Research Article/Civil and Sanitary Engineering http://doi.org/10.15446/ing.investig.112017

Statistical Study on Domestic Gas Boiler Failures Using Various Software Platforms

Estudio estadístico de fallos en calderas de gas domésticas utilizando diversas plataformas de *software*

Pavel Shcherban [1] and Reda Abu-Khamdi [2]

ABSTRACT

Modern public utilities require a high level of reliability, especially with regard to systems used for heating and hot water supply, such as gas boilers. This paper collected and studied statistical information on the failures of gas boilers of four brands. The data were provided by three service companies that repair gas equipment in Kaliningrad, in the Russian Federation. The specifics of these failures were studied, determining their possible causes, and they were stratified by severity. Service companies typically have to use existing platforms (supplementing them with add-ons) or develop their own software to analyze failure statistics. In this regard, they are interested in the emergence of simple and effective tools for monitoring the quality of gas boiler maintenance and repair work. In this study, we used both the publicly available Scikit-learn library of the Jupyter Notebook environment and a custom program to perform data clustering. The main goal was to conduct a comparative assessment of the reliability of gas boilers of various brands based on the analysis of their failure statistics, as well as to develop a software product that enables such an assessment.

Keywords: gas equipment failures, quality management, equipment reliability, cluster analysis, statistical data processing software

RESUMEN

Las empresas de servicios públicos modernas requieren un alto nivel de fiabilidad, especialmente en lo que respecta a los sistemas utilizados para calefacción y suministro de agua caliente, como las calderas de gas. Este artículo recopiló y estudió información estadística sobre las fallas de calderas de gas de cuatro marcas. Los datos fueron proporcionados por tres empresas de servicio que reparan equipos de gas en Kaliningrado, en la Federación de Rusia. Se analizaron las características específicas de estas fallas, determinando sus posibles causas, y se clasificaron según su gravedad. Las empresas de servicio generalmente deben utilizar plataformas existentes (complementándolas con extensiones) o desarrollar su propio *software* para analizar estadísticas de fallos. En este sentido, están interesadas en el desarrollo de herramientas simples y efectivas para monitorear la calidad del mantenimiento y las reparaciones de calderas de gas. En este estudio, se utilizó tanto la biblioteca de código abierto Scikit-learn del entorno Jupyter Notebook como un programa personalizado para realizar la agrupación de datos. El objetivo principal fue realizar una evaluación comparativa de la fiabilidad de las calderas de gas de diferentes marcas en función del análisis de sus estadísticas de fallos, así como desarrollar un producto de *software* que permita dicha evaluación.

Palabras clave: fallos en equipos de gas, gestión de calidad, fiabilidad de equipos, análisis de conglomerados, software de procesamiento de datos estadísticos

Received: November 5th, 2023
Accepted: October 25th, 2024

Introduction

The reliability of gas equipment is achieved through service and repair work throughout its lifecycle. This is particularly vital for household appliances such as gas boilers. During their operation, various failures may occur due to the violation of installation and usage rules, the quality of gas and water, the stability of the connection to the electrical network, the schedule of diagnostic and service tasks, and the use of non-original spare parts.

In large enterprises of the energy complex, specialized structural divisions address the diagnostics and maintenance of gas boilers. As a rule, they have laboratories and diagnostic resources at their disposal, as well as a significant database of potential failures associated with the equipment in operation. This allows them to maintain the reliability of gas energy equipment at a high level.

Small companies operating in the market for servicing and repairing domestic gas equipment usually have scattered data on the statistics of gas boiler failures. This is due to the large variety of boiler brands and models, the lack of a structured data storage system, staff turnover, and users

[1] Immanuel Kant Baltic Federal University, the Branch scientific cluster, Institute of High Technologies, Kaliningrad, 236009, Russia. Email: ursa-maior@yandex.ru
[2] Moscow Institute of Physics and Technology. Faculty of Applied Mathematics and Informatics, Moscow, 117303, st. Kerchenskaya, 1 A, bldg. 1. Email: rabou-hamdi@gmail.com

turning to various service companies when new failures occur (which makes tracking repair history impossible). One significant problem in this regard is the lack of specialized systems for analyzing the statistics of gas power equipment failures.

In small service companies, this often leads to the irrational planning of spare parts purchases, the impossibility of planning the work volume of repair teams, and difficulties in forecasting maintenance activities for gas boilers of various brands (which should be corrected according to local usage conditions, rather than simply copying the provisions specified by the manufacturer). In general, these problems reduce companies' competitiveness in the technical services market.

A number of studies, such as [1] and [2], have already raised these issues and proposed various ways to solve them. In particular, an idea was put forward to restructure the gas boiler maintenance system, introduce passports for assessing technical conditions and repair work, create a common database of failures, develop mobile applications for contacting service companies, and install analytical units on gas boilers in order to track the characteristics of its operation (with the ability to subsequently extract the data). The proposed measures have their strengths and weaknesses. In some cases, they are organizationally complex, costly, and the promotion of such services in the market can be problematic. In solving these issues, the standardization of software and hardware and the establishment of uniform regulatory requirements are important. These proposals assume greater control over the technical conditions of the gas boilers used in everyday life, and, as a result, they increase the quality of technical maintenance and reliability.

It should be noted that the possible ways to improve companies' quality of service are not limited by the aforementioned measures; an alternative is the creation of a software product intended for the analysis of gas boiler failures, hence the necessity of developing a software package that enables the preparation of statistical data and the comprehensive analysis of the causes of failures.

Thereupon, two goals were set in this study: firstly, to develop a prototype software product for analyzing statistical data on the operation of gas boilers, using the cluster analysis method with stratification by failure severity; and secondly, using this program to form the main statistical groups of reliability for each type of boiler (less reliable, more reliable, and standard). These three equipment groups are generally consistent with the theory of reliability. A number of parameters for said groups (distribution density, and frequency) should indicate greater or lesser reliability in boilers of a particular brand.

To test the software product, statistical data on gas boiler failures, as recorded by three service companies in Kaliningrad in 2022, were used. The study also included a comparison of the results of data processing using our own developed program and the existing scikit-learn library within the Jupyter Notebook interactive computing environment.

