❏   1439

Comparison of structural analysis and principle component analysis for leakage prediction on superheater in boiler

Eko Mursito Budi[1], Estiyanti Ekawati[1], Bobby Efendy[2]
[1]Engineering Physics Research Group, Faculty of Industrial Technology, Institut Teknologi Bandung, Bandung, Indonesia
[2]Instrumentation & Control Master Program, Faculty of Industrial Technology, Institut Teknologi Bandung, Bandung, Indonesia

Article Info	ABSTRACT
	Leakage is one of the failures which commonly happens in boiler operation. Moreover, a continuous unsettled anomaly in a boiler could lead to leakage failure. An algorithm has been developed to predict the failure, consisting of three general procedures: feature selection, followed by hierarchical clustering, and naïve Bayes classification. The hierarchical clustering changes unlabeled data into labeled data, and naïve Bayes classification calculates the probability to justify anomaly occurrence. Meanwhile, this research focused on the effect of the feature selection method on the result of leakage prediction. Two different feature selection methods, namely the structural analysis and the principal component analysis (PCA), were deployed separately and then compared. The result showed that leakage prediction using the structural analysis method gave 13 hours 40 minutes of prediction time, and the PCA method gave 25 hours of prediction time. However, the PCA feature selection method caused more false alarms than feature selection with structural analysis, which only triggered five false alarms a week before leakage. Moreover, the structural analysis offered better traceability than PCA to understand the leakage occurrence.
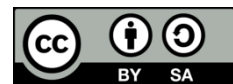
Corresponding Author:

Eko Mursito Budi
Engineering Physics Research Group, Faculty of Industrial Technology, Institut Teknologi Bandung
Jl. Ganesha 10, Bandung 40132, Indonesia
Email: mursito@itb.ac.id

1. INTRODUCTION

Process monitoring is one of the important aspects in industries. It helps the operators to evaluate the equipment or component. The information received from process monitoring can be appropriately managed to bear helpful information, such as predicting failure [1]. However, several aspects need to be considered before predicting a component's failure, such as the parameters describing the related failure, the characteristics of the failure itself, and the method to specify the trend of the degradation [2].

The research object is the boiler failure prediction, specifically boiler leakage prediction. Leakage on the boiler usually happens in the superheater, where the process of extreme combustion takes place [3]–[6]. The extended anomalous trend in the process monitoring, especially the unresolved anomaly, tends to cause the related component or equipment failure if it is not well controlled [2], [7]–[9]. For that reason, superheater leakage can be predicted by detecting the early anomaly in the related area.

A related study about principal component analysis (PCA) and the clustering process was also conducted to detect anomalies in a marine engine system without comparison to the maintenance logs [10]. It used PCA to reduce 24 variables into seven principal components. However, the implementation of PCA and the absence of maintenance logs resulted in difficulty interpreting the clustered data. Because of that, maintenance logs were fully considered in this practice to help process failure understanding.

This study aims to acknowledge how different feature selection methods could make a significant difference in the leakage prediction result. Two methods were chosen, which are structural analysis and PCA. Combined with the clustering and classification processes, these methods were used to detect the superheater anomaly. This research used hierarchical clustering and naïve Bayes classification, respectively, as the methods for clustering and classification. These methods successfully have been used in the same case of the leakage in boiler superheater and predicted the leakage 13 hours 40 minutes earlier using three selected variables by structural analysis in the previous research [11].

This research will offer a clear view about the comparison of the "black box" (PCA) and the "grey box" (structural analysis) feature selection would make a significant impact on the method ability (hierarchical clustering and naïve Bayes classification) to predict superheater leakage. The structural analysis depends on understanding the combustion process in a boiler to select process variables and determine how the variables contribute to a failure [5]. On the other hand, PCA offers a more straightforward approach, where the operator can choose any variables and does not have to ponder the principle of the boiler process [12]. Both feature selection methods have each strength and weaknesses. Briefly, this research will provide some comparative studies on structural analysis and PCA in selecting features for leakage prediction in superheater boilers. The comparison includes how long each method can predict the leakage, how many false alarms are triggered during the process, and how they interpret the leakage.

2. RESEARCH METHOD

The flow diagram of this research method is shown in Figure 1. This section will discuss the primary procedure in the flow diagram, such as structural analysis, PCA, hierarchical clustering, and naïve Bayes classification. The first step was acquiring the dataset. The dataset was taken during a similar load of the boiler to increase the consistency of the measurement because the different loads on the boiler will cause dissimilar the mean of regular operation [4]. Then the data underwent two processes separately: structural analysis and PCA. In the first processs, some variables were selected as input variables for the clustering and classification process in structural analysis. On the other process the data went through PCA and transformed into lower-dimensional data as the input and then processed for clustering and classification tasks. Finally, both calculations were compared and analyzed.

