IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 12 September 2024, accepted 12 October 2024, date of publication 21 October 2024, date of current version 6 November 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3484000

RESEARCH ARTICLE

Classification of Distributions of the Values of Impulses Forcing Vibrations of an Oscillator

NATALIA FRANKOWSKA, PRZEMYSŁAW FRANKIEWICZ, AND AGNIESZKA OZGA
Faculty of Mechanical Engineering and Robotics, AGH University of Krakow, 30-059 Kraków, Poland

Corresponding author: Natalia Frankowska (frankows@agh.edu.pl)

This work was supported by the Ministry of Science and Higher Education in Poland under Grant 16.16.130.942 and Grant 10.16.130.7990.

ABSTRACT The present work discusses the classification of distribution of values of impulses. We solve the problem in which random impulses occur before the vibrations caused by previous random impulses expire. The goal was to achieve a correct classification of three groups of distributions characteristic of three different processes: correct, transitory and flawed. The starting moment of the analysis as well as the length of time interval are the objects of study of this article. In order to determine the significance of the extracted features of the random time series under study, SHapley Additive exPlanations method was applied. The paper discussed the stage of study which is focused on finding the first three stochastic raw moments that allow for the most precise classification of distributions characterizing different processes in the shortest time. The analysis using the random forest classifier made it possible to distinguish the distributions characteristic for different types of processes, that is, it is possible to distinguish correct processes and transitory ones, etc. It is hardest to distinguish correct processes from one another, since they are highly similar to one another, but different from the other types.

INDEX TERMS Stochastic mechanics, random series of impulses, stationary time series, non-stationary time series, classification, feature engineering.

I. INTRODUCTION

The theory of stochastic dynamic systems is an interdisciplinary field which is also applied in mechanics. In the case of mechanic non-deterministic systems there are no mathematical instruments that would allow for unequivocal description of the condition of a system at any given moment on the basis of its previous conditions. This issues from the random character of forces acting on this system. Hence, in the 20th century several research centers initiated studies on application of stochastic equations [1], [2], [3], [4], [5], [6], [7], [8] in the analysis of dynamics of mechanical systems. Instruments are searched for, which would make it possible to foresee future states of dynamic systems with the least level of uncertainty. The researchers work on assessment of the risk, analysis of stability and optimization of functioning of the system in the presence of random disruptions.

Among other tasks there were investigations facilitating the analysis of actions of the systems whose vibrations were forced by series of random impulses. First, mathematical models were developed [9], [10], [11], [12], [13], [14], [15]. In the subsequent years the first attempts at verification of the models with the help of simulation methods were made [16], [17], [18].

The development of Artificial Intelligence started the new stage of studies on stochastic mechanics [19], [20]. Application of the methods of machine learning is also a new approach to classification of distributions of values of stochastic impulses forcing vibrations of discrete systems. The innovatory approach [20] to solution of some problems of classification of discrete signals were presented in the work. The investigations discussed in the present paper consider diffrent time intervals and the methods of classification are based on decision trees. These classification techniques

The associate editor coordinating the review of this manuscript and approving it for publication was Diego Bellan.

are extensively applied in time series classification [21], [22].

II. DESCRIPTION OF THE RESEARCH PROBLEM AND DEFINING THE GOAL OF RESEARCH

Production of energy in a coal power plant requires control over the burning process. Huge coal particles and small coal particle that can be found in the dust pipe indicate incorrections in the milling process and cause incorrect burning. Looking through the data base of the Web of Science we can see that merely twenty seven papers deal with the problem of granulation of coal particles. Scientists discuss fine-grained waste materials [23], creation of coal and peat fuel compositions for burning in solid-fuel boilers [24] or enhancement of coal reactivit [25]. There are also studies on atmospheric emissions of carbon compounds depending on the burnt solid fuel, including the size of coal grains [26].

Diagnostics of the size of coal particles is still an open problem.