This study comprises several sections. The first section is analytical and outlines the collection of data on gas boiler failures as well as a description of their causes. Then, the methodological section presents the principles through which the data are ranked by equipment failure severity using a risk matrix. This section also describes primary data normalization according to the established ranking rule. The software section is based on the outlined methodological apparatus. The resulting array was processed using publicly available software and our own program. Both of these tools were developed to provide a clustering of failures by severity for each gas boiler brand. Finally, the results obtained with various software programs are compared against the data presented in scientific periodicals, and, based on them, the reliability issues of gas boilers are examined, possible reasons for the obtained data are discussed, and directions for further research are defined.

Failures of double-circuit boilers – grouping and pre-processing of statistical data

When thoroughly assessing the technical conditions of equipment, as a rule, one must first determine the prediction strategy and prognostic background and then develop a system of parameters reflecting the nature and structure of the technical object. Then, one must develop normative and search models for the equipment, perform simulations, and evaluate the accuracy of the models (*i.e.*, the rate of equipment failure, the correspondence between the initial and final parameters). The models must be verified using statistics, experimental data, or expert methods. Afterwards, recommendations should be made to optimize decision making with regard to the planning and management of equipment operation based on the obtained models.

However, this method is acceptable either for equipment developers (who provide a warranty for trouble-free operation) or for large service centers responsible for high-tech industrial devices. Small and medium-sized services, which account for the main share of the semi-industrial and domestic equipment repair market, should rather use the most informative parameters that allow assessing technical conditions in a comprehensive and rapid manner, aiming for the operational processing of failure statistics. For the case of double-circuit gas boilers, we selected data on the time until failure of boiler equipment, the efficiency factor of double-circuit boilers, and the severity of the failure that preceded the repair [3]. In general, any set of technical condition parameters can be used for analysis, but the highest informativeness will be obtained if the processed data are complete (*i.e.*, they fully reflect the results of equipment failures), reliable (*i.e.*, they are experimentally verifiable), accurate, holistic, interrelated, measurable, and thus able to reflect the different technical and technological characteristics of the object over time [4].

The data used in this study (mean time until failure, efficiency, and severity of the failure that preceded the repair) were processed and prepared step by step. The sample was formed from four groups of double-circuit gas boilers (Ariston, Buderus, Bosch, and Wisseman), with 25 boilers in each group, which were examined after failure and repair. The first parameter considered was the MTBF, a technical parameter that characterizes reliability of a repaired device, equipment, or technical system [5]. The reliability function can be defined using non-failure operation probability, whose formula is as follows:

$$P(t) = exp^{\left(-\frac{t}{T_o}\right)} \quad (1)$$

where:

- t – operating time of the machine
- T_o – mean time until failure

The MTBF of a piece of equipment can also be estimated through the failure rate formula [6]. The failure rate (2) is a parameter that determines the reliability of a piece of equipment:

$$\lambda = \frac{1}{MTBF} \quad (2)$$

where:

- λ – failure rate
- MTBF – mean time between failures

However, in general, the MTBF is determined statistically over time (3):

$$T = \frac{\sum_1^m t_i}{m} \quad (3)$$

where:

- t_i – mean time until failure *i*
- m – number of failures

It should be noted that this study deals with a MTBF measured from the moment of a machine's first start-up to the first breakdown before repair [7]. Thus, this parameter indicates the non-failure operation time of the new equipment purchased and used by the consumer. This information is extremely important for the consumer to plan their expenses, as well as for the service organization to schedule incoming orders depending on the previous sales volume of certain equipment [8].

We used the efficiency of boilers before failure as the second parameter. In this case, this is an effective analytical parameter because, on the one hand, from the manufacturer and the service company's perspective, it indicates the technical system's dynamics of degradation over time. On the other hand, from the consumer's perspective, it shows the efficiency of the equipment's operation before failure. Boiler efficiency mainly corresponds to the ratio of fuel consumed to heat emitted [9]. This parameter can be calculated through several methods, the first of which is direct:

$$\eta_k = \frac{Q_k}{Q_H^p} \quad (4)$$

where:

η_k – boiler efficiency

Q_k – useful energy transferred to the coolant

Q_H^p – thermal energy released as a result of the chemical reaction of combustion

There is also the reverse method [10]:

$$Q_H^p = Q_1 + Q_2 + Q_3 + Q_4 + Q_5 \quad (5)$$

where:

- Q_H^p – thermal energy released as a result of the combustion reaction
- $Q_1 = D \cdot (h_s - h_{fw}) / B$ – heat used for steam generation
  - D – boiler steam capacity (kg/s)
  - B – fuel flow rate per second (kg/s or m3/s)
  - h_s and h_{fw} – enthalpy of the steam and feed water (kJ/kg)
- Q_2 heat loss caused by the exhaust gases in the boiler unit
- Q_3 – chemical heat loss (underburning) caused by the incomplete combustion of the fuel
- Q_4 – mechanical heat loss (underburning) caused by incomplete combustion
- Q_5 – heat loss to the environment through the boiler's external enclosures

The third parameter analyzed was failure severity, based on the complexity of repair work and the frequency of occurrence. In general, failure severity can also be regarded

as the total criticality of the *j*-th failure of the *i*-th piece of equipment [11]. It can be calculated as:

$$\sum Cm_{ij}^{k} = \beta_{ij}^{k} \cdot \alpha_{ij} \cdot \lambda_{i} \cdot (T_{work})_{i} \quad (6)$$

where:

- Cm_{ij}^{k} – criticality of the *j*-th type of failure of the *i*-th element
- β_{ij}^{k} – the probability of consequences of a certain category of severity for the j-th type of failure of the i-th element
- α_{ij} – share of the *j*-th type of failure of the *i*-th element
- λ_{i} – failure rate of the *i*-th element
- $(T_{work})_{i}$ – mean time until failure of the *i*-th element

Visually, this parameter can also be represented by a risk matrix (Fig. 1).

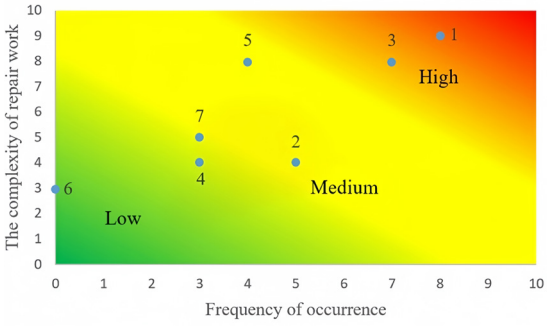

Figure 1. Severity of boiler equipment failure
Source: Authors

This figure includes the following items:

1. Corrosion of gas burners
2. Removal of fouling products
3. Switch failure
4. Breakage of the temperature sensor
5. Gas combustion noise
6. Inoperative burners
7. Inoperative traction sensor

We prepared the initial data on all four types of double-circuit gas boilers according to all the given parameters. Moreover, we identified the relationship between these parameters (which was obvious from a technical and technological perspective), but we did not know its extent or how evident it would be in the groups [12]. In general analytics, it is better to employ cluster analysis, noting that clustering can be performed in different ways. However, before clustering, it was necessary to normalize the initial data, given that, in their original form, which exhibited a high degree of numerical variation, the results were more difficult to formalize and visualize.