2.1. Data acquisition

This research uses primary data taken from the real-time data logging in the hydroelectric power plant. This research used two datasets: a dataset containing normal data operation (with a minor disturbance) and another dataset containing major leakage failure in the pipeline that caused the unit to be shut down. Each dataset has 1,008 samples with an interval of 10 minutes.

2.2. Structural analysis

The structural analysis relies on understanding boiler process operation to pick reliable variables for the following process. A brief illustration of the boiler can be seen in Figure 2, and the available measurements are listed in Table 1. How the leakage was triggered becomes valuable information to choose the potential variables to detect the anomaly that initiates the leakage. In many boiler leakage cases, some were mainly caused by overheating on the piping that disturbs the heat transfer to the steam and consequently causes the pipe to rupture due to the overheating process [5], [13]. Overheating occurs in areas where the combustion process plays a huge role, such as superheater, reheater, and so on [14]. Based on the available measurement data and the physics phenomenon within the process, three variables were selected based on structural analysis from the list in Table 1. PSH (primary superheater) inlet temperature, PSH outlet temperature, and SSH (secondary superheater) inlet temperature. Figure 3 shows a brief diagram of how the three selected variables integrate with the process component in the boiler.

2.3. PCA

The principal component analysis is a statistical method reducing high dimensional data into the lower one with keeping the essential variation and any important information in it [15], [16]. Moreover, by using PCA in-process monitoring, any knowledge before the system is unnecessary because of its ability to handle the high degree of correlation for each variable. It can also reduce the computational cost by decreasing data dimensions [17].

Research related to PCA was conducted to perform anomaly detection for a boiler generally during a load changing in boiler operation and successfully detect abnormality during the load transition [18]. However, it can not give a clear physical or cause-effect explanation about what happened during the transition. This research used 12-dimensional variables of boiler operation and then was transformed into the

lower dimensional dataset, which principal components will represent. The number of principal components were determined by how many components required to obtain at least 90% variance of the actual data [12]. The list of the 12 boiler operative variables used in this research is shown in Table 1. It shows that the efficient number of principal components is 6, which can obtain at least 90% of the variance. Table 2 gives the information of principal components and the variance.
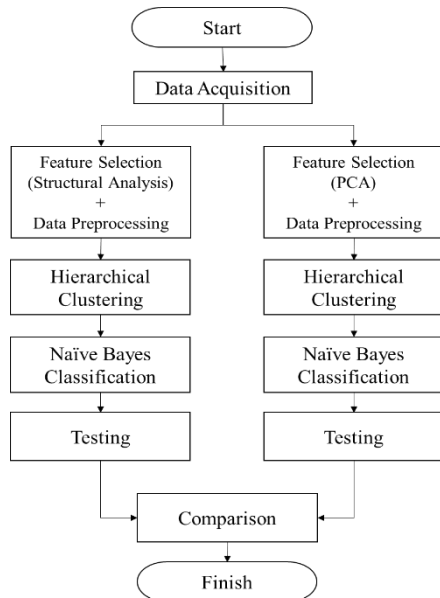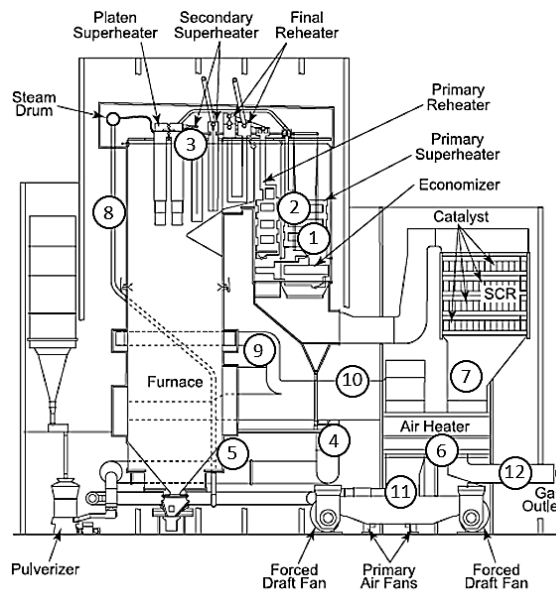

