline and detection methods specifications. Moreover, the proposed approach can be used during early stages of a pipeline design and, also, for online health management of existing pipelines since data pre-processing (Steps 1-9 in Figure 5), and corresponding optimization problems can be solved in a timely manner using a computerized scheme (see Table 7).

Using the approach, a pipeline is divided into multiple segments, each having a uniform longitudinal and non-uniform circumferential density of damages. For each realization, the location and size of the localized damages are simulated and placed on each pipeline segment (see [20], [28]). However, for non-uniform longitudinal damage density, one can use a method like the one discussed in [29]. Moreover, clustering is conducted for each realization of spatially distributed damages to facilitate the optimization work and improve the choice of detection methods and their placement over pipeline surface. Additionally, to account for the stochasticity of localized damages when there is a limited budget on computational time, the proposed approach uses a modified version of a well-known probabilistic sampling method, i.e., the Wilks method [30].

Figure 5 shows an overview of the various steps of the proposed approach, as shown with a step number to the right side of each block. These steps are detailed next.
**Step 1 - Obtain Modified Vulnerability Distributions (MVDs):** Collect prior maintenance and degradation data of a similar pipeline. Map the available data over the pipeline surface and perform data analysis, including data simulation (augmentation) for model selection and fitting. Subsequently, obtain three distributions for each localized damage type: longitudinal, circumferential spatial, and size distributions. Assume the three distributions are independent. Modify the distributions to have a higher intensity for larger size damages in the pipeline areas with a higher risk of failure.

**Step 2 - Determine Segments and Mesh Cell Size:** Based on longitudinal MVD for each damage type, divide the pipeline into segments with an (approximately) uniform longitudinal density of damages. Mesh each pipeline segment into circumferential strips (as mesh cells) with their width laid out along the longitudinal direction (x-axis) of the pipeline (recall Figure 3). Considering constant longitudinal density of damages and using Poisson distribution, determine the width of each strip so that the probability of having more than a damage of each type in each strip is negligible.

**Step 3 – Generate a Pipeline Segment Realization:** Assuming a constant longitudinal density of damages along a pipeline segment, use the Poisson distribution to calculate the probability of having a single damage of an arbitrary type in each arbitrary strip (mesh cell). Use this value in a Bernoulli test as a success probability. If the simulated outcome of a Bernoulli test is one, place a damage on the centerline of the corresponding strip and determine its circumferential location based on the circumferential MVD.

**Step 4 - Define Type-Size Indicators:** Once a segment realization is simulated, determine a type-size class vector for each of the damages of that realization. For each damage type, define type-size classes which refer to the intervals of the corresponding size MVD with the same cumulative probability. Use a Monte Carlo simulation to assign one of these classes to each of the simulated damages of the realization at hand. As an example, consider Figure 6 that presents a size distribution of pitting corrosion damages.
MVD for pitting corrosion damage [31]. For this size MVD, the number of classes is set to four. If a (randomly generated) class of damage \( i \) (\( d_i \)) is equal to 3, \( d_i \) will be a pit damage with a depth in the interval 0.33-0.55 mm, and its binary class vector will be \( C_i = [0,0,1,0] \).

In the case of multiple damage types (e.g., pitting corrosion and stress corrosion cracking damages), extend the above-explained technique so that classes 1 to 4, for example, represent pitting corrosion size MVD while classes 5 to 8 represent stress corrosion cracking.

**Step 5 - Determine Nodes:** Place a single node near each damage of a segment realization. The longitudinal location of the node is the same as that of the corresponding damage. In contrast, its circumferential location is randomly generated to make it possible for the optimization solver to distinguish detection methods with different coverage and detection capabilities.

**Step 6 - Develop IP, POD, and ME Matrices:** Follow the instructions below and develop matrices to be used for modeling probabilistic detection in an optimization context.

To calculate the IP term (\( \sum \log(\delta_{ij}) \)) of LPOND of each damage (Eq.3) in an optimization context (see Eq. 6.2), develop a \( \log(6^{k-m}) \) matrix for each combination of damage type-size classes \( k \) and detection methods type \( m \). On the other hand, to calculate the POD term of Eq. 3, develop \( \Lambda_i \) as the POD matrix of damage \( i \) with element \( \lambda_i^{m,j} \). Use Eq. 4 to calculate \( \lambda_i^{m,j} \) that denotes POD for damage \( i \) of \( a \ priori \) known class \( k \), which will be recognized by detection method type \( m \) (if used) at node \( j \). Assume uncorrelated POD dependency on damage size \( (\lambda_i^{m}) \) and damage-to-node distance \( (\lambda d_i^{m,j}) \).

\[
\lambda_i^{m,j} = \lambda_i^{m} \times \lambda d_i^{m,j}
\]

Lastly, calculate tabularized ME values to be used in the optimization problem (Eq. 6.4). At first develop a matrix \( \Delta \) with its element \( \delta_{k,m} \) defined to be the ME value (Section 3.3) corresponding to the detection of a damage of class \( k \) by detection method type \( m \). Next, considering particular placement of damages in the realization at hand, use Eq. 5 to calculate the mean ME (\( \delta_{j}^{m} \)) corresponding to the detection of all damages \( i \) in the coverage range of detection method \( m \), if it is used at node \( j \). The number of corresponding damages of interest is denoted by \( N(i,j) \) and \( c_{ij} \) is the element of type-size class vector for damage \( i \) (Step 4).

\[
\delta_{j}^{m} = \frac{\sum_{i} \delta_{k,m} \times c_{ik}}{N(i,j)}
\]

**Step 7 - Cluster Damages:** Cluster damages of the realization at hand to smooth the optimization work. To do that, define a Minimal Spanning Tree (MST) [32] with damages at its nodes. Determine the clustering distance limit to be equal to minimum of the inference distance limit (Section 3.2) and the length of the longest edges of the corresponding MST. Then, determine the number of clusters accordingly. Next, define a similarity metric based on the model in Eq. 1 and Eq. 6.10. Use the constrained k-means [33] algorithm to cluster damages where the defined similarity metric is considered, and damages at a distance greater than clustering distance limit are not assigned to the same cluster.