Data normalization for subsequent cluster analysis

Normalization is the process of converting data into a single, dimensionless range of values. Depending on the purpose of this process, there are various functions that match the original data of the row with their normalized value. One of these is the scaling function, which maps a value from the range [min(X),max(X)] (where $X = x_1, x_2, \ldots, x_n$, representing a row of experimental data) to a value from a given range $[a,b]$. If $[a,b]=[0,1]$, there are several options for the scaling function, *e.g.*, min-max normalization, z-estimation, and scaling relative to a unit vector. Min-max normalization, the simplest of the scaling functions, has the following form: $f(x_i)=\frac{x_i-\min(X)}{\max(X)-\min(X)}$. Z-estimation allows transforming a data row such that its mean value becomes 0 and its variance becomes 1. Scaling relative to a unit vector implies representing a row of *n* data as an *n*-dimensional vector, after which the formula for normalizing the vector $(X'=\frac{X}{|X|})$ is applied based on the selected norm function.

In this study, the purpose of normalization was to bring three-dimensional data, which had various units of measurement and differed in absolute value by orders of magnitude, to relative dimensionless values. To this effect, the simplest and most effective method was a scaling function, *i.e.*, min-max scaling [13]. Accordingly, each component of the three-dimensional data $A=\{(x_i, y_i, z_i),\ i=1..n\}$ was represented as a separate data row ($X=\{x_i\}$, $Y=\{y_i\}$, $Z=\{z_i\}$), applying min-max normalization to each of them. As a result, the normalized data rows $X'=\{f(x_i)\}$, $Y'=\{f(y_i)\}$, and $Z'=\{f(z_i)\}$ were obtained. Afterwards, these rows were gathered into a single set, restoring the original three-dimensional row while including the normalized values $A'=\{(f(x_i), f(y_i), f(z_i)),\ i=1..n\}$. The dimensionlessness achieved was a result of data normalization using the min-max scaling formula and the division of two values with the same dimensions. This facilitated the analysis, as well as the graphical representation of the data's dependence on the values of other rows, especially when the values differed by orders of magnitude in the original dimensions. This was necessary because the dimensionality of the analyzed data complicated their graphical representation. In general, this procedure can be applied to any kind of data used in cluster analysis.

Comparison of cluster analysis based on the Scikit-learn library with our own custom implementation

There are different types of cluster analysis: the probabilistic approach, which deals with the relationship between each studied object and one of the cluster classes; the logical approach, which builds dendrograms using decision trees; and the hierarchical approach, which requires different types of clusters combined into a single one. In this case, we applied the K-means method [14], as it allows distributing

events into groups both qualitatively and quantitatively. The idea of the method is to reduce the cluster points' sum of quadratic deviations from the clustering centers [15].

As a rule, K-means method involves setting an initial number of clusters [16]. In the case of this study, it can be logically assumed that three main groups of clusters will be formed:

- *Unreliable devices*. These had a high failure rate or exhibited low efficiency and MTBF values, *i.e.*, the equipment failed before its warranty period ended or underperformed before failing.
- *Devices with normative reliability*. For these devices, all the parameters related to failure or performance loss behaved as expected within the warranty period, which, as a rule, entailed no severe consequences.
- *Devices exhibiting a high degree of reliability*. These devices showed low failure severity while remaining in operation longer than specified in the warranty and without significant loss of performance.

For a more accurate analysis, we used two approaches for clustering: the publicly available Scikit-learn package of the Jupyter Notebook environment and our own custom program. We then compared the accuracy of clustering on the same raw data, which were previously normalized. Jupyter Notebook is a development environment that immediately displays the results when executing code [17]. The advantage of this environment is that the code can be segmented and executed in any order. The Scikit-learn library is a Python library that includes clustering and regression analysis methods.

In our custom program, also written in Phyton, the initial parameters of the clustering process can be set more accurately. The steps of this process are presented in Fig. 2 in the form of flowcharts for both methods.

The first flowchart describes the clustering algorithm of the Scikit-learn library. After initialization, the data are entered into the first block, after which they are delivered to the *k_means* function. This function performs clustering according to the K-Means algorithm and writes the result