Figure 1. Flow diagram of the research

Figure 2. Position of the sensors in the boiler

Table 1. List of operative variables in the boiler

| No	Variable
1	PSH: inlet temperature
2	PSH: outlet temperature
3	SSH: inlet temperature
4	Primary air heater: outlet pressure
5	Primary air duct: pressure
6	Primary air heater: inlet air temperature
7	Primary air heater: flue gas temperature
8	Riser to steam drum: water temperature
9	Secondary air duct: pressure
10	Secondary air heater: inlet air temperature
11	Secondary air heater: inlet flue gas temperature
12	Secondary air heater: outlet flue gas temperature
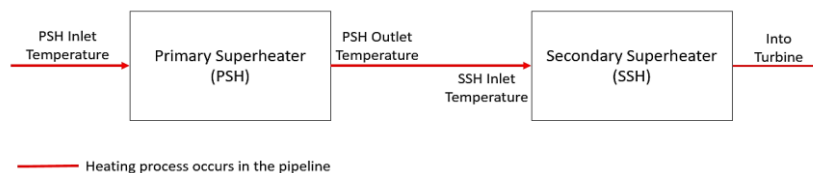
Figure 3. Block diagram of the superheater

Table 2. Accumulative variance on principal components

Principal Component (PC)	Accumulative Variance (%)
PC 1	46.5
PC 2	60.9
PC 3	71.6
PC 4	79.6
PC 5	86.9
PC 6	92.5

2.4. Hierarchical clustering

After the input variables are obtained from both feature selection methods, the next step is the clustering process. This research will use a hierarchical method for the clustering process. Hierarchical clustering is an unsupervised machine learning that can assemble a group of data into some clusters with high similarity without the necessity of labeled data [19]. Euclidian distances were used to rate the similarity between the data in this research.

The clustering process has been used to detect a failure in a boiler in some research. These researches showed that the clustering process could detect a failure and even warned the failure before it happened [20], [21]. So, the clustering process has the potential to detect anomalies on a boiler, which is followed by leakage failure. Moreover, the more definitive the input variables (how the variables and related failure are interconnected), the more possible the user can interpret the component's condition [11].

During the clustering process, the number of clusters must be defined. Research, which was trying to evaluate the performance of an exhaust fan by the vibration, showed that the effective number of clusters was three, which indicates normal, warning, and failure [7]. Even though the three clusters rule seems convincing, the data must be tested by making a dendrogram to determine the number of clusters [11]. A dendrogram is a graph that represents the distance of each data and shows how close they are by their distances.

2.6. Naïve Bayes classification

Naïve Bayes classification is a classification method based on the Bayes probability theorem. This classification method has a wide range of applications, such as detecting brain tumors [22], predicting purchase [23], text classification [24], and so on. This procedure has also been mentioned as a method to prognosis in industrial cases [2]. Label data can also utilize naïve Bayes classification to infer the category of the new data [25]. In this research, naïve Bayes classification is combined with hierarchical clustering. Hierarchical clustering will transform the unlabeled data into labeled data to distinguish the normal data from the anomaly. Subsequently, the naïve Bayes classification will calculate the probability of the data being categorized as normal or anomaly based on the labeled data from the clustering process. The probability calculation of data is considered normal can be seen in (1)-(3).

$$P\left(N\middle|Y_{1,j}, Y_{2,j}, \ldots, Y_{n,j}\right) = \left(\prod_{i=1}^{n} P(Y_{i,j}|N)\right). P(N) \tag{1}$$

Where,

$$P(N) = \frac{S_{norm}}{S} \tag{2}$$

$$P\left(Y_{i,j}\middle|N\right) = \frac{e^{-\frac{(Y_{i,j}-\overline{X_{iN}})^2}{2\sigma_{iN}{}^2}}}{\sigma_{iN}\sqrt{2\pi}} \; ; i = 1,2,3,\ldots,n \tag{3}$$

$P\left(N\middle|Y_{1,j}, Y_{2,j}, \ldots, Y_{n,j}\right)$ is the probability of normal condition with the appearance of jth row in testing data Y, which contains n input variables. S_{norm} is the amount of normal sample, S is the total sample, $\overline{X_{iN}}$ is the mean of the ith variable in the normal cluster of training data X, and σ_{iN} is the standard deviation of the ith variable in the normal cluster of training data X.

The next step is setting minimum probability for the data determined as a normal condition. (4) describes how to obtain the minimum normal probability P_{min}. P_{min} is the minimum probability of anomaly training data (X') considered a normal cluster. If the value of $P\left(N\middle|Y_{1,j}, Y_{2,j}, \ldots, Y_{n,j}\right) > P_{min}$, then the data is determined as normal. Conversely, if $P\left(N\middle|Y_{1,j}, Y_{2,j}, \ldots, Y_{n,j}\right) < P_{min}$, then the data is determined as an anomaly.

$$P_{min} = f_{max}\left(P\left(N\middle|X'_{1,j}, X'_{2,j}, \ldots, X'_{n,j}\right)\right) \tag{4}$$

3. RESULTS AND DISCUSSION

This section presents the training data, then the testing results from the feature selection methods, the structural analysis, and PCA. It compares the results regarding the interpretability, false alarms occurrence, and longevity of the prediction. The results were quite difference and able to expose the strength and weakness of each method.