The mathematical model for calculating of distributions of impulses (that is, particles of different sizes) was developed in the previous studies [15]. The model is independent of the number of impulses, their values and intensity of occurrence as well as the distributions of impulse values. It is only in the case when time tends to infinity [27], [28] that time series calculated for subsequent raw moments are ergodic and the distributions calculated from the model at any time moment are flawless. Therefore it cannot be applied in technology, where the solution to a reverse problem, that is, recognition of distributions of impulse values ought to be reached within a few seconds or a few minutes.

Earlier studies [28] showed that for oscillators with strong damping and high frequency, when the impulses occur extremely rarely, there is a chance of calculating the distributions of values with a small uncertainty threshold. Yet, there is no chance of applying the model in several other cases - the waiting time for the time series to become ergodic is too long. Having applied the methods of Artificial Intelligence, in the current paper we are seeking a possibility of classifying time series when the intensity of hits is so high that subsequent impulses occur before the vibrations caused by the previous one die out. The analyzed time series are not ergodic, for a majority of trials they will not be even stationary.

III. MATHEMATICAL BACKGROUND

A. THE DISCRETE SYSTEM VIBRATING UNDER THE INFLUENCE OF A RANDOM SERIES OF IMPULSES

The authors analyse the vibrations of an oscillator with damping (1)

$$\frac{d^2x}{dt^2} + 2b\frac{dx}{dt} + a^2x = f(t) \tag{1}$$

where a and b are parameters of the oscillator.

The vibrations are forced with the help of random series of impulses (2),

$$f(t) = \sum\nolimits_{t_i<t} \eta_i \delta(t - t_i) \tag{2}$$

where η is the value of the i-th impulse, $i = 1, 2, \cdots, n$, t_i is i-th moment of excitation of the movement, $\delta(t - t_i)$ is Dirac distribution at time t_i [29].

In the domain of stochastic mechanics that deals with the response of systems to being forced by a random series of impulses, it is hard to find a uniform notation. Taking into consideration equation (2) only, we should suggest the description like coefficient occurring at Dirac delta. In the most frequently cited work [12] dealing with impulse forcings, the coefficient η was called strength of the impulse while in the subsequent work [30] η is described as random vectors representing the direction of the impulses. These descriptions do not reflect the character of random variables that perform the key role in the discussed problem. In the further part of the paper the name the impulse value has been used [9], [31].

B. RANDOM VARIABLES DEFINING THE RANDOM SERIES OF IMPULSES

A random variable is defined by the set of values and by the distribution of probability determining the probability of occurrence of the value. A random series of impulses (2) have been described by two random variables – one discrete and the other one continuous. Random variable ϕ, of discrete character, defines the values of impulses η, which in the case under consideration are independent realisations of a random variable with finite mean value and finite variance. In turn, time intervals between impulses $t_i = (t_i - t_{i-1})$ are independent realisations of the random variable for which the function of probability density is in the form of exponential distribution (3).

$$t(\tau) = \begin{cases} \lambda e^{-\lambda\tau} & \text{for} \quad \tau \geq 0 \\ 0 & \text{for} \quad \tau < 0 \end{cases} \tag{3}$$

For the assumptions formulated in this way the solution of the equation (1), (2), (3) assumes the form:

$$x(t) = \frac{1}{c}\sum\nolimits_{0<t_i<t} \eta_i e^{-b(t-t_i)} sin(c(t - t_i)) \tag{4}$$

where b is damping coefficient, c is frequency of damped vibrations, η is the value of i-th impulse, η_i is a sequence of independent identically random variables with finite expectation.

C. RAW MOMENTS COMPUTED FROM THE MOVEMENT OF THE SYSTEM

Moments of occurrence of impulses as well as their values are independent random variables. By making the signal discrete we can establish the estimators of raw moments of the order $k = 1, 2, \cdots, m$.

$$\bar{m}_i = \frac{1}{[t/h]}\sum\nolimits_{n<t/h} x^i(nh) \tag{5}$$

TABLE 1. Parameters of distributions characterizing a correct processs.