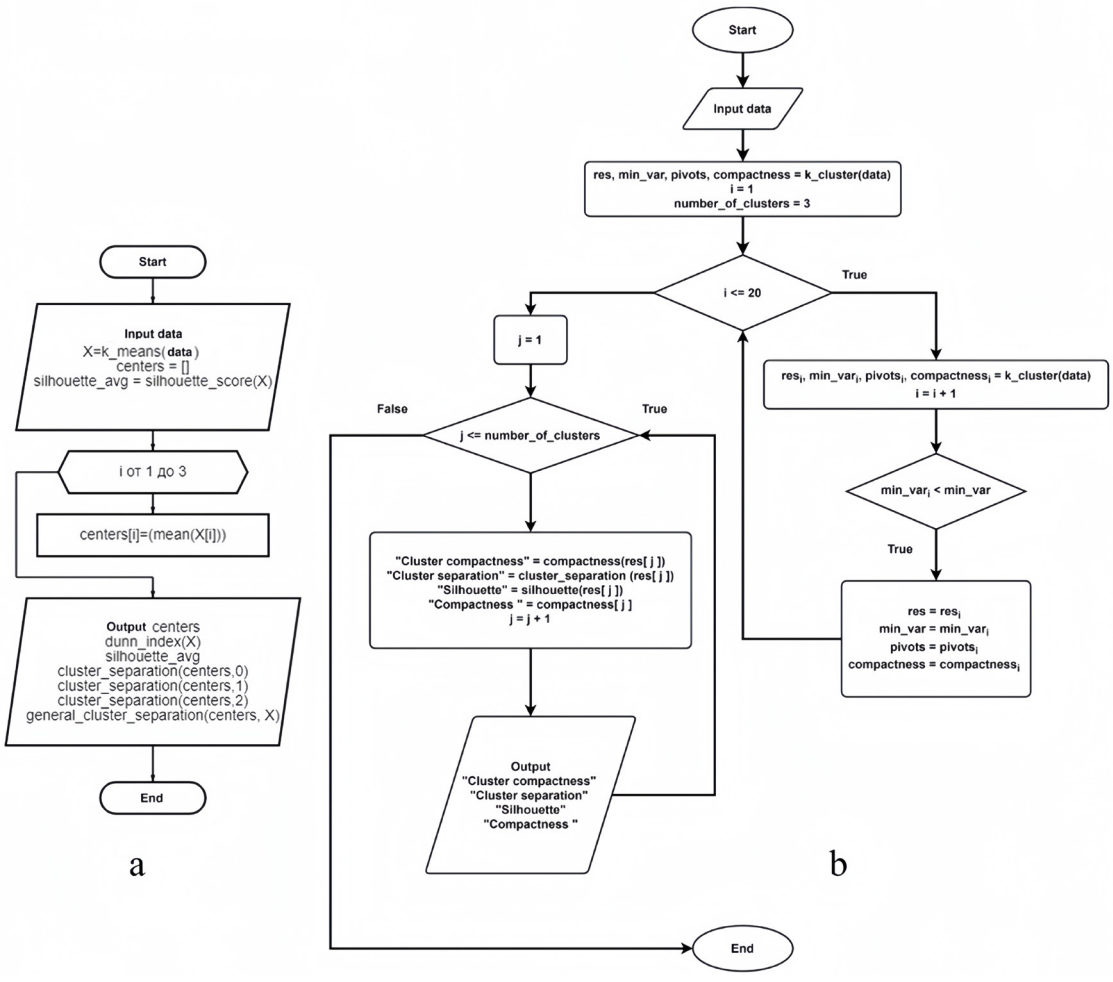

Figure 2. Clustering using the k-means method via a) the Scikit-learn library of the Jupyter Notebook environment and b) our custom program for cluster analysis
Source: Authors

into the X variable, after which an empty array of centers is created, with the purpose of storing the coordinates of the cluster centers. The *silhouette_score* function calculates the average clustering silhouette and writes the result into the *silhouette_avg* variable. Next, within the loop, the centers are calculated as the average point in the cluster using the mean function, after which they are recorded in the array. In the last block, before the end, the results are obtained, namely the cluster centers, the Dunn index, the silhouette, the separability of each individual cluster, and the total separability.

The second flowchart describes the algorithm of our program. The first block after initialization is data entry. The clustering algorithm is executed using the *k_means* function, which accepts the data as input and returns the following values:

- res_i : an array containing the three clusters
- $\min\_var_i$: the sum of all clusters' compactness scores
- $pivots_i$: an array containing the central point of each cluster
- $compactness_i$: an array containing the compactness score for each cluster

The process was repeated 20 times, selecting the clustering with the lowest sum of quadratic deviations. Afterwards, the algorithm estimated the following parameters for each cluster: compactness, separability, silhouette, and the Dunn index. Based on the two presented algorithms for the four groups of normalized data, we conducted a

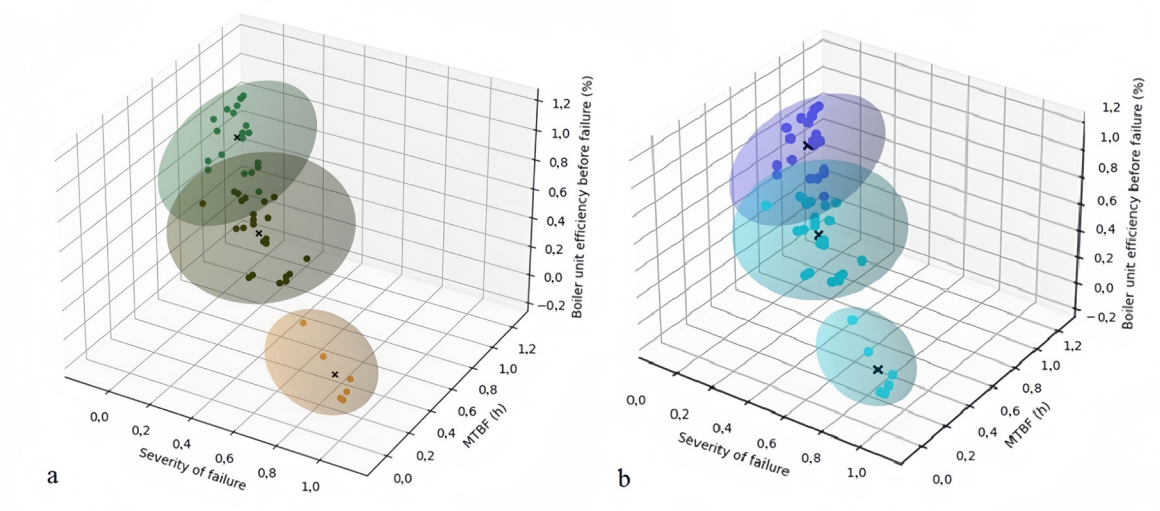

Figure 3. Clustering of Ariston boiler failure data, as obtained a) using the Scikit-learn library in Jupyter Notebook and b) using our own program
Source: Authors

