Leveraging Machine Learning for Predictive Maintenance of Lock and Dam Infrastructure Using Real-World Data

DANIA AMMAR[1], DAN AO[1], SARAH MIELE[1],
GUGA GUGARATSHAN[1] and BRIAN EICK[2]

ABSTRACT

Structural health monitoring (SHM) of water resource infrastructure, such as locks and dams, requires innovative approaches due to their complex environments, accessibility challenges, and variable loading conditions. Machine learning (ML) plays a significant role in detecting anomalies and enabling predictive maintenance capabilities. ML provides a data-driven approach to identify potential structural issues before failure occurs. This study enhances an ensemble-based anomaly detection framework by integrating advanced feature extraction techniques and sensor fusion methods to improve the monitoring and assessment of lock and dam health. A time-series segmentation approach is utilized to divide sensor data into discrete periods, which are then characterized by a set of statistical features. These features serve as inputs to a Principal Component Analysis (PCA), allowing for effective dimensionality reduction. Advanced feature-level fusion is achieved by leveraging multi-channel strain sensor data, leading to improved anomaly detection and system reliability. This technique enables the integration of data from multiple sensors into a unified framework, offering a comprehensive assessment of structural conditions. An ensemble of ML-based anomaly detection methods is applied to recognize deviations in sensor data patterns, providing an early warning system for potential structural concerns. To refine classification accuracy, the anomaly detection system is supplemented with existing models. This hybrid strategy enhances the interpretation of detected anomalies, reducing false positives and improving the reliability of SHM systems. By merging data-driven ML techniques with well-established physics models, the proposed approach offers a more robust method for predictive maintenance in lock and dam infrastructure. The framework is validated using real-world strain sensor data from a lock gate, where an existing crack is a primary concern for the maintenance team. This study presents a flexible and scalable predictive maintenance framework tailored for lock and dam infrastructure. The proposed approach emphasizes the efficient deployment of SHM methods in real-world scenarios, addressing a critical challenge in the field by providing a practical pathway for implementing SHM solutions effectively.

[1]Hottinger Bruel & Kjaer Solutions LLC, 100 Research Blvd., Suite 216, Starkville, MS 39759
[2]Construction Engineering Research Laboratory, US Army Corps of Engineers, Champaign, IL 61822

INTRODUCTION

The U.S. Army Corps of Engineers (USACE) currently oversees a total of 237 locks and dams, which are distributed across 192 commercial lock sites [1]. These locks and dams are pivotal structures that facilitate the transportation of various types of goods by providing a medium for cargo navigation in inland waterways [2]. Maintenance is essential to ensure their uninterrupted operation and to avoid unexpected closures resulting from events such as structural damage, ultimately leading to the impediment of the supply chain and severe economic losses.

To protect locks and dams and prevent shutdowns, it is important to maintain their lock gates, which play the role of a damming surface across the lock chamber. Numerous gate designs exist including vertical lift gates, tainter gates, sector gates, rolling gates, tumbler gates, and miter gates [3]. Miter gates possess high operational efficiency and require lower maintenance requirements as compared to other types of gates [3] and are the most commonly used in the U.S. inland waterway system's locks and dams [4].

To ensure the continued reliability of these gates, engineers employ various maintenance strategies, including field observations and structural health monitoring (SHM). Field observations are associated with high cost, prolonged durations, and high risks [4], and as a result, SHM has gained attention as an improved alternative for human inspections. SHM is performed to continuously monitor the structure's health and to detect any possible structural damage using sensors [5]. This real-time monitoring process has the advantages of detecting potential issues early on, enabling informed decision makings and rapid interventions to prevent additional damage [4].

There are numerous examples of USACE lock and dam structures employing SHM on their miter gates, including Lock and Dam 27 of the Mississippi River, and the Bonneville and Dalles Lock and Dam of the Columbia River [6]. In Lock and Dam 27, SHM was capable of identifying contact degradation between the lateral edge of the lock gate and the wall of the lock chamber [7]. In contrast, SHM results in the Greenup lock gate on the Ohio River captured the presence of a gap damage-type developing in the quoin, indicating a loss of bearing contact between the lock wall and a portion of the gate [6, 8]. With the benefit of an SHM system and preventative maintenance, the Meldahl miter gate on the Ohio River suffered an extreme case of quoin gap, leading to extensive cracking in the pintle region where the hydrostatic load was concentrated upon its redistribution [9].