**Steps 8 and 9 - Perform Partial Coverage Inspection Analysis and Determine Detection Limits and Inference Confidence:** Use partial coverage inspection analysis [22] to estimate the state of the entire pipeline segment being studied with minimal monitoring coverage (i.e., minimum health monitoring cost) and a certain confidence level. In the case of locally damaged pipelines, use partial coverage inspection to determine the damage detection percentage required for having the worst-case damage of the entire segment under investigation being detected with \( p \% \) confidence. Define an overall detection constraint as well as a lower limit on detection percentage for each damage cluster considering the required detection percentage.

**Step 10 – Perform Health Monitoring Layout Optimization:** Feed the results of the pre-processing stage (i.e., Steps 1 through 9) as well as available detection methods and their specifications into the optimization model. (This step and the corresponding optimization formulation are detailed in Section 5.)

**Steps 11 – Implement Probabilistic Sampling:** Use a modified version of the Wilks method [34] to determine the minimum number of realizations (\( N_W \)) of a stochastic phenomenon (e.g., localized corrosion) needed to capture at least \( p_1 \% \) of variations (e.g., spatial and size variations of localized damages) with \( p_2 \% \) confidence. Repeat Steps 1 to 10 of the proposed approach \( N_W \) times and obtain \( N_W \) optimal health monitoring layouts.

**Steps 12 – Aggregate All Layouts:** Aggregate all \( N_W \) layouts to form an aggregated layout where \( p_1 \% \) of variations are considered with \( p_2 \% \) confidence. Use K-means clustering to find clusters of detection methods in the aggregated layout. Determine the number of clusters considering the value of \( N_W \) and the total number of detection methods in the aggregated layout. Then, place the detection methods at the center of the corresponding clusters to form the final layout.

**V. HEALTH MONITORING LAYOUT OPTIMIZATION PROBLEM FORMULATION**

The health monitoring layout optimization problem is formulated as a mixed-integer nonlinear programming problem where the objective function (Eq. 6.1) is formed by a weighted sum of an average utility and an average LPOND function. The utility function is formulated based on key features of available detection methods, including data acquisition cost, coverage range, and an information metric. The utility of the information metric is defined as a linear combination of the utility of data acquisition frequency, ME, and comprehensiveness of the gathered data. The utility is maximized in order to maximize the information comprehensiveness and data acquisition frequency while the cost of the health monitoring layout and ME are minimized. Moreover, the average LPOND is maximized to, equivalently, maximize
the geometric mean of POND of all damages in the realization of the pipeline segment under study.

The variables being optimized in the optimization problem are \( dm_j \) (Eq. 6.19), i.e., an integer indicator for the type of detection method used at each node \( j \). All other variables (i.e., binary detection indicators and detection method variables in Eqs. 6.10-12) are either determined in the preprocessing stage of the approach or are a function of \( dm_j \).

The formulation of the optimization problem is shown by Eqs. 6.1-6.21, as shown at the bottom of this page, with Table 1 providing a definition for all symbols in the problem. Note that the bold letters in Table 1 and formulated

\[
\text{Minimize } \sum_{j=1}^{N} \left( w_{\text{obj}} \frac{E(\Omega') - G_L}{G_U - G_L} + (1 - w_{\text{obj}}) \frac{-U_{\text{tot}}}{b_u \times N} \right)
\]

where:

\[
\Omega' = \sum_{m=0}^{M} \sum_{j=1}^{N} \sum_{k=1}^{Q} \{ \rho_j \times c_{i,k} \times id_{j,m} \times \log(\theta_{ij}^{km}) \} + \sum_{j=1}^{N} \log \left( 1 - \eta_{ij} \times \mathbf{A}^t_i \times \mathbf{D}_j \right)
\]

\[
\sum_{i=1}^{N} \frac{\Omega'}{N} = E(\Omega')
\]

\[
U(\text{id}_{mj}) = \sum_{m=0}^{M} [ \text{id}_{ij,m} \times (w_{fr} \times U(fr_{m}) + w_{iv} \times U(iv_{m}) + w_{ME} \times U(\delta_{m}^{ij}))]
\]

\[
U_j = w_{co} \times U(\mathbf{D}_j \cdot \mathbf{CO}^t) + w_{cr} \times U(\mathbf{D}_j \cdot \mathbf{CR}^t) + U(\text{id}_{mj})
\]

\[
U_{\text{tot}} = \sum_{j=1}^{N} (U_j \times (1 - \text{id}_{j,m=0}))
\]

\[
0 \leq w_{\text{obj}} \leq 1
\]

\[
w_{\text{cost}} + w_{\text{cov}} + w_{\text{Fr}} + w_{\text{IV}} + w_{\text{ME}} = 1
\]

\[
w_{\text{cost}}, w_{\text{cov}}, w_{\text{Fr}}, w_{\text{IV}}, w_{\text{ME}} \geq 0
\]