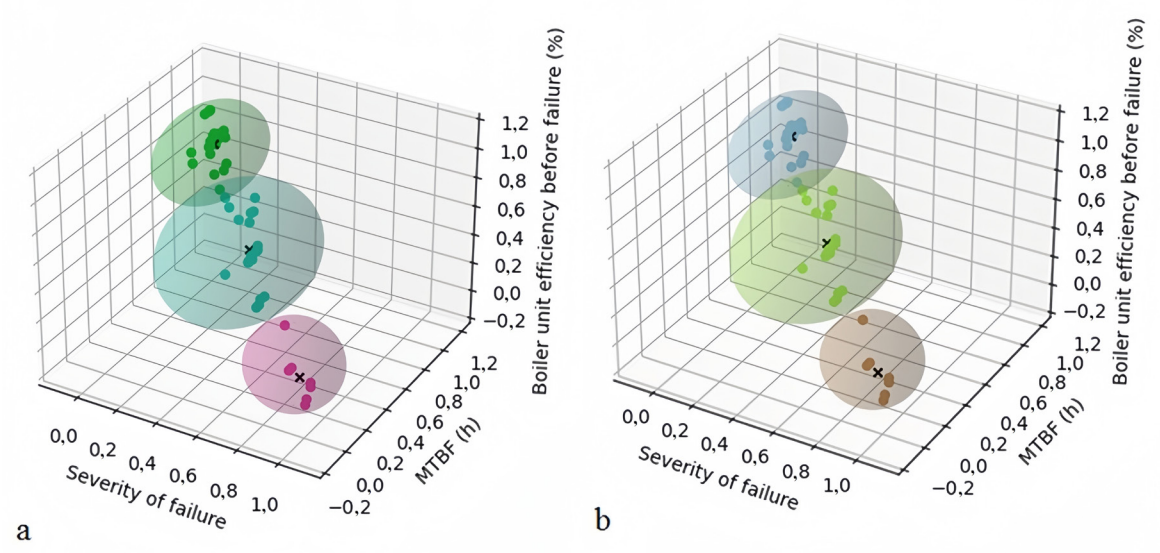

Figure 4. Clustering of Bosch boiler failure data, as obtained a) using the Scikit-learn library in Jupyter Notebook and b) using our own program
Source: Authors

cluster analysis. The results of this analysis are presented in the form of graphs (Figs. 3 to 6). In this case, the visual differences are minimal. Each cluster center is fixed, and the data group is clearly delineated. By comparing the results obtained, the visual differences between algorithms are not as clear as those between the sizes of the clusters for the boiler brands. We initially assumed three main groups of clusters: unreliable devices, devices with normative reliability, and devices exhibiting a high degree of reliability. In the figures, the first category is closest to the reader, the second is in the middle of the clustering cube, and the third is in the upper left corner. In the case of Ariston boilers, the largest clusters correspond to devices exhibiting high and normative reliability (Fig. 3). As for Bosch boilers (Fig. 4), the largest cluster is in the normative reliability zone, and, for the Buderus and Viessmann boilers (Figs. 5 and 6), the largest clusters are in the low reliability zone.

Even a visual assessment of the results of the cluster analysis provides a lot for technicians and managers. In our case, we identified the need to study the causes of the increased failure frequency and severity of Buderus and Viessmann boilers. Measures should also be taken to shift the failure distribution statistics from the zone of low reliability towards those of high and normative reliability. However, a visual assessment was obviously not enough to determine the effectiveness of data clustering by means of the two studied approaches. In order to determine exactly what method is more accurate and efficient in processing the data, several key clustering parameters had to be calculated. These structurally indicated the density of the clusters obtained, their separability, and the closeness of the links between the grouped events [18], making it possible to compare the results obtained for each of the four failure datasets and for both approaches.

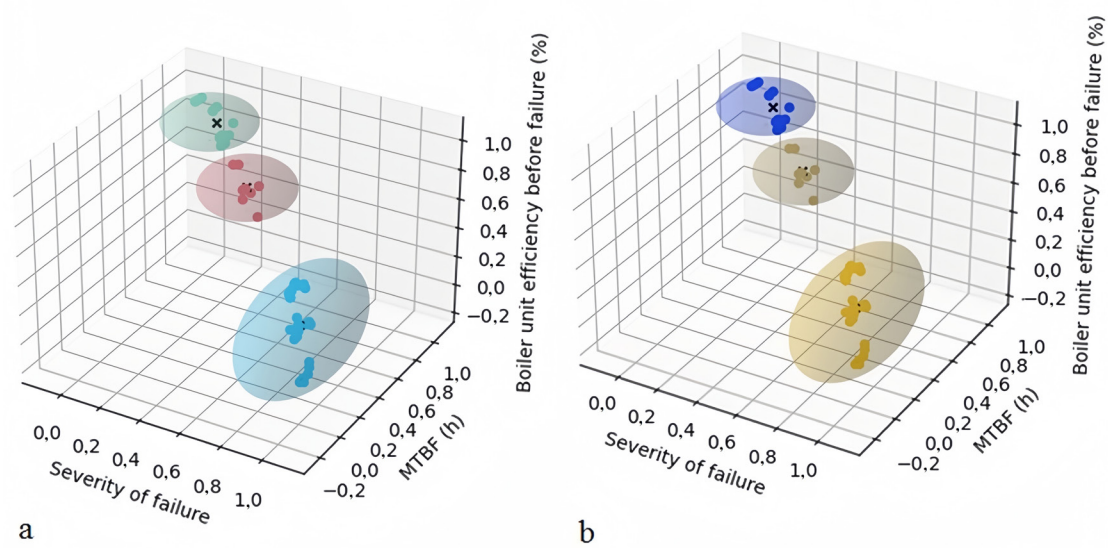

Figure 5. Clustering of Buderus boiler failure data, as obtained a) using the Scikit-learn library in Jupyter Notebook and b) using our own program
Source: Authors

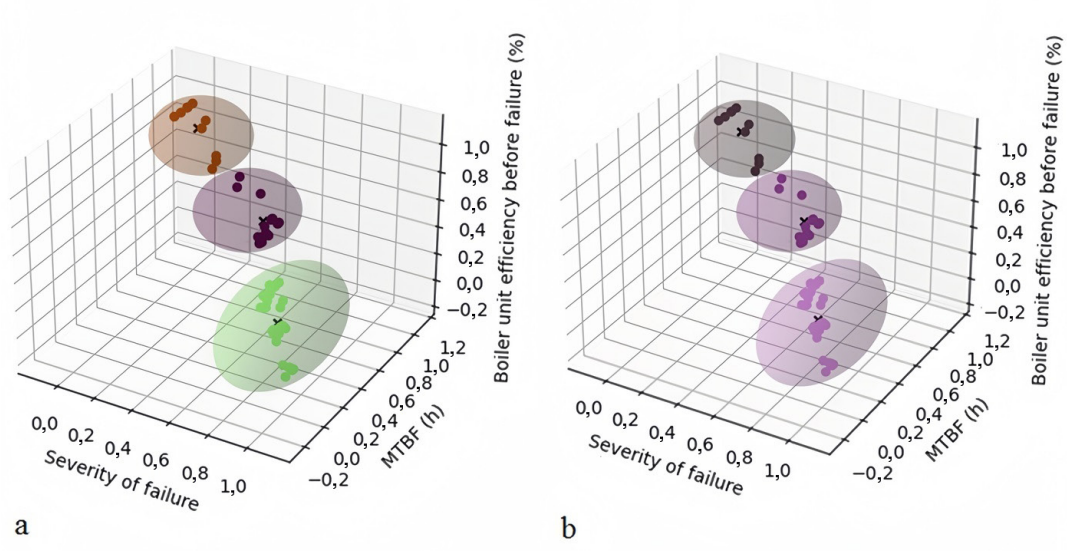

Figure 6. Clustering of Viessmann boiler failure data, as obtained a) using the Scikit-learn library in Jupyter Notebook and b) using our own program
Source: Authors