The analysis of data from SHM systems is typically conducted using Predictive Maintenance (PdM) algorithms. PdM involves the transformation of raw sensor data into valuable information. The insights, trends, and patterns drawn from raw sensor data using PdM are at the essence of scheduling preemptive maintenance repairs when needed before the issues escalate to failures. In this regard, PdM algorithms can be classified into three groups: physics-based, data-driven, and hybrid [10]. In the area of data-driven algorithms, PdM employs advanced Artificical Intelligence (AI) and ML models to detect deviations from established patterns and expected behaviors, also known as anomalies [11]. These models are divided into two main categories, supervised and unsupervised. In supervised PdM, the input data encompasses labeled events, unlike unsupervised PdM, which lacks labeled events for evaluation. Applying unsupervised PdM on miter gates, Eick et al. [6] utilized a PCA to capture significantly different slope measures in the Greenup lock and dam and validated the approach using a finite element model. In [6], the slope measure, derived as a gap damage-sensitive feature, is calculated as the change in strain values as a function of load resulting from chamber water level during filling or emptying events. Another study then used Bayesian neural networks to estimate the gap depths for the same structure [9].

This paper is focused on the application of a fully data-driven PdM algorithm on a miter gate structure using an ensemble anomaly detection technique. While different anomaly detection methods have varying strengths and weaknesses, combining results from multiple algorithms reduces the risk of false positives, making the anomaly detection framework more robust and reliable. Furthermore, this work integrates sensor fusion, such that the records of multiple sensors are analyzed jointly after the extraction of time-segmented features, designating key characteristics of raw signals. The ensemble anomaly detection results aim to provide a single, intuitive metric for maintenance engineers to extract actionable insights from the data. The daily predictions resulting from anomaly detection are further analyzed and compared with the original sensor data and other exogenous variables, providing additional insights into the reliability of the developed ensemble learning.

METHODOLOGY

The methodology for this research is split into five steps. (1) Data is preprocessed to retain strain responses that correspond to either filling or emptying events of the lock chamber. (2) Several statistical features are extracted for each sensor after applying a fixed length window time series segmentation. (3) Features are first normalized and are then reduced in dimension through the application of PCA to optimize the performance of the ML algorithms. (4) Five different anomaly detection methods are unified in an ensemble learning framework and applied to the selected joint sensor data components. (5) The data-driven results of the test set undergo an in-depth evaluation in attempt to support the PdM decisions by comparing with individual original strain sensor data, measured exogenous variables, and previously established physics-driven damage methods. Figure 1 illustrates the top-level block diagram of the data analysis pipeline.

Before training ML algorithms for anomaly detection, various preprocessing steps are applied to clean the data. These steps include handling missing values, removing duplicates, standardizing formats, eliminating unnecessary or corrupt columns, and addressing anomalies caused by data quality issues. Next, the data undergo a filtering process to remove time periods where no filling or emptying events are happening. This step aims to remove strain values corresponding to time periods outside of lockage events when the change in strain values is minimal, thus yielding negligible relevance.