Subject to:

\[
\text{if } \left[ \left( x_i - x_j \right)^2 + \min \left( |y_i - y_j|, |y'_i - 2\pi R - y'_j| \right)^2 \leq \mathbf{ID}_j \cdot \mathbf{CR} \right], \text{ then } \eta_{ij} = 1; \text{ otherwise } \eta_{ij} = 0 \quad (6.10)
\]

\[
\text{if } \sum_{j=1}^{N} \eta_{ij} \geq 1, \text{ then } \pi_i = 1; \text{ otherwise } \pi_i = 0 \quad (6.11)
\]

\[
\text{if } dm_j = m, \text{ then } id_{j,m} = 1; \text{ otherwise } id_{j,m} = 0 \quad (6.12)
\]

\[
\sum_{m=0}^{M} \text{id}_{ij,m} = 1
\]

\[
b_i \times N \leq \sum_{i=1}^{N} \tau_i \leq b_u \times N \quad (6.14)
\]

\[
\sum_{i} \tau_i \geq d_{ic}, i \in \text{cluster#ic} \quad (6.15)
\]

\[
E(\text{cost}_j) \leq c_{oa} \quad (6.16)
\]

\[
\sum_{j=1}^{N} \eta_{ij} \leq \text{Re}_{\text{max}} \quad (6.17)
\]

\[
E \left( \sum_{j=1}^{N} \eta_{ij} \right) \leq \text{Re}_U \quad (6.18)
\]

\[
0 \leq dm_j \leq M \quad (6.19)
\]

\[
1 \leq i_c \leq N_{\text{cl}} \quad (6.20)
\]

\[
G_L \leq \Omega'_i \leq G_U \quad (6.21)
\]
### TABLE 1. Symbols used in the optimization model (Eqs. 6.1 – 6.21).

| Symbol | Definition |
|--------|------------|
| $b_1$  | Lower bound of overall detection percentage |
| $b_2$  | Upper bound of overall detection percentage |
| $c_{i,k}$ | Damage type-size vector with element $c_{i,k}$ |
| $c_{m}$ | Upper bound of expected cost of detection |
| $C_m$  | Detection methods cost vector (its $m+1^{th}$ element is cost of detection method $m$) |
| $CR$   | Detection methods coverage vector (its $m+1^{th}$ element is coverage radius of detection method $m$) |
| $d_{i,nc}$ | Minimum detection for damages in cluster $i$ |
| $d_{m}$ | Integer indicator of type of the detection method assigned to node $j$ |
| $E(.)$ | Expected value of ($.) |
| $f_{im}$ | Data acquisition frequency of detection method type $m$ |
| $G_L$  | Lower bound of LPOND |
| $G_U$  | Upper bound of LPOND |
| $I$    | Damage number superscript |
| $i_{m}$ | Information metric of detection method used at node $j$ |
| $i_{v_{m}}$ | Information comprehensiveness/value for detection method type $m$ |
| $ID_{ij}$ | Detection method type vector corresponding to node $j$ |
| $J$    | Node number superscript |
| $m$    | Type indicator for detection method used at each node |
| $M$    | Number of available detection methods |
| $N_d$  | Number of damages/nodes in the segment realization at hand |
| $N_{cl}$ | Number of clusters of damages in the segment realization at hand |
| $Q$    | Number of damage type-size classes |
| $RE_{max}$ | Upper limit of detection redundancy for each damage |
| $RE_{u}$ | Upper limit of expected detection redundancy |
| $t$    | Vector superscript for transpose of a vector |
| $U(.)$ | Input utility for ($.) |
| $U_{ij}$ | Utility value of the detection method used at node $j$ |
| $U_{tot}$ | Total utility value of a layout |
| $(x,y)$ | Location coordinates of a damage or node (see section 3.1) |
| $y'$  | Circumferential coordinate of the replica of a damage (see section 3.1) |
| $w_{ij}$ | Weighting factor of ($.$) |
| $\delta_m$ | Mean ME corresponding to detection method $m$ used at node $j$ (see Eq. 5) |
| $\eta_{ij}$ | $P_{x_{ij}}$ binary indicator (Eqs. 1 and 6.10) |
| $\theta_{k,m}^{i,j}$ | Probability of missing damage $i$ of type $k$ through inference by detection method $m$ at node $j$ (Eq. 6.2) |
| $A_{ij}'$ | Column $j$ of POD matrix of damage $i$ (Eq. 6.2) |
| $\pi'_{i}$ | Detection indicator of damage $i$ (Eq. 6.11) |
| $\Omega_{i}'$ | LPOND of damage $i$ (Eq. 6.2) |

In the problem formulation, Eq. 6.2 represents LPOND of damage $i$ (see Section 3.2). Eq. 6.4 defines the information metric corresponding to data of all the damages detected by the detection method used at node $j$. Moreover, Eq. 6.5 defines the input utility value corresponding to the detection method used at each node, while the detection method type vector (Eqs. 6.12-13) is used to show the type of detection method used at each node. On the other hand, Eq. 6.10-6.21 represent...
VI. EXAMPLES

Two notional examples are solved in this section using the proposed approach. For both examples, it is assumed that the pipeline segment under consideration is used for transporting crude oil. Except for the assumed pipeline radius ($R = 1$ unit of length, i.e., 1 meter), other pipeline specifications, and their effect on the performance of detection methods, are not considered. Moreover, in both examples, only localized damages resulting from internal pitting corrosion, but not any other failure mechanism, are considered. Furthermore, pit depth is considered for damage size while pit length is neglected following the analyses presented in [37]. Due to lack of access to real-world data and unavailability of benchmark datasets for pipeline health monitoring, synthetic internal pitting corrosion data are generated considering values, models, and relations reported in the literature for pitting corrosion or similar degradation processes [26], [38], [39]. Generated data are then used as the input in both examples. Stochasticity of the localized damages is also considered through probabilistic sampling.