To make comparisons, we used the Dunn index, a metric for evaluating clustering algorithms. In these contexts, the goal is to identify sets of clusters that are compact, exhibit little variance between elements, and are well separated. The mean values of the clusters must be distant from each other when compared to the variance within them. For a given cluster distribution, a higher Dunn index indicates better clustering [19]. This can be calculated in various ways, such as

$$D(C) = \frac{\min_{Ck\in} C\left\{\min_{Cl\in C\setminus Ck}\left\{\delta(Ck,Cl)\right\}\right\}}{\max_{Ck\in C}\left\{\Delta(Ck)\right\}} \quad (7)$$

where:

- Ck, Cl – clusters from set C
- C – set of clusters
- δ – inter-cluster distance
- Δ – cluster diameter

The Euclidean norm was used as a distance function in this and all subsequent estimations. We considered the distance between clusters to be the minimum distance between their points.

We also used the silhouette index, which indicates how similar an object is to its cluster when compared to other clusters. The silhouette index can be calculated as follows:

$$il(C) = \frac{1}{N}\sum_{Ck\in C}\sum_{x_i\in C_k}\frac{b(x_i,c_k)-a(x_i,c_k)}{max\left\{a(x_i,c_k),b(x_i,c_k)\right\}} \quad (8)$$

where:

- x_i – element of the cluster
- a, b – average distance of point x_i
- N – number of elements in the set C

Cluster separability is evidently an important parameter, as it indicates the degree of overlap between clusters and the distance from each other in space. In general, it confirms the correctness of the initial hypothesis regarding the calculated number of clusters according to their remoteness from each other. Separability is determined using the following formula:

$$BSS = n\cdot\sum_{j=1}^{M}\left(\overline{x_j}-\overline{x}\right)^2 \quad (9)$$

where:

- n – a cluster number
- M – a number of clusters
- j – an index

Another important parameter is cluster compactness, which evaluates the distance between cluster elements. This property can be expressed through the distance between the elements in the cluster, the cluster's density, or the volume occupied by the cluster in the multidimensional space. Compactness (10) is calculated as follows:

$$WSS = \sum_{j=1}^{M}\sum_{i=1}^{|C_j|}\left(x_{ij}-\overline{x}_j\right)^2 \quad (10)$$

The obtained data on cluster centers, silhouettes, cluster separability, compactness, and the Dunn index are summarized in Table I. The results presented therein show that the use of our own program allowed for a more accurate organization of the data in clusters. This is especially important for large sample sizes. Thus, in most cases, the results of statistically processing all considered parameters are higher by the hundredths. This indicates a clearer clustering, which has to do with the specific characteristics of Jupyter Notebook [20]. In particular, the Scikit-learn library was originally written as a universal tool for mathematical data analysis, which is why, in applying individual information processing methods, it can yield slightly larger deviations [21].

It can be stated that, for the analysis of small samples with an accuracy of one hundredth of a unit, it is possible to employ clustering via Scikit-learn. In other cases, to study large amounts of statistical data with the purpose of assessing the quality of maintenance and repair, it is better to create custom clustering software.

Discussion

Regardless of the clustering program used, an analysis of the results (Figs. 3 to 6 and Table I) allows drawing several important technical conclusions. This section discusses them and compares them against facts previously presented in scientific periodicals. The shift in the density function of the failures is most visible in gas boilers of the Buderus and Viessmann brands (within the data sample, these boilers failed more often, exhibiting sharp drops in efficiency and total equipment failure).

It is clear that the reliability indicator is higher for Bosch and especially Ariston boilers. This is evidenced in the smaller size of the lower right cluster of events in each of the graphs. This cluster shows the number of boilers of a certain brand from the sample which exhibited failures requiring complex repairs.

Similar conclusions regarding Bosch and Ariston boilers' increased time until failure and greater reliability have been presented in other publications [22, 23]. It is also worth noting that, in comparison, the technical data sheets of the

Table I. Summary of the clustering results

	Ariston		Bosch		Buderus		Viessmann	
	Jupiter	Our program	Jupiter	Our program	Jupiter	Our program	Jupiter	Our program
Dunn index	0.195	0.200	0.259	0.260	0.223	0.220	0.235	0.240
Silhouette index	0.509	0.841	0.628	0.965	0.627	0.923	0.594	0.911
Separability	0.497	0.498	0.528	0.528	0.502	0.503	0.478	0.478
Compactness	2.043	2.043	1.463	1.463	1.62	1.62	1.639	1.639
Cluster centers	0.938 0.069 0.113	0.938 0.069 0.113	0.528 0.383 0.609	0.528 0.383 0.609	0.156 0.855 0.951	0.156 0.855 0.951	0.846 0.250 0.172	0.846 0.250 0.172
	0.432 0.354 0.644	0.432 0.354 0.644	0.116 0.817 0.913	0.116 0.817 0.913	0.872 0.256 0.197	0.872 0.256 0.197	0.580 0.619 0.576	0.580 0.619 0.576
	0.106 0.751 0.908	0.106 0.751 0.908	0.937 0.073 0.114	0.937 0.073 0.114	0.431 0.608 0.750	0.431 0.608 0.750	0.100 0.866 0.930	0.100 0.866 0.930

Source: Authors

Buderus and Viessmann boilers requested longer periods between diagnostics and maintenance work (*i.e.*, six months longer), which could lead to the untimely detection and elimination of early-stage defects, which could subsequently become more critical [24].

The separation of clusters is clearly observed in the case of the Buderus and Viessmann boilers. This may be due to differences in assembly quality (*i.e.*, separation of the upper left cluster of a more reliable group) or to an abrupt transition of boilers from a standard condition to serious failure (*i.e.*, gap between the central and lower right cluster). Such a sharp transition (cluster separability) is more widely associated in scientific research [25, 26] with the peculiarities of gas distribution networks' wear than with the insufficient reliability of the equipment itself. This is also more likely to occur in countries where the gas transmission system mainly consists of steel pipes (which is more typical for the countries of the former COMECON) [27].