3.1. Data

Figure 4(a) shows the plot of training data used in structural analysis: PSH inlet temperature, PSH outlet temperature, and SSH inlet temperature. Furthermore, Figure 4(b), for principal component 1-3, and Figure 4(c), for principal component 4-6, shows the plot of standardized data of the 12 variables in Table 1, which was done by PCA and then compressed into six principal components. All data in Figure 4(a), Figure 4(b) and Figure 4(c) were taken in the same period in same boiler unit.

Figure 5 plots the testing data used in this research. Figure 5(a) shows the plot of testing data with the same variables in Figure 4(a). Figure 5(b), for principal component 1-3, and Figure 5(c), for principal component 4-6, uses the same variables and is also standardized by PCA as in Figure 4(b) and Figure 4(c). Both variables in Figures 4 and 5 were taken in the same boiler unit but on different occasions. Testing data contains a significant leakage failure that occurred in the 911th sampling, which is marked by a big dot.
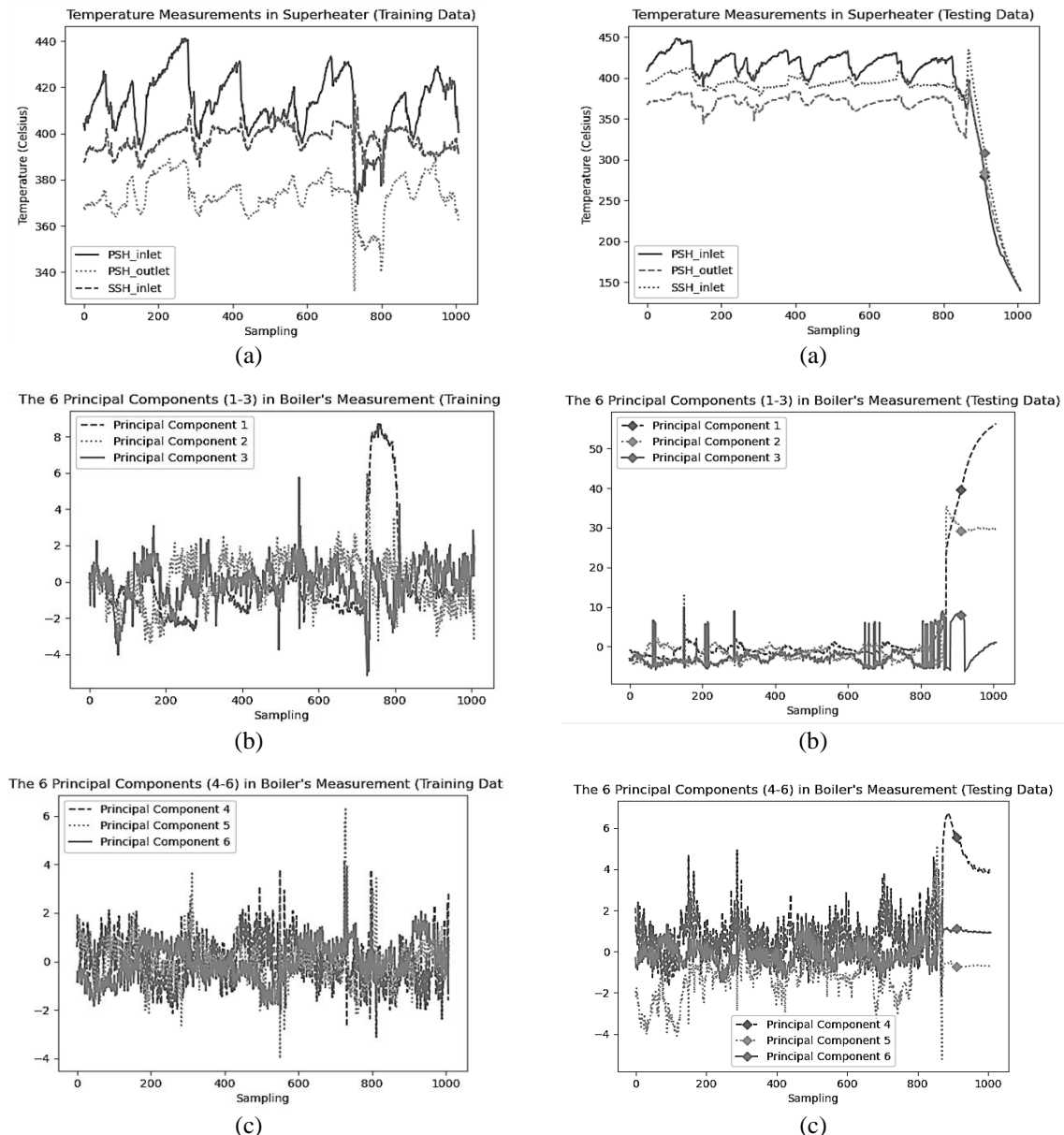