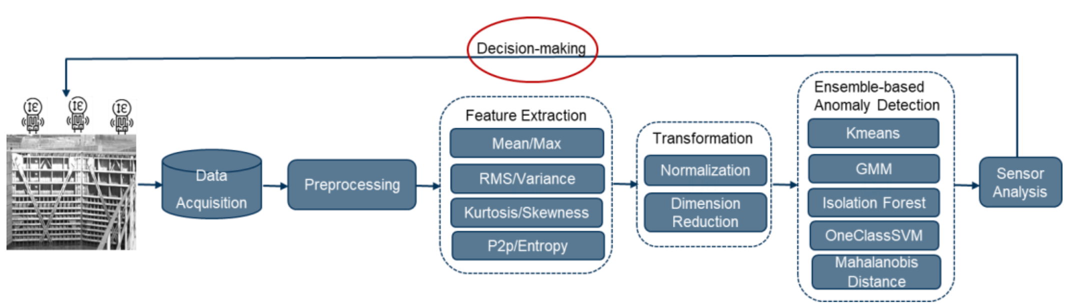

Figure 1. Visualization of the data analysis pipeline

Feature extraction is performed in the time domain across each time interval to make these data more manageable and informative for model training. A moving window is applied to segment the time-series signals and extract statistical features from each column within every segmented interval. For each column, eight statistical features are computed: absolute mean, absolute maximum, standard deviation, root mean square, kurtosis, skewness, peak-to-peak amplitude, and entropy. The moving window is configured to use non-overlapping time windows in this study to streamline the process and maintain clarity in our approach. To reduce the dimension of the extracted features, PCA is applied. PCA transforms the large set of features into a smaller set of variables while retaining important information.

In this study, five popular anomaly detection algorithms are explored: K-Means Clustering, Gaussian Mixture Model (GMM), Isolation Forest, one-class Support Vector Machine (SVM), and Mahalanobis Distance. K-Means Clustering and GMMs group data points based on similarities, either by distance or density, with outliers identified by their deviation from the cluster. Although Isolation Forest and one-class SVM are derived from supervised learning, they are applied in an unsupervised context. Isolation Forest focuses on the "few and different" nature of anomalies [12], while one-class SVM defines a hypersphere to separate normal data points from anomalies. Mahalanobis Distance is a robust multivariate metric that quantifies the distance between a data point and a distribution, making it effective for identifying outliers that differ significantly from the normal distribution. However, each of these anomaly detection algorithms comes with its own limitations. For example, K-Means Clustering struggles when clusters have irregular shapes [13], and Isolation Forest tends to detect only local anomaly points [14]. The performance of any algorithm is highly dependent on the underlying structure of the data. To avoid the dilemma in choosing the optimal method in the absence of evaluation data and ensure reliable and robust results, an ensemble approach is proposed.

A critical component of the anomaly detection system is the decision-level fusion process. It combines the outputs of the above anomaly detection algorithms—each contributing with varying confidence levels—into a unified anomaly indicator. This aggregated signal is then normalized to produce a final anomaly score. To optimize the ensemble's performance, individual algorithm outputs are weighted according to their respective confidence levels. The system generates a data frame containing timestamps and the corresponding ensemble-based anomaly scores. Figure 2 illustrates how each anomaly detector assigns a binary label to each data instance: '0' for 'no anomaly detected' and '1' for 'anomaly detected.' To compute the normalized anomaly score at a given instance, a weighted average of these binary labels is calculated across all detectors, as detailed below:

$$Score = \left(\sum_{i=1}^{n} w_i f_i\right) \Big/ \left(\sum_{i=1}^{n} w_i\right) \tag{1}$$

Here, for a total of n anomaly detectors, w_i represents the user-defined weight assigned to the i'th detector, and f_i denotes the binary label produced by the i'th detector. The resulting normalized anomaly score ranges from 0 to 1.

To validate the described methodology, a thorough analysis is performed to uncover any underlying correlation between the scores and the measured sensor data. The analysis involves comparisons between the anomaly scores and the strain data, along

Schema of the Anomaly Detection System

Anomaly Detector	Outlier Assessment										
K-means Clustering	0	0	…	0	1	…	1	1	…	1	1
Gaussian Mixture Model	0	0	…	0	0	…	0	1	…	1	1
Isolation Forest	0	0	…	0	0	…	1	1	…	1	1
OneClassSVM	0	0	…	0	0	…	1	1	…	1	1
Mahalanobis Distance	0	0	…	0	0	…	0	0	…	0	1
Normalized Score	**0**	**0**	**…**	**0.125**	**0.25**	**…**	**0.5**	**0.75**	**…**	**0.875**	**1**