In the first example, Section 6.1, a problem corresponding to a short pipeline segment, is solved to illustrate the proposed approach step-by-step. The corresponding results are used to evaluate the performance of the proposed approach. The final health monitoring layout is obtained for this example through the aggregation of layouts corresponding to different realizations of pitting corrosion damages over the pipeline segment. For this example, in the layout optimization formulation, there are 142 continuous variables representing utility of the detection method at each node and 198 binary variables representing indicators for probabilistic detection, inference, and the types of detection method at each node. Moreover, there are 345 constraints following the optimization formulation discussed earlier. For the second example, a longer pipeline segment is considered in Section 6.2 (Example 2), where only the results corresponding to a single realization of damages are presented. This example is presented to show the applicability of the proposed approach to a larger sized problem with greater density of localized damages, for which there are 490 continuous variables, 670 binary variables, and 11,065 constraints in the layout optimization problem. While both examples of this section consider a single pipeline segment, the proposed approach can be easily extended for use for an entire pipeline to obtain an optimal health monitoring layout separately for each pipeline segment of interest.

A. EXAMPLE 1

In this example, an oil pipeline segment with a 50-meter length ($L$) and a 1-meter radius ($R$) is considered. It is assumed that there are only two detection options: AE sensors (the high-frequency and low-coverage detection method type 1 ($m_1$)) and human inspection with an ultrasonic tool (the low-frequency, high-coverage, and expensive detection method type 2 ($m_2$)). An assignment of no detection method to the nodes is the other option ($m_0$). Table 2 shows assumed specifications of the detection methods. The specifications values are chosen considering the literature recommendation [20], [21], [26] and assumed pipeline segment geometry.

### TABLE 2. Specifications of detection methods.

|                  | AE sensor | Human Inspection with ultrasonic tool |
|------------------|-----------|---------------------------------------|
| Coverage radius (meter) | 0.400     | 20.0                                  |
| Cost              | 1.000     | 10.0                                  |
| Utility (cost)    | 1.000     | 0.10                                  |
| Utility (coverage)| 0.013     | 1.00                                  |
| Utility (frequency)| 1.000   | 0.01                                  |
| Utility (information value) | 0.020 | 1.00                                  |

Constraints of the formulated problem, including (1) detection limits (Eq. 6.14-15): these constraints are determined in Steps 8 and 9 of the proposed approach using partial coverage inspection analysis. (2) Overall inference limit (Eq. 6.21): this constraint is also determined in Steps 8 and 9 of the proposed approach. (3) Cost limit (Eq. 6.16): the expected cost of detection methods along a segment cannot be higher than a pre-specified limit.

The above-formulated optimization problem is modified, using linear and integer programming techniques such as big M method [35], so that the corresponding problem can be solved using a mixed integer non-linear programming [36] solver.

In the first example, Section 6.1, a problem corresponding to a short pipeline segment, is solved to illustrate the proposed approach step-by-step. The corresponding results are used to evaluate the performance of the proposed approach. The final health monitoring layout is obtained for this example through the aggregation of layouts corresponding to different realizations of pitting corrosion damages over the pipeline segment. For this example, in the layout optimization formulation, there are 142 continuous variables representing utility of the detection method at each node and 198 binary variables representing indicators for probabilistic detection, inference, and the types of detection method at each node. Moreover, there are 345 constraints following the optimization formulation discussed earlier. For the second example, a longer pipeline segment is considered in Section 6.2 (Example 2), where only the results corresponding to a single realization of damages are presented. This example is presented to show the applicability of the proposed approach to a larger sized problem with greater density of localized damages, for which there are 490 continuous variables, 670 binary variables, and 11,065 constraints in the layout optimization problem. While both examples of this section consider a single pipeline segment, the proposed approach can be easily extended for use for an entire pipeline to obtain an optimal health monitoring layout separately for each pipeline segment of interest.

A. EXAMPLE 1

In this example, an oil pipeline segment with a 50-meter length ($L$) and a 1-meter radius ($R$) is considered. It is assumed that there are only two detection options: AE sensors (the high-frequency and low-coverage detection method type 1 ($m_1$)) and human inspection with an ultrasonic tool (the low-frequency, high-coverage, and expensive detection method type 2 ($m_2$)). An assignment of no detection method to the nodes is the other option ($m_0$). Table 2 shows assumed specifications of the detection methods. The specifications values are chosen considering the literature recommendation [20], [21], [26] and assumed pipeline segment geometry.

1) STEP-BY-STEP APPLICATION OF THE PROPOSED APPROACH

Considering the modified Wilkes method by Pourgol-Mohammad [34], forty-six realizations ($N_w$ in Step 11 of Figure 5) of the corresponding pipeline segment are generated to guarantee capturing 95% of spatial and size variations of localized damages with 90% two-sided confidence. There are 5 to 15 damages in each of these realizations. The 22nd realization has a relatively dense placement of damages (with 13 damages) that is good for illustration purposes. Hence, a brief discussion of the steps in the proposed
approach (as shown in Figure 5) for the 22nd realization of pipeline damages is followed.

**Step 1** – Based on the existing literature [29], [40], [41], it is assumed that the longitudinal MVD for internal pitting corrosion is 0.2 pits per one-meter length of the pipeline. Also, considering the higher likelihood of corrosion at the lower quadrant of oil pipelines [28], it is assumed that the circumferential MVD of localized damages follows a normal distribution with zero mean at the bottom of the pipeline. Moreover, assuming that the risk of failure is consistent with damage size over the pipeline segment of interest, the distribution shown in Figure 6 is considered as the size MVD.

**Step 2** – Using Poisson distribution with identically, independently, and uniformly distributed damages along the pipeline segment, the mesh strip width is found to be 0.5 meters, so that the probability of having more than one localized damage at each mesh strips is less than 0.01.