(a)

(b)

(c)

Figure 4. Data plotting of (a)training data using actual measurement for structural analysis (3 variables); (b) training standardized data using PCA (6 principal components) from principal component 1-3; and (c) training standardized data using PCA (6 principal components) from principal component 4-6

(a)

(b)

(c)

Figure 5. Data plotting of (a)testing data using actual measurement for structural analysis (3 variables); (b) testing standardized data using PCA (6 principal components) from principal component 1-3; and (c) testing standardized data using PCA (6 principal components) from principal component 4-6. The major leakage failure is marked by dots in each variable (911th sample)

3.2. Interpretability

Interpretable machine learning has become an exciting topic these days. It gives valuable information and helps the user in the decision-making process. However, to understand the result from machine learning, the users should understand the basic concept of the related object and compare the observation from the machine learning to the real to get the real insight of the event. Figure 6 shows the comparison of dendrogram in the training data, using structural analysis and PCA. Figure 6(a) is a dendrogram constructed by structural analysis. The data was divided into three optimal clusters, which can be observed deeply using a scatter plot.

On the other hand, Figure 6(b) was constructed by PCA with six principal components, which differed the data into two clusters. The differences occur because of different selected variables in each method. Structural analysis intensely focused on the superheater components variables, which is considered the leading cause of the leakage. At the same time, PCA took broader aspects of the boiler, which generalized its observation, resulting in different diagnoses.

The comparison of scatter plots using the training dataset between structural analysis and PCA can be seen in Figure 7(a) and Figure 7(b). Figure 7(a) shows the clustering result of structural analysis feature selection, while Figure 7(b) shows the clustering result of PCA feature selection. Feature selection method using structural analysis offered simplicity to explain the character of the possible anomaly. From the structural analysis perspective, the anomaly, followed by leakage failure, showed unstable measurement (higher or lower than the usual). The abnormality of the heat transfer process caused instability during the combustion. Slugging in the pipeline could cause the cause of higher steam temperature. In comparison, the lower temperature could be caused by the defect of the pipeline material (such as corrosion, deteriorating material, and so on) that kept the pipeline absorbing the heat without transferring it to the steam until the pipeline ruptured. Even though PCA covered more variables in the process, PCA didn't show any sign of interpretability because the principal components did not have any physical meaning or explanation
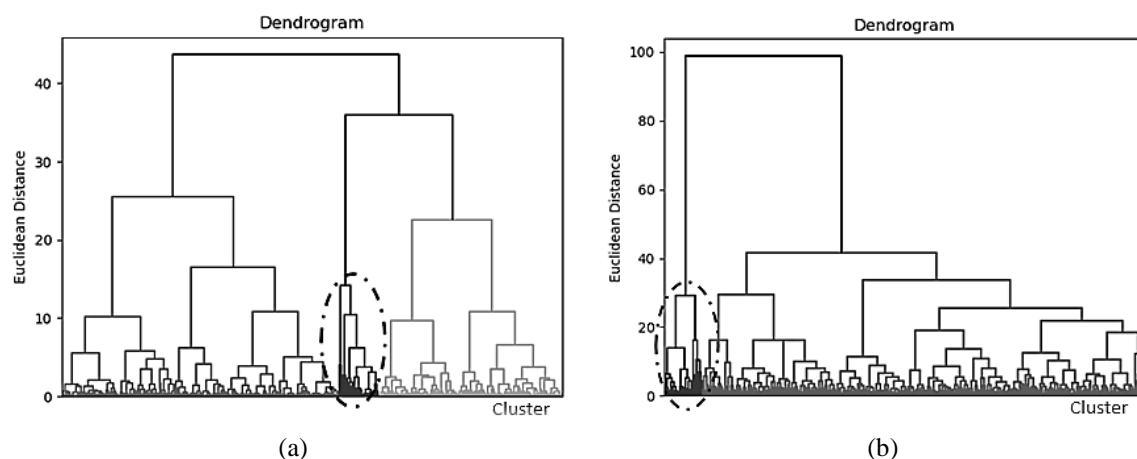

(a)   (b)

Figure 6. Comparison of dendrogram in the training data using (a) structural analysis, and (b) PCA (6 principal components). The anomaly cluster is marked by black circle; the normal cluster is marked by grey or light grey

3.3. False Alarm Occurrence and Longevity of the Prediction

Both methods showed significant differences in the false alarm and length of prediction categories, as shown in Figure 8. A blue dot in the figure represented the event when the actual leakage was confirmed, and the red line indicates how long the prediction was before the leakage began. The anomaly state was confirmed when the status value on that time equals 1. Alternatively, it is normal when the status value is 0.