To verify this fact, additional information is required, particularly concerning the frequency of boiler failures depending on the distribution network type (plastic, steel). The materials of these brands' gas burners may require additional reliability testing.

It can only be stated with certainty that there is a group of factors (external or internal) that leads to the emergence of such a sharp separation in the case of the central and right lower clusters for the Buderus and Viessmann brands boilers. Perhaps more clarity on this issue will be obtained through further research.

The slight separability of the central and upper right clusters for the Bosch and Ariston boilers indicates a higher quality of the assembled equipment, the majority of which has either the declared characteristics or an additional strength margin. This is also confirmed by the increased time until failure of the aforementioned brands [23], as well as by the fact that smart work management systems were used in these products first [28]. In general, these factors contribute to the accumulation of the largest clusters in the upper right corner and in the center (high and normative reliability).

In addition to the technical component (*i.e.*, the reliability and failure rate of double-circuit gas boilers of various brands), there are slight differences in the results of data clustering using Jupyter Notebook and our program. As stated earlier, this is due to the heterogeneity and greater breadth of the Jupyter Notebook libraries used. Nevertheless, the use of two different programs with a common methodological basis allowed verifying the calculations.

However, despite the correctness of the applied methods and the presented results, the most accurate conclusions regarding the studied boilers could only be obtained by conducting a broader statistical analysis.

It is necessary to mention that the data samples presented in this study are small (25 boilers of each brand) and were taken from one region, *i.e.*, the Kaliningrad region of the Russian Federation (even though the data were provided by different companies carrying out repairs). It is possible that, with a larger dataset, the reliability indicators might paint a different picture. A larger dataset could eliminate bias in the

recorded values, which might occur due to both regional influence (usage and temperature conditions as well as the quality of gas purification) and the quality of the supplied batches of this equipment. The influence of these factors and the analysis of larger databases are issues for further research.

The desirability of such an approach to assessing the quality of using equipment, especially under different external conditions, is indicated by [29] and [30].

Conclusions

To process the results of the observations of service companies on the failure of gas boilers of various brands, this study proposes an approach that allows assessing the severity of equipment failure based on frequency and repair complexity. A cluster analysis was conducted while considering the equipment's time until failure as well as its efficiency. Four brands were analyzed: Bosch, Ariston, Buderus, and Viessmann. Clustering was carried out using our own software and the publicly available Scikit-learn library of Jupyter Notebook. As a result, a convenient method for processing statistical information on gas boiler failures was obtained, which allows verifying the results through two software products prior to visualization.

Studying the data on gas boiler failure statistics in Kaliningrad, as recorded by three service companies in 2022, allows drawing a number of analytical conclusions. It was found that, for the studied samples, three groups of clusters are clearly formed: devices with increased reliability, devices with normative reliability, and unreliable devices. The most reliable gas boilers were of the Ariston brand, which exhibited lowest failure frequency and severity. This was also the case with Bosch boilers, which showed a smaller share of severe failures. Buderus and Viessmann boilers are less reliable. This may be due to issues in software or hardware, the specifics of the production technology, or the specifics of service work (*i.e.*, lower frequency). The efficiency parameter of the Ariston and Bosch boilers before failure was also higher.

The results show that Buderus and Viessmann boilers require more frequent monitoring by service companies. Furthermore, considering the frequency and severity of the reported failures, service companies are advised to pay attention to the quality of their repair work, use original spare parts, and ensure their timely availability. A more detailed understanding of the results, as well as their confirmation based on the statistics of other regions' gas boiler failures could be provided in further research. We propose conducting a comparative analysis of the results with data from other cities, as this will allow determining the influence of natural factors. These parameters could vary and have a significant impact on equipment reliability. Secondly, it is necessary to compare data not only based on the brand, but also on individual models, or to link them to batches from manufacturing plants (equipment reliability can also vary from model to model and from batch to batch, even within the same manufacturer).

The obtained results can be used, on the one hand, as practical information in planning the activities of service companies regarding diagnostics and maintenance work (*e.g.*, in determining the volume and type of the work planned). On the other hand, the presented research materials and the proposed methodological and software apparatus should serve as a good basis for conducting further studies on the reliability of gas boilers of different brands during operation.

CRediT author statement

Shcherban, P. Conceptualization, investigation, methodology, project administration, writing (review and editing).

Abu-Khamdi, R. Data curation, software, validation, visualization, writing (original draft).

Conflicts of interest

The authors declare no competing interests.

References

[1] S. Liu, "Corrosion failure analysis of the heat exchanger in a hot water heating boiler," *Eng. Fail. Anal.*, vol. 142, no. 1, pp. 106847, Jan. 2022. https://doi.org/10.1016/j.engfailanal.2022.106847

[2] F. Samanlioglu, "Evaluation of gas-fired combi boilers with HF-AHP-MULTIMOORA," *Appl. Comput. Intell. Soft Comput.*, vol. 2022, no. 1, pp. 1–10, Jan. 2022. https://doi.org/10.1155/2022/9225491

[3] K. Simic, "Modelling of a gas-fired heating boiler unit for residential buildings based on publicly available test data," *Energy Build.*, vol. 253, no. 1, art. 111451, Jan. 2021. https://doi.org/10.1016/j.enbuild.2021.111451

[4] K. Ritosa, "A probabilistic approach to include the overall efficiency of gas-fired heating systems in urban building energy modelling," *J. Phys. Conf. Ser.*, vol. 2069, no. 1, art. 012105, Jan. 2021. https://doi.org/10.1088/1742-6596/2069/1/012105

[5] Y. Chang, "Statistics and analysis of test data of industrial boiler approved products in 2020," in *Proc. Int. Conf. Adv. Manuf. Technol. Manuf. Syst.* (ICAMTMS), vol. 12309, pp. 604–610, Jan. 2022. https://doi.org/10.1117/12.2645720

[6] V. Prabhu and D. Chaudhary, "Machine learning-enabled condition monitoring models for predictive maintenance of boilers," in *Proc. 4th Int. Conf. Recent Dev. Control Autom. Power Eng.* (*RDCAPE*), Nov. 2021, pp. 426–430. https://doi.org/10.1109/RDCAPE52977.2021.9633534

[7] A. R. Aikin, "The process of effective predictive maintenance," *Tribol. Lubr. Technol.*, vol. 77, no. 2, pp. 34–40, Feb. 2021