Ensemble: Weighted Indicator Averaging

Figure 2. Schema of the ensemble-based anomaly detection system

with other exogenous variables. First, the anomaly scores are compared with the measured strain signal of all selected sensors. This step is indispensable for a preliminary visual investigation, exploring whether a high anomaly score detected across the components is associated with a deviation in the original data. More specifically, the daily resampled strain data is inspected for particular trends or patterns associated with the evolution of the daily anomaly score with time. Second, the relationship between strain sensors and water and air temperatures is also explored to determine whether the evolution of anomalies could be partly driven by environmental changes. The correlation between the resampled daily strain and temperature values is considered. Finally, the slope depicting the variation in strain value as a function of chamber water level is calculated in accordance with Eick et al. [6]. In this step, strain values for each filling and emptying events are isolated. For each event, the slope designating the rate of change in strain with respect to chamber water level is calculated by fitting a first-degree polynomial using the least square method. The slope is also examined at a day level by considering the absolute mean and absolute maximum values. Similar to previous efforts, the correlation between the anomaly score and calculated slope is then inspected.

RESULTS

To test the developed framework, we used a real example miter gate data from an in-service lock and dam. The miter gate is equipped with several strain sensors in addition to other sensors measuring exogenous variables (e.g., chamber water level, water temperature, air temperature). Since a higher load is typically exerted at the bottom of the miter gate, six strain sensors, corresponding to this region, were included

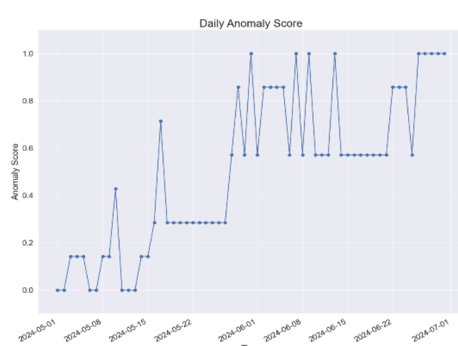

Figure 3. Results of the ensemble-based anomaly detection model

in this study. The sensor locations cover the pintle region and girders at the bottom of the gate. Remaining sensors in the same region were excluded due to severe data quality issues as captured by checks during data preprocessing. For the selected six sensors, data between November 2023 and June 2024 was employed for this analysis. After running PCA, five eigenvectors corresponding to five eigenvalues were selected such that they captured 93% of the variability. The selected components corresponding to November through April were used to train the ensemble algorithm, whereas the selected components for the months of May and June were used for testing.

The resultant anomaly score for each day during the months of May and June are displayed in Figure 3. During the first half of May, the daily anomaly score is relatively low, with a few days exhibiting slightly higher scores. In the second half of May, the daily anomaly score escalates to become relatively higher than in the first half of the same month, with the score being almost constant across multiple days. This could be indicative of some sensors beginning to pick up an anomalous pattern in the structural response for consecutive days. After that, the anomaly score significantly increases, hitting a value of 1 at multiple days during the month of June. This suggests that all of the sensors are picking up the anomalous behavior as it propagates through the structure with time. Although there is a general amplification in the score, there are some fluctuations that can be observed. Such instances could be related to multiple factors, including data quality issues, data traces during preprocessing, or other external factors.

Upon comparing the anomaly score with daily resampled strain sensor data, we examined that some strain sensors experienced an incremental drift with an upward trend starting mid-May (Figure 4). This pattern is in parallel with an increase in the anomaly score during the same period of time. The sensors exhibiting this relationship are Sensor1, Sensor3, Sensor4, and Sensor6.