**Steps 3, 4, 5** – The problem setup corresponding to the 22nd realization is provided in Figure 7, where axes are defined based on Eq.1 and Figure 3. Moreover, diamonds of different sizes represent localized damages of classes 1 through 4 (see Figure 6), where classes are randomly assigned to each damage. Moreover, thirteen nodes are located along the pipeline segment for the 22nd realization (stars represent nodes in Figure 7).

The longitudinal locations of the nodes are set to coincide with those of corresponding pits. In contrast, the circumferential location of the nodes is determined by adding a random value from the interval (-0.5, 0.5) to that of the corresponding damage. This interval is chosen considering AE sensor coverage radius (See Table 2).

**Step 6** – In this step, to develop the aggregate POD model (see Eq. 3) for AE sensors, a decaying power relation is defined for the dependency of the POD on damage-node distance based on the distance metric in Eq. 1, the assumed coverage radius of AE sensors (0.4 m), and the model provided in [26]. However, this dependency for human inspection with the ultrasonic tool is neglected since the inspector takes the inspection tool to the exact location of damages in the inspection area. On the other hand, dependency of POD on damage size is calculated for AE sensors and human inspection with an ultrasonic tool using probabilistic relations provided in [7], [20]. As such, POD as a function of type-size class of damages in 22nd realization is presented in Table 3.

| Class (k) | Acoustic Emission (m=1) | Human Inspection (m=2) |
|----------|-------------------------|------------------------|
| 1        | 0.32                    | 0.02                   |
| 2        | 0.54                    | 0.07                   |
| 3        | 0.72                    | 0.25                   |
| 4        | 0.90                    | 0.77                   |

**Step 7** – Following the proposed clustering approach (Section 4, Step 7), damages of the 22nd realization are
clustered to 3 groups where the clustering distance limit is determined to be 15 meters. In addition, the minimum number of detected damages of each cluster (Table 5) is determined considering the overall detection lower limit (Eq. 6.14).

Steps 8, 9- In parallel, it is assumed that the overall detection lower limit is 50% for an estimation of the entire segment with 90% confidence. Considering the overall detection lower limit, minimum detection of each of clusters is determined in step 7.

Steps 10- Subjective utility weights (Table 6) are used to form the utility function that is defined to be maximized (see Table 1 for abbreviations). Moreover, the cost limit is set to be 25 (see Table 2 for cost metric values). Also, LPOND limits (Eq. 6.21) are set to be $-1.5$ and $-12$.

2) SINGLE REALIZATION RESULTS AND DISCUSSION

Two health monitoring layouts and results corresponding to the 22nd realization of the pipeline segment of this example are discussed in this section. All the results presented in this section, including layouts illustrated in Figure 8,

| Cluster # | Damage # | Minimum Detected Damages |
|-----------|----------|-------------------------|
| 1         | 1,2,3,4  | 2                       |
| 2         | 5,6,7,8,9| 3                       |
| 3         | 10,11,12,13| 2                     |

|  | $W_{co}$ | $W_{cr}$ | $W_{fr}$ | $W_{iv}$ | $W_{ME}$ |
|---|----------|----------|----------|----------|----------|
|   | 0.3      | 0.3      | 0.1      | 0.15     | 0.15     |

TABLE 4. Class intervals and corresponding ME probabilities ($\delta_k$, m).

| Class (k) | AE sensors (m=1) | Human Inspection (m=2) | Class Interval (mm) |
|-----------|------------------|------------------------|---------------------|
| 1         | 0.130            | 0.147                  | [0 0.19]            |
| 2         | 0.060            | 0.151                  | [0.19 0.32]        |
| 3         | 0.045            | 0.142                  | [0.32 0.55]        |
| 4         | 0.015            | 0.145                  | [0.55 1.50]        |

TABLE 5. Clustered damages of realization #1.

TABLE 6. Utility weights used in utility function.

FIGURE 8. Optimal health monitoring layout corresponding to the 22nd realization when (a) damage clustering and ME are not considered, (b) damage clustering and ME are considered.
are obtained using GAMS® and a BARON® optimization solvers [42], using a desktop computer with a 64-bit Windows® 10 operating system, an Intel® Core™ i7-2760QM CPU at 2.40 GHz, and 16.0 GB of RAM. Also, the damage simulation is conducted using the R programming language. (The corresponding GAMS® and R code package is available at https://github.com/aminaria/SHM2Opt.) Note that the weighting factor of the objective function \( w_{\text{obj}} \) in Eq. 6.1 is equal to 0.5 for all the layouts in Section 6.1 unless mentioned otherwise.

In Figure 8, dashed circles are the coverage boundary of the AE sensors, and hatched areas represent the inspection areas. Moreover, double hatched areas are inspected twice (Note that only inspection location, but not inspection time, is discussed in this paper and inspections can be done at different time instances). Average LPOND, which is considered in the objective function of the formulated optimization problem (Eq. 6.1) is considered in both layouts of Figure 8. However, clustering and ME are not considered in layout (a) to illustrate improvements achieved upon their consideration in layout (b). Hence, dashed lines in Figure 8(b) represent cluster boundaries.

Layouts of Figure 8 indicate that the consideration of clustering and ME in layout (b) has led to a better average utility (Table 7). This improvement can be due to a lower health monitoring cost in layout (b) with four AE sensors, instead of five, as in layout (a). Moreover, an efficient assignment of AE sensors to nodes 2 and 8 (with extensive damages and smaller ME values) in layout (b) instead of nodes 4 and 6, as is done in the ME-insensitive layout (a), can be the other reason of this improvement.