Feature selection with PCA improved the longevity of the prediction by up to 25 hours. However, the alarm kept being triggered for the whole week because it detected too many anomalies in other components even though they were not significantly related to the leakage event. On the other hand, feature selection with structural analysis could only predict 13 hours and 40 minutes before the leakage. However, the false alarm rate was so low that only five false alarms were activated during the testing. The low false alarm rate in structural analysis happened because it only focused on the superheater area where most of the leakage failure occurred.
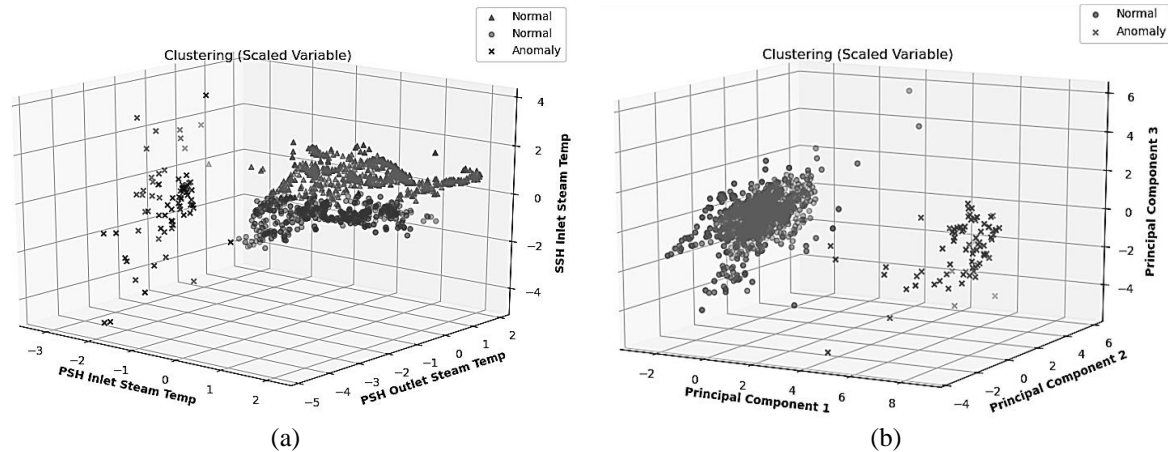
Figure 7. Scatter plot of training data using (a) structural analysis (3 variables), and (b) PCA (6 principal components, only 3 principal components are shown). The anomaly cluster is marked by "x", the normal cluster marked by "o"
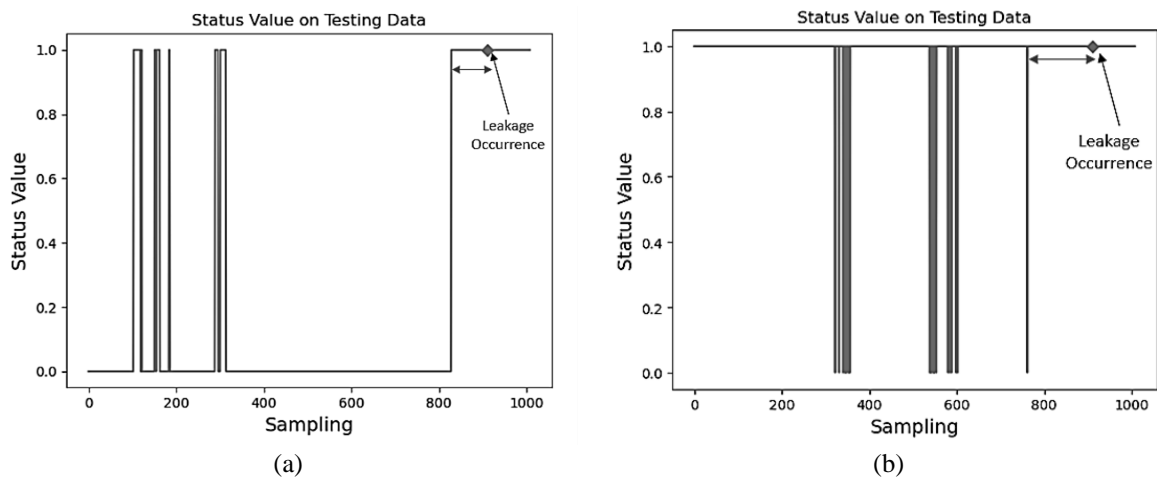

Figure 8. The status value of anomaly detection on testing data using (a) structural analysis and (b) PCA

4.   CONCLUSION

The implementation of different feature selection methods brought out remarkable results. For example, PCA serves an extended prediction period of up to 25 hours with no requirement to understand the process cycle in the boiler. Still, the alarm was triggered during the whole week, and the observation was not interpretable. PCA yields these problems because it indirectly monitors more variables. As a result, it indicates the abnormality in different sections as the suspected anomaly, despite not being related to the failure. Conversely, the structural analysis could predict the leakage 13 hours 40 minutes (much later than PCA) before the occurrence. However, the result from the clustering process was possible to be interpreted. Furthermore, steam temperature variation in the superheater represented the anomaly in the heat transfer between the pipeline and steam, which led to the leakage. Therefore, the false alarm rate by structural analysis was lower than the PCA. Hopefully, the combination of PCA and structural analysis can be developed to predict leakage in future research.