[8] S. Mohanty, K. C. Rath, and O. P. Jena, "Implementation of total productive maintenance (TPM) in the manufacturing industry for improving production effectiveness," in *Industrial Transformation*, 1st ed. Boca Raton, FL, USA: CRC Press, 2022, pp. 45–60. https://doi.org/10.1201/9781003229018-3

[9] N. Xodjiev *et al.*, "Analysis of the resource-saving method for calculating the heat balance of the installation of hot-water heating boilers," in *AIP Conf. Proc.*, vol. 2432, no. 1, art. 020019, Jan. 2022. https://doi.org/10.1063/5.0090455

[10] W. Ma *et al.*, "Energy efficiency indicators for combined cooling, heating, and power systems," *Energy Convers. Manag.*, vol. 239, art. 114187, Jan. 2021. https://doi.org/10.1016/j.enconman.2021.114187

[11] M. Larbi Rebaiaia and D. Aït-Kadi, "Maintenance policies with minimal repair and replacement on failures: Analysis and comparison," *Int. J. Prod. Res.*, vol. 59, no. 23, pp. 6995–7017, Dec. 2021. https://doi.org/10.1080/00207543.2020.1832275

[12] J. Leukel, J. González, and M. Riekert, "Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review," *J. Manuf. Syst.*, vol. 61, pp. 87–96, Jan. 2021. https://doi.org/10.1016/j.jmsy.2021.08.012

[13] C. Acuña-Soto, V. Liern, and B. Pérez-Gladish, "Normalization in TOPSIS-based approaches with data of different nature: Application to the ranking of mathematical videos," *Ann. Oper. Res.*, vol. 296, no. 1, pp. 541–569, Jan. 2021. https://doi.org/10.1007/s10479-018-2945-5

[14] A. Chabane, S. Adjerid, and I. Meddour, "Dependability analysis in systems engineering approach using the FME-CA extracted from the SysML and failure modes classification by K-means," *Int. J. Dyn. Cont.*, vol. 10, no. 3, pp. 981–998, Jul. 2022. https://doi.org/10.1007/s40435-021-00855-8

[15] F. Aksan *et al.*, "Clustering methods for power quality measurements in virtual power plants," *Energies*, vol. 14, no. 18, p. 5902, Sep. 2021. https://doi.org/10.3390/en14185902

[16] W. Peiyi *et al.*, "Analysis and research on enterprise resumption of work and production based on K-means clustering," in Proc. 6th Int. Conf. Big Data Anal. (IC-BDA), pp. 169–174, Dec. 2021. https://doi.org/10.1109/ICBDA51983.2021.9403217

[17] L. Quaranta, F. Calefato, and F. Lanubile, "KGTorrent: A dataset of Python Jupyter notebooks from Kaggle," in *Proc. IEEE/ACM 18th Int. Conf. Mining Softw. Repos. (MSR)*, May 2021, pp. 550–554. https://doi.org/10.1109/MSR52588.2021.00072

[18] H. Fangohr, T. Kluyver, and M. Dipierro, "Jupyter in computational science," *Comput. Sci. Eng.*, vol. 23, no. 2, pp. 5–6, Mar. 2021. https://doi.org/10.1109/MCSE.2021.3059494

[19] C. J. Weiss and A. Klose, "Introducing students to scientific computing in the laboratory through Python and Jupyter notebooks," in *ACS Symp. Ser. Teach. Program. Across Chem. Curric.*, Jan. 2021, pp. 57–67. https://doi.org/10.1021/bk-2021-1387.ch005

[20] E. Mazur, P. Shcherban, and V. Mazur, "Research and evaluation of the operating characteristics of used ship engine oil using the process parameter matrix method," *FME Trans.*, vol. 51, no. 4, pp. 497–503, Dec. 2023. https://doi.org/10.5937/fme2304497M

[21] J. Piazentin Ono, J. Freire, and C. T. Silva, "Interactive data visualization in Jupyter notebooks," *Comput. Sci. Eng.*, vol. 23, no. 2, pp. 99–106, Mar. 2021. https://doi.org/10.1109/MCSE.2021.3052619

[22] P. Shcherban, A. Sokolov, and R. A. Hamdi, "Study of failure statistics of cavitators in the fuel oil facilities through the application of regression and cluster analysis," *Proc. Eng. Sci.*, vol. 5, no. 1, pp. 39–48, Jan. 2023. https://doi.org/10.24874/PES05.01.004

[23] J. Zhao *et al.*, "Root cause analysis of a cracked primary heat exchanger in a gas wall-mounted boiler," *Eng. Fail. Anal.*, vol. 153, art. 107583, Jan. 2023. https://doi.org/10.1016/j.engfailanal.2023.107583

[24] R. Oliver, A. Duffy, and I. Kilgallon, "Statistical models to infer gas end-use efficiency in individual dwellings using smart metered data," *Sustain. Cities Soc.*, vol. 23, pp. 1–10, Nov. 2016. https://doi.org/10.1016/j.scs.2016.01.009

[25] Z. Liao, M. Swainson, and A. L. Dexter, "On the control of heating systems in the UK," *Build. Environ.*, vol. 40, no. 3, pp. 343–351, Mar. 2005. https://doi.org/10.1016/j.buildenv.2004.05.014

[26] A. V. Gubarev, "Determination of the thermotechnical measures of the condensing water heating boiler's high-temperature section," in *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 791, no. 1, art. 012011, Jan. 2020. https://doi.org/10.1088/1757-899X/791/1/012011

[27] M. Bukurov, S. Bikić, and B. Marković, "Efficiency and management of gas boilers in public buildings in Vojvodina," *J. Process. Energy Agric.*, vol. 20, no. 2, pp. 87–92, Jun. 2016.

[28] J. P. Xu, X. Q. Ma, and H. Shen, "Remote control system design for domestic wall-mounted gas boilers based on GSM modem," *Adv. Mater. Res.*, vol. 650, pp. 559–564, Mar. 2013. https://doi.org/10.4028/www.scientific.net/AMR.650.559

[29] M. S. Mahmud *et al.*, "A survey of data partitioning and sampling methods to support big data analysis," *Big Data Min. Anal.*, vol. 3, no. 2, pp. 85–101, Jun. 2020. https://doi.org/10.26599/BDMA.2019.9020015

[30] A. M. Ikotun *et al.*, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci.*, vol. 622, pp. 178–210, Jan. 2023. https://doi.org/10.1016/j.ins.2022.11.139