LPOND (Eq. 3 and Eq. 6.1) is a log-linear relation that is not sensitive to decimal probabilistic changes. Hence, insufficient monitoring (e.g., LPOND = \( \log(0.11) = -0.95 \)) of some areas/damages of a segment is likely while other areas are excessively monitored (e.g., LPOND \( = \log(0.11/10) = -1.95 \)). Thus, damage clustering is utilized here to provide guidelines for placement of detection methods consistent with the distribution of damages over all parts of a pipeline segment. As such, layout (b) of Figure 8 has a slightly worse average LPOND, better overall coverage, and a better average redundancy (Table 7) in comparison to layout (a).

Considering the discussion above, it can be concluded that utilizing clustering and ME has resulted in a better performance in the case of 22\textsuperscript{nd} realization. Moreover, Table 7 reveals that usage of clustering not only has not led to sub-optimality but also has led to an improved optimization runtime. As such, impacts of the utilization of clustering and ME on the optimal layouts of all 46 realizations of this example are explored next to check if these improvements are seen for other realizations as well.

3) MULTIPLE REALIZATION RESULTS AND FINAL AGGREGATE LAYOUT

Statistics of improvements achieved upon consideration of clustering and ME corresponding to all 46 realizations of this notional example (recall the beginning of Section 6.1.1) are reported in Table 8. Reported statistics indicate that the average utility is improved upon consideration of clustering and ME for the majority of the realizations. Nonetheless, the average LPOND gets slightly worse for a considerable number of the realizations, while average detection redundancy is improved. These observations can be due to more efficient placement of detection methods following utilization of clustering, as was discussed earlier. On the other hand, reported statistics for run time reveal a significant improvement for almost all the realizations when damage clustering is considered. This observation can be due to the shrinkage of the combinatorial feasible region of the optimization problem

| TABLE 7. Comparison of performance of the layouts of Figure 8. |
|---------------------------------------------------------------|
| Without clustering and ME | With clustering and ME | Improvement % |
| Ave (Utility) | Ave (Redun.) | Ave (LPOND) | Run Time (s) |
| 0.29 | 1.30 | -9.10 | 106 |
| 0.45 | 1.46 | -8.73 | 68 |
| 55% | 12% | -4% | |

| TABLE 8. Statistics of improvement percentage in layout performance for 46 realizations of example 1 when clustering and ME are considered (run time values increase with number of damages and cover the span of 11 to 445 seconds when clustering and ME are considered.) |
|---------------------------------------------------------------|
| Ave(Utility) | Ave(Reudn.) | Ave(LPOND) | Run Time |
|---------------------------------------------------------------|
| 10% quantile | -5.0% | -7.5% | 0.0% |
| 90% quantile | 25.0% | 3.0% | 210.0% |
| Average | 49.2% | 8.8% | 93.4% |
| Standard Deviation | 11.0% | 13.8% | 109.7% |

90610 VOLUME 8, 2020
following removal of cases that violate the detection limits of any damages cluster.

Considering Table 8, utilizing damage clustering in more computationally complex cases of layouts with ME reduces the run time considerably, while leading to improvements in terms of health monitoring performance (e.g., average utility and average detection redundancy). Thus, we may claim that the proposed approach provides a computationally tractable solution for consideration of probabilistic detection metrics in optimal health monitoring of pipelines and the final layout of each realization should be the one that considers clustering and ME.

4) THE FINAL AGGREGATE LAYOUT AND EVALUATION OF ITS PERFORMANCE

An aggregation of all 46 optimal health monitoring layouts of the pipeline segment of this example (Section 4, Step 12) is illustrated in Figure 9, where there are 176 AE sensors (triangles) and 54 inspection nodes (plus signs). As a result, it is expected to have 4 AE sensors and 1 human inspection with the ultrasonic tool in each health monitoring layout on average. Consequently, the final health monitoring layout is obtained (Figure 10), where sensors and human inspection are located at the center of 4 sensor clusters and 1 inspection node cluster corresponding to Figure 9.

All 46 pipeline realizations of this example, i.e., a spectrum of localized damages with different sizes and at various locations, are considered in obtaining the final layout of Figure 10. Moreover, the placement of sensors and human inspection is consistent with the uniform longitudinal and non-uniform circumferential spatial distribution of localized damages (see Section 6.1.1). Nonetheless, to verify the choice of this layout as the final layout, the performance of this layout is compared with that of (Pareto) optimal designs corresponding to the most and least expensive layouts (in terms of health monitoring total cost) of all 46 realizations. Pareto optimal designs are different layout designs corresponding to the various values of weight factor $w_{obj}$ in the objective function (Eq. 6.1) of the formulated optimization problem. In Figure 11, a longitudinal projection of Pareto designs corresponding to realizations 31 (most expensive layouts with redundant inspection) and 45 (least expensive layouts with no human inspection) are presented. Also, the design “final” represents the longitudinal projection of layout of Figure 10. In Figure 11, the triangles denote AE sensors while shaded strips denote inspection area. In addition, double shaded strips represent regions that are inspected twice.

Using the Analytic Hierarchy Process [43], the performance of the projected layouts of was compared for two newly generated test realizations (filled and hollow diamonds in Figure 12) of the pipeline segment of this example. It was concluded that the final health monitoring layout (Figure 10) outperforms all projected layouts of Figure 11 for both test realizations in terms of average utility and average LPOND.
FIGURE 11. Longitudinal projection of final health monitoring layout and Pareto designs corresponding to most and least expensive layouts of realizations of Example 1. Here, \( w \) refers to \( w_{\text{obj}} \) in Eq. 6.1.

FIGURE 12. Two newly generated realizations used as test realizations.

FIGURE 13. Health monitoring layout corresponding to one realization of the longer pipeline segment.

B. EXAMPLE 2
To show the applicability of the proposed approach for larger and more densely corroded pipelines, a pipeline segment of 200-meter length and 1-meter radius is considered. It is assumed that the longitudinal intensity of localized damages is 0.4 pit per meter. Data generation process and all other assumptions, conditions, and relations used in this example are the same as those in the Example 1 (see Section 6.1.1).