REFERENCES

[1] J. Yan, Y. Meng, L. Lu, and C. Guo, "Big-data-driven based intelligent prognostics scheme in industry 4.0 environment," in *2017 Prognostics and System Health Management Conference (PHM-Harbin)*, Jul. 2017, no. 2, pp. 1–5, doi: 10.1109/PHM.2017.8079310.

[2] Z. Zhou, D. Liu, and X. Shi, "Research on combination of data-driven and probability-based prognostics techniques for equipments," in *Proceedings of 2014 Prognostics and System Health Management Conference, PHM 2014*, 2014, pp. 323–326, doi: 10.1109/PHM.2014.6988187.

[3] N. A. Harish, R. Ismail, and T. Ahmad, "State space modeling of a reheater system in power generation plant," in *2013 IEEE Symposium on Computers & Informatics (ISCI)*, Apr. 2013, pp. 104–107, doi: 10.1109/ISCI.2013.6612384.

[4] P. Madejski and P. Żymełka, "Calculation methods of steam boiler operation factors under varying operating conditions with the use of computational thermodynamic modeling," *Energy*, vol. 197, p. 117221, Apr. 2020, doi: 10.1016/j.energy.2020.117221.

[5] S. Xue *et al.*, "Analysis of the causes of leakages and preventive strategies of boiler water-wall tubes in a thermal power plant," *Engineering Failure Analysis*, vol. 110, p. 104381, Mar. 2020, doi: 10.1016/j.engfailanal.2020.104381.

[6] D. Egeonu, C. Oluah, and P. Okolo, "Thermodynamic optimization of steam boiler parameter using genetic algorithm," *Innovative Systems Design and Engineering*, vol. 6, no. December, pp. 53–73, 2015.

[7] N. Amruthnath and T. Gupta, "A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance," in *2018 5th International Conference on Industrial Engineering and Applications (ICIEA)*, Apr. 2018, pp. 355–361, doi: 10.1109/IEA.2018.8387124.

[8] J. D. Kelleher, B. Mac Namee, and A. D'Arcy, *Fundamentals of machine learning for predictive data analytics*. Massachusetts: The MIT Press, 2015.

[9] N. R. Prasad, S. Almanza-Garcia, and T. T. Lu, "Anomaly detection," *CMC-Computers, Materials & Continua*, vol. 14, no. 1, pp. 1–22, 2009, doi: 10.3970/cmc.2009.014.001.

[10] E. Vanem and A. Brandsater, "Cluster-based anomaly detection in condition monitoring of a marine engine system," in *2018 Prognostics and System Health Management Conference (PHM-Chongqing)*, Oct. 2018, pp. 20–31, doi: 10.1109/PHM-Chongqing.2018.00011.

[11] B. Efendy, E. M. Budi, E. Ekawati, M. Soleh, and H. Nugraha, "Leakage prediction on superheater in boiler with hierarchical clustering and naïve Bayes classification," 2021.

[12] E. L. Russell, L. H. Chiang, and R. D. Braatz, "Data-driven methods for fault detection and diagnosis in chemical processes," *Journal of Chemical Information and Modeling*, vol. 53, no. 9, pp. 1689–1699, 2000, [Online]. Available: http://link.springer.com/10.1007/978-1-4471-0409-4.

[13] L. Jiao-Jiao, X. Liang, W. Ke-Yang, and C. Zhi, "Cause analysis of leakage of supercritical (ultra-supercritical) boiler water wall slag screen pipe," in *Proceedings - 11th International Conference on Intelligent Computation Technology and Automation, ICICTA 2018*, 2018, pp. 388–390, doi: 10.1109/ICICTA.2018.00094.

[14] K. Sun, J. Wang, Y. Yan, J. Jiang, L. Deng, and D. Che, "Analysis of frequent tube burst of high temperature reheater in 600 MW unit," in *iSPEC 2020 - Proceedings: IEEE Sustainable Power and Energy Conference: Energy Transition and Energy Internet*, Nov. 2020, pp. 2429–2434, doi: 10.1109/iSPEC50848.2020.9351120.

[15] M. Z. Sheriff, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, "Fault detection using multiscale PCA-based moving window GLRT," *Journal of Process Control*, vol. 54, pp. 47–64, 2017, doi: 10.1016/j.jprocont.2017.03.004.