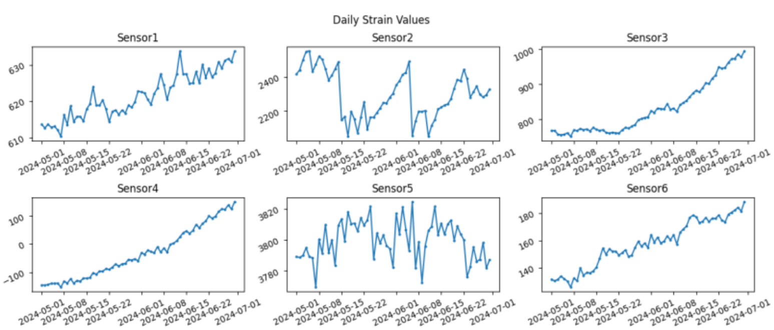

Figure 4. Daily resampled strain values

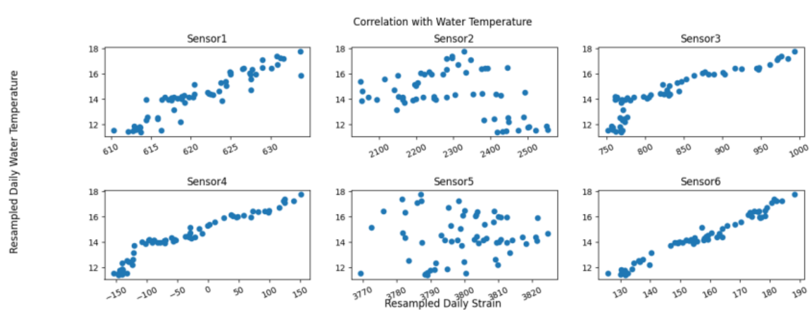

Figure 5. Relationship between daily resampled strain values and water temperature

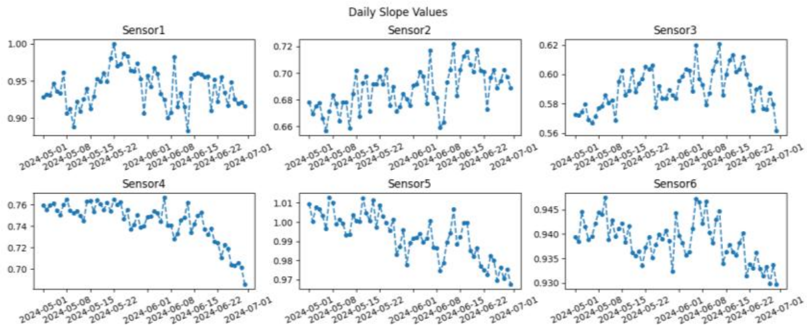

Figure 6. Absolute mean daily slope damage-sensitive feature

A consistent pattern between four individual strain sensors and the water temperature was identified. As shown in Figure 5, Sensor1, Sensor3, Sensor4, and Sensor6 demonstrate a positive correlation with water temperature. Such results may point to the need for retraining the anomaly detection model during each season to reduce variability pertaining to environmental factors.

As for the relationship between the anomaly score and the slope measure, being a damage-sensitive feature, the computed daily absolute mean slope for each strain sensor is shown in Figure 6. Results did not reveal a clear relationship between the daily anomaly score and the daily slope calculated for each individual strain sensor. Moreover, we did not observe a consistent pattern across the slopes of individual sensors, except for a strong correlation between Sensor4 and Sensor5 and a moderate correlation between Sensor6 on one end and Sensor4 and Sensor5 on the other end.

CONCLUSIONS

This study presents efforts to detect anomalous multi-sensor behavior through a comprehensive data-driven approach encompassing ensemble learning and validation through physics-based relationships. The anomaly detection results exhibited an anomalous behavior becoming more salient with time, eliciting its consistency across the distinct sensors and its propagation along the isolated region. Moreover, the correlation between raw sensor behavior and the anomaly score as a function of time underscores the effectiveness of PCA's multi-sensor data fusion technique in making anomaly detection algorithms efficient and reliable rather than conforming to large quantities of raw data. While the current examination of the relationship between strain and some exogenous variables uncovered some relationship between water temperature and strain sensors exhibiting a strong correlation with anomaly scores, the conformity between the anomaly score and the incremental drift observed in some sensor data could also indicate a localized plastic deformation or fatigue cracking on the upstream/skin plate, where these sensors are located. Moreover, different failure mechanisms have different physical randomness. Therefore, the validation of this PdM framework could be further enhanced using more complex physics-based models, such as an FE model or a fast emulator of an FE model being a surrogate model, capable of accurately assessing the damage mode in the miter gate at hand. Nevertheless, the physics model could benefit from a reduction in its computational effort by using data that corresponds to sensors exhibiting a consistent anomalous behavior throughout the testing time period. As a result, the surrogate model could be limited to the upstream/skin plate.