The health monitoring layout corresponding to one realization of the segment for this example is shown in Figure 13. (All the symbols in are similar to those of Figure 8.) There are 22 AE sensors and 6 inspection nodes in the optimal layout of Figure 13. This is reasonable considering the long pipeline segment and higher localized damage density. Similar to the previous example, the final health monitoring layout of this example (with 95% coverage of localized damage variations with 90% confidence) can be attained by aggregation of 46 optimal layouts corresponding to 46 realizations of the longer pipeline segment of this example.

VII. CONCLUDING REMARKS
This paper presents a new approach for health monitoring layout optimization for pipelines, with a focus on probabilistic detection of localized damages. The approach considers different detection methods (e.g., different types of sensors and human inspection tools) simultaneously. The optimal
layout obtained by the approach includes the location and type of sensors and human inspection tools/areas.

With the proposed approach, this paper makes new contributions to the existing literature on optimization-based health monitoring of locally damaged pipelines. In particular, the proposed approach has several features. (i) It uses three metrics, which together, can be used for better probabilistic detection of damages with different levels of severity. As shown in the paper, the use of these metrics can reduce the cost of health monitoring. (ii) It is based on an optimization objective function which is formed based on a weighted sum of two functions: a health monitoring utility function and a function for probabilistic detection of damages. By changing the weights for these functions, the user can explore tradeoffs between the two functions and thus obtain different optimized layout solutions as desired by the decision-maker. (iii) It considers a significant number of key attributes of the detection methods, such as detection cost, coverage capability, data acquisition frequency, and measurement error. Compared to the existing literature, these will provide a more detailed account of the detection methods considered. (iv) It uses probabilistic sampling methods for simulating and placing damages on a pipeline surface, considering relevant available data. In this way, the proposed approach can be used for pipelines with stochastic localized damages while not only the expected value of damage specifications (e.g., damage location and size), but also their variations can be accounted for.

The proposed approach is demonstrated with two notional examples. For the first example, a step-by-step demonstration of the proposed approach is provided. For this example, considering the severity level of different damages, it is shown that the proposed layout optimization approach can obtain a better solution than a single detection method approach. For the second example, a longer pipeline segment with more densely localized damages is considered to show the applicability of the proposed approach to larger sized problems.