[16] Y. J. Cui, L. W. Xia, Y. Huang, and X. C. Ma, "Research on fault diagnosis and early warning of power plant boiler reheater temperature deviation based on machine learning algorithm," in *2020 IEEE 6th International Conference on Control Science and Systems Engineering, ICCSSE 2020*, 2020, pp. 212–216, doi: 10.1109/ICCSSE50399.2020.9171950.

[17] L. Liu, J. Liu, X. Yu, H. Wang, and Z. Chen, "A multivariate monitoring method based on PCA and dual control chart," in *Proceedings of the 31st Chinese Control and Decision Conference, CCDC 2019*, 2019, pp. 5914–5918, doi: 10.1109/CCDC.2019.8832679.

[18] W. Wenbiao, L. Chuanjin, Z. Jie, and W. Siyuan, "Research of optimal operation method based on cross and piecewise PCA for industrial boilers," in *Proceedings of the 30th Chinese Control and Decision Conference, CCDC 2018*, 2018, pp. 819–824, doi: 10.1109/CCDC.2018.8407243.

[19] Nisha and P. J. Kaur, "Cluster quality based performance evaluation of hierarchical clustering method," in *Proceedings on 2015 1st International Conference on Next Generation Computing Technologies, NGCT 2015*, 2016, vol. September, pp. 649–653, doi: 10.1109/NGCT.2015.7375201.

[20] K. H. Kim, H. Seok Lee, J. H. Kim, and J. Ho Park, "Detection of boiler tube leakage fault in a thermal power plant using machine learning based data mining technique," in *Proceedings of the IEEE International Conference on Industrial Technology*, 2019, vol. 2019-Febru, pp. 1006–1010, doi: 10.1109/ICIT.2019.8755058.

[21] J. Yu, J. Jang, J. Yoo, J. H. Park, and S. Kim, "A clustering-based fault detection method for steam boiler tube in thermal power plant," *Journal of Electrical Engineering and Technology*, vol. 11, no. 4, pp. 848–859, 2016, doi: 10.5370/JEET.2016.11.4.848.

[22] H. T. Zaw, N. Maneerat, and K. Y. Win, "Brain tumor detection based on naïve Bayes classification," in *Proceeding-5th International Conference on Engineering, Applied Sciences and Technology, ICEAST 2019*, 2019, pp. 1–4, doi: 10.1109/ICEAST.2019.8802562.

[23] F. Harahap, A. Y. N. Harahap, E. Ekadiansyah, R. N. Sari, R. Adawiyah, and C. B. Harahap, "Implementation of naïve Bayes classification method for predicting purchase," in *2018 6th International Conference on Cyber and IT Service Management (CITSM)*, Aug. 2018, pp. 1–5, doi: 10.1109/CITSM.2018.8674324.

[24] J. Chen, Z. Dai, J. Duan, H. Matzinger, and I. Popescu, "Naive Bayes with correlation factor for text classification problem," in *Proceedings - 18th IEEE International Conference on Machine Learning and Applications, ICMLA 2019*, 2019, pp. 1051–1056, doi: 10.1109/ICMLA.2019.00177.

[25] D. P. He, Z. L. He, and C. Liu, "Recommendation algorithm combining tag data and naive bayes classification," in *Proceedings-2020 3rd International Conference on Electron Device and Mechanical Engineering, ICEDME 2020*, 2020, pp. 662–666, doi: 10.1109/ICEDME50972.2020.00156.

BIOGRAPHIES OF AUTHORS

Eko Mursito Budi 🆔 ⅷ sc Ⓟ is a senior lecturer in the Faculty of Industrial Technology, Institut Teknologi Bandung. He is a multi-disciplinary engineer with various innovations, including virtual control, real-time object-oriented kernel, angklung robot, and batik photonics. He had undergraduate and doctoral degrees in the Engineering Physics Program and a master's degree in Informatics. Moreover, he also holds a professional engineer certificate. He can be contacted at email: mursito@itb.ac.id.

Estiyanti Ekawati 🆔 ⅷ sc Ⓟ is a senior lecturer in the Faculty of Industrial Technology, Institut Teknologi Bandung. She had a bachelor's degree in Engineering Physics, a Master's Degree in Instrumentation and Control, and a Doctor of Philosophy in Engineering. Her main interest is in systems identification and control, data engineering, and data science. She can be contacted at email: esti@itb.ac.id.

Bobby Efendy 🆔 ⅷ sc Ⓟ earned his bachelor's degree in the Engineering Physics Program and master's degree in the Instrumentation and Control Program, both from Institut Teknologi Bandung, Indonesia. He can be contacted at email: bobby.efendy@students.itb.ac.id.