ACKNOWLEDGEMENTS

This work has been funded by the United States Army Corps of Engineers through the U.S. Army Engineer Research and Development Center's (ERDC) Construction Engineering Research Laboratory. The authors would especially like to thank Clayton Thurmer from the ERDC, The University of California San Diego (UCSD), and Professor Michael Todd for their support on this project. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the UCSD or ERDC.

REFERENCES

1. USACE (United States Army Corps of Engineers). 2023. "The U.S. Coastal and Inland Navigation System-Transportation Facts & Information," *Navigation and Civil Works Decision Support Center*, Alexandria, VA.
2. Casavant, K., M. Denicoff, E. Jessup, A. Taylor, D. Nibarger, D. Sears, H. Khachatryan, V. McCracken, E. Jessup, M. Prater, J. O'Leary, N. Marathon, B. McGregor, and S. Olowolayemo. April 2010. "Study of Rural Transportation Issues," *U.S. Department of Agriculture, Agricultural Marketing Service*.
3. USACE (United States Army Corps of Engineers). September 1995. "Planning and Design of Navigation Locks," (Engineer Manual No. EM 1110-2-2602)
4. Dixon, D.L. and W.D. Hodo. June 2021. "Finite element analysis of quoin block deterioration and load transfer mechanisms in miter gates: Pintle and pintle connections," (Report No. ERDC TR-21-9). US *Engineer Research and Development Center*.
5. Tian, W., S. Weng, Y. Xia, H. Zhu, F. Gao, Y. Sun, and J. Li. 2019. "An iterative reduced-order substructuring approach to the calculation of eigensolutions and eigensensitivities," *Mech. Syst. Signal. Pr.*, 130, 361-377.
6. Eick, B. A., Z.R. Treece, B.F. Spencer, M.D. Smith, S.C. Sweeney, Q.G. Alexander, and S.D. Foltz. 2018. Automated damage detection in miter gates of navigation locks. *Structural Control and Health Monitoring*, *25*(1).
7. Treece Z., M. Smith, N. Wierschem, S. Sweeney, B. Spencer. 2015. "USACE SMART gate: structural health monitoring to preserve America's critical infrastructure," *Structural Health Monitoring*
8. Vega, M. A., Z., Hu, T.B., Fillmore, M.D., Smith, and M.D. Todd. 2021. "A novel framework for integration of abstracted inspection data and structural health monitoring for damage prognosis of Miter Gates," *Reliability Engineering & System Safety*, 211, 107561.
9. Hoskere, V., B. Eick, B.F. Spencer, M.D. Smith, and S.D. Foltz. 2019. "Deep Bayesian neural networks for damage quantification in miter gates of navigation locks," *Structural Health Monitoring*; 19(5):1391-1420.
10. Kan, M.S., A.C.C. Tan, and J. Mathew. 2015. "A review on prognostic techniques for non-stationary and non-linear rotating systems", *Mechanical Systems and Signal Processing*, 62, 1-20.
11. Wen Y., M.F. Rahman, X. Lu, and T-L.B Tseng. 2022. "Recent Advances and Trends of Predictive Maintenance from Data-driven Machine Prognostics Perspective," *Measurement*, 187 110276.
12. Xu, H., et al., 2023. "Deep isolation forest for anomaly detection," *IEEE Transactions on Knowledge and Data Engineering.*
13. Singh, K., D. Malik, and N. Sharma, 2011. "Evolving limitations in K-means algorithm in data mining and their removal," *International Journal of Computational Engineering & Management*, 12(1): p. 105-109.
14. Gao, R., et al. 2019. "Research and improvement of isolation forest in detection of local anomaly points," *Journal of Physics: Conference Series*. IOP Publishing.