REFERENCES

[1] E. Rabiei, E. L. Drogueut, M. Modarres, and M. Amiri, “Damage precursor based structural health monitoring and damage prognosis framework,” Saf. Rel. Complex Eng. Syst., vol. 23, pp. 2441–2449, Sep. 2015.
[2] H. Bruske and S. Thöns, “Fact sheet on the domains of the value of information in structural health monitoring,” Tech. Univ. Denmark, Lyngby, Denmark, Tech. Rep. TU1402, 2017. [Online]. Available: https://www.researchgate.net/profile/Henning_Brueske/publication/333776045_Fact_sheet_on_the_domains_of_The_Value_of_Information_in_structural_Health_Monitoring/links/5d037e8d458515b055cf0902/0902_Fact-sheet-on-the-domains-of-the-Value-of-Information-in-Structural-Health-Monitoring.pdf
[3] H. Iqbal, S. Tesfamariam, H. Haider, and R. Sadig, “Inspection and maintenance of oil & gas pipelines: A review of policies,” Struct. Infrastruct. Eng., vol. 13, no. 6, pp. 794–815, 2017.
[4] S. Alaswad and Y. Xiang, “A review on condition-based maintenance optimization models for stochastically deteriorating system,” Rel. Eng. Syst. Saf., vol. 157, pp. 54–63, Jan. 2017.
[5] S. N. Zulkifli, H. A. Rahim, and W.-J. Lau, “Detection of contaminants in water supply: A review on state-of-the-art monitoring technologies and their applications,” Sens. Actuators B, Chem., vol. 255, pp. 2657–2689, Feb. 2018.
[6] R. Duan, Y. Lin, and T. Feng, “Optimal sensor placement based on system reliability criterion under epistemic uncertainty,” IEEE Access, vol. 6, pp. 57061–57072, 2018.
[7] S. Zhang and W. Zhou, “Cost-based optimal maintenance decisions for corroding natural gas pipelines based on stochastic degradation models,” Eng. Struct., vol. 74, pp. 74–85, Sep. 2014.
[8] F. Aldurabi, N. Lasla, and M. Younis, “Coverage-based node placement optimization in wireless sensor network with linear topology,” in Proc. IEEE Int. Conf. Commun. (ICC), May 2016, pp. 1–6.
[9] Y. Sun and S. Halgamuge, “Minimum-cost heterogeneous node placement in wireless sensor networks,” IEEE Access, vol. 7, pp. 14847–14858, 2019.
[10] C. Argyris, S. Chowdhury, V. Zabel, and C. Papadimitriou, “Bayesian optimal sensor placement for crack identification in structures using strain measurements,” Struct. Control Health Monitor., vol. 25, no. 5, p. e2137, May 2018.
[11] A. Verzobio, D. Bolognani, D. Zonta, and J. Quigley, “Quantifying the benefit of structural health monitoring: Can the value of information be negative?,” Struct. Health Monit., p. e32266, 2019. [Online]. Available: http://www.dpi-proceedings.com/index.php/shm2019/article/view/32266, doi:10.12783/shm2019/32266.
[12] D. Straub et al., “Value of information: A roadmap to quantifying the benefit of structural health monitoring,” in Proc. 12th Int Conf. Struct. Syst. Rel. (ICOSAR), Vienna, Austria, Aug. 2017.
[13] L. Iannacone, P. Gardoni, P. F. Giordani, and M. P. Limongelli, “Decision making based on the value of information of different inspection methods,” in Proc. 12th Int Workshop Struct. Health Monit., Enabling Intell. Life-Cycle Health Manage. Ind. Internet Things (IOT), Lancaster, PA, USA: DEStech Publications, 2019, pp. 1479–1486.
[14] W. J. Klerk, T. Schweekendiek, F. den Heijer, and M. Kok, “Value of information of structural health monitoring in asset management of flood defences,” Infrastructures, vol. 4, no. 3, p. 56, 2019.
[15] S. Abdollahzadeh and N. J. Navimipour, “Deployment strategies in the wireless sensor network: A comprehensive review,” Comput. Commun., vol. 91–92, pp. 1–16, Oct. 2016.
[16] B. W. Brunton, S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Sparse sensor placement optimization for classification,” SIAM J. Appl. Math., vol. 76, no. 5, pp. 2099–2122, Jan. 2016.
[17] Y. L. Guo, Y. Q. Ni, and S. K. Chen, “Optimal sensor placement for damage detection of bridges subject to ship collision,” Struct. Control Health Monit., vol. 24, no. 9, p. e1963, Sep. 2017.
[18] A. P. Berens, “NDE reliability data analysis,” in ASM Handbook: Nondestructive and Quality Control, vol. 17. Materials Park, OH, USA: ASM International, 1989, pp. 689–701.
[19] X. Liu, J. Zheng, J. Fu, Z. Nie, and G. Chen, “Optimal inspection planning of corroded pipelines using BN and GA,” J. Petroleum Sci. Eng., vol. 163, pp. 456–555, Apr. 2018.
[20] K. Chatterjee and M. Modarres, “A probabilistic approach for estimating defect size and density considering detection uncertainties and measurement errors,” Proc. Inst. Mech. Eng., O, J. Risk Rel., vol. 227, no. 1, pp. 28–40, Feb. 2013.
[21] M. Abyani and M. R. Baharai, “A comprehensive reliability study of corroded pipelines based on Monte Carlo simulation and latin hypercube sampling methods,” Int. J. Pressure Vessels Piping, vol. 181, Mar. 2020, Art. no. 104079.
[22] D. Benstock and F. Ceglá, “Sample selection for extreme value analysis of inspection data collected from corroded surfaces,” Corrosion Sci., vol. 103, pp. 206–214, Feb. 2016.
[23] M. Shafee and J. D. Sorensen, “Maintenance optimization and inspection planning of wind energy assets: Models, methods and strategies,” Rel. Eng. Syst. Saf., vol. 192, no. 1, pp. 1–19, Dec. 2019.
[24] R. Hou, Y. Xia, Q. Xia, and X. Zhou, “Genetic algorithm based optimal sensor placement for L1-regularized damage detection,” Struct. Control Health Monit., vol. 26, no. 1, p. e2274, 2019.
[25] W. Ostachowicz, R. Soman, and P. Malinowski, “Optimization of sensor placement for structural health monitoring: A review,” Structural Health Monit., vol. 18, no. 3, pp. 963–988, May 2019, doi:10.1177/1475927119825601.
[26] M. G. R. Sause, F. F. Linscheid, and M. Wiehler, “An experimentally accessible probability of detection model for acoustic emission measurements,” Int. J. Pressure Vessels Piping, vol. 74, no. 5, pp. 493–506, Apr. 1999.
[27] K. Knop, E. Olejarcz, and R. Ulewicz, “Evaluating and improving the effectiveness of visual inspection of products from the automotive industry,” in Proc. Int. Sci.-Tech. Conf. Manuf. Cham, Switzerland: Springer; May 2019, pp. 231–243.
AMIN ARIA received the B.Sc. degree in mechanical engineering from the Sharif University of Technology, Tehran, Iran. He is currently pursuing the Ph.D. degree in reliability engineering with the University of Maryland, College Park, USA. His research interests include prognosis and health management, operations research, stochastic optimization, statistical analysis, and machine (deep) learning. His articles on applications of operations research and deep learning in prognosis and health management of machinery have been published in journals and conference proceedings.

MOHAMMAD MODARRES received the M.S. degree in mechanical engineering and the Ph.D. degree in nuclear engineering from MIT. He is currently the Nicole Kim Eminent Professor and the Director of the Center for Risk and Reliability with the A.J. Clark School of Engineering, University of Maryland, College Park. He is also an Expert in probabilistic risk assessment, probabilistic physics of failure and probabilistic fracture mechanics. His interests in risk, reliability, structural integrity and prognosis and health management include both experimental and probabilistic model development efforts. His recent researches involve multi-unit risk assessment of nuclear plant sites, integration of data analytics with physics of failure for condition monitoring. He has extensive publications in archival journals and proceedings of conferences, including textbooks and book chapters in various areas of risk and reliability engineering. He is a University of Maryland Distinguished Scholar-Teacher and a fellow of the American Nuclear Society.

SHAPOUR AZARM received the Ph.D. degree from the University of Michigan, Ann Arbor. He is currently a Professor of mechanical engineering and an affiliate Professor of applied mathematics and statistics and scientific computation program. His research interests are in engineering optimization and decision-making models and methods. He is a fellow of ASME and a Senior Member of AIAA. Among numerous awards, he received the ASME Design Automation Award for his sustained and meritorious contributions to research in design automation, specifically in computational design optimization and engineering decision making. He is a Review Editor and a Senior Advisor of the Structural and Multidisciplinary Optimization. He was the Editor-in-Chief of the Journal of Mechanical Design (ASME Transactions) and an Associate Editor of the Mechanics-Based Design of Structures and Machines.