Distributions characterizing a correct process	
$\phi_A : p(\eta_1 = 80) = 0.5,\ p(\eta_2 = 70) = 0.5$	
$\phi_G : p(\eta_1 = 85) = 0.5,\ p(\eta_2 = 65) = 0.5$	
$\phi_K : p(\eta_1 = 90) = 0.5,\ p(\eta_2 = 60) = 0.5$	

TABLE 2. Parameters of distributions characterizing a transitory process that is the one transforming in a wrong direction.

Distributions characterizing a transitory process that is the one transforming in a wrong direction	
$\phi_I : p(\eta_1 = 100) = 0.5,\ p(\eta_2 = 50) = 0.5$	
$\phi_R : p(\eta_1 = 110) = 0.5,\ p(\eta_2 = 40) = 0.5$	
$\phi_O : p(\eta_1 = 120) = 0.5,\ p(\eta_2 = 30) = 0.5$	

TABLE 3. Parameters of distributions characterizing a flawed process.

Distributions characterizing a flawed process	
$\phi_E : p(\eta_1 = 130) = 0.5,\ p(\eta_2 = 20) = 0.5$	
$\phi_C : p(\eta_1 = 140) = 0.5,\ p(\eta_2 = 10) = 0.5$	
$\phi_M : p(\eta_1 = 145) = 0.5,\ p(\eta_2 = 5) = 0.5$	

where h is value of the coefficient responsible for discretization of the signal, t is time, $n = 1,2, \cdots$ is number of the subsequent sample in the signal discrete.

In the study we check the possibility of classification of random movement on the basis of the analysis of raw moments (5). Such methods of classification are searched for, which recognition of the distribution of a random series of impulses should depend only on the values of impulses included in the series. The impact on the raw moments of time in which a subsequent impulse occurs should be reduced to a minimum. Therefore, like in the analysis of continuous random variables, an individual point, that is, a certain moment of time, cannot be analysed. The analysis should cover a certain time interval and its starting point as well as its duration are the object of study in the current paper.

Due to the fact that experimental analysis of Dirac delta function (2) is impossible, and an attempt at its implementation in experimental investigations would transform the mathematical model (4), a decision was made to conduct simulation studies focused mostly on engineering of features describing time series of raw moments. Features that could be used in classification of distributions are searched for. Experimental studies taking into account the difference between the model of a hit as described by Dirac delta function and an actual hit will be possible to execute at a later date thanks to the directions of analysis determined at the present stage of investigations.

IV. METHOD

A. PREPARATION OF THE DATASET

There are infinitely many possible distributions of a discrete random variable. In actual technological systems, however, the goal is not to distinguish all possible distributions. The research is focused on possible methods of recognizing correct processes and those marked with errors.

TABLE 4. Statistical parameters of distributions characterizing a correct process.

Parameter	A	G	K
Mean value	75	75	75
Standard deviation	5	10	15
Coefficients of variation	7%	13%	20%

TABLE 5. Statistical parameters of distributions characterizing a transitory process hat is the one transforming in a wrong direction.

Parameter	I	R	O
Mean value	75	75	75
Standard deviation	25	35	45
Coefficients of variation	33%	47%	60%

TABLE 6. Statistical parameters of distributions characterizing a flawed process.

Parameter	E	C	M
Mean value	75	75	75
Standard deviation	65	70	55
Coefficients of variation	87%	93%	73%

TABLE 7. Parameters of the simulations.

Parameter	Parameter ‘s value
l	10
c	20
b	10
period of time	1800 s
number of random samples generated for each distribution	1000

In the current study, we define a correct process as one that involves impulses of similar values (Table 1). A process marked with errors (flawed one) is the one (Table 2), which includes impulses of high and low values. There are also transitory processes (Table 3), that transform from one to the other, that is from a correct process to flawed one or vice versa.

Legends in the tables show the colours that mark particular classes in all visualizations. The classes representing correct processes are shown in various shades of green, classes representing flawed processes are shown in various shades of red while the classes representing transitory processes are shown in various shades of yellow and orange.

Nine distributions including just two impulses with the occurrence probability equal to 0.5 were selected for the study of the features of time series. Distributions of impulse values were selected so that they had the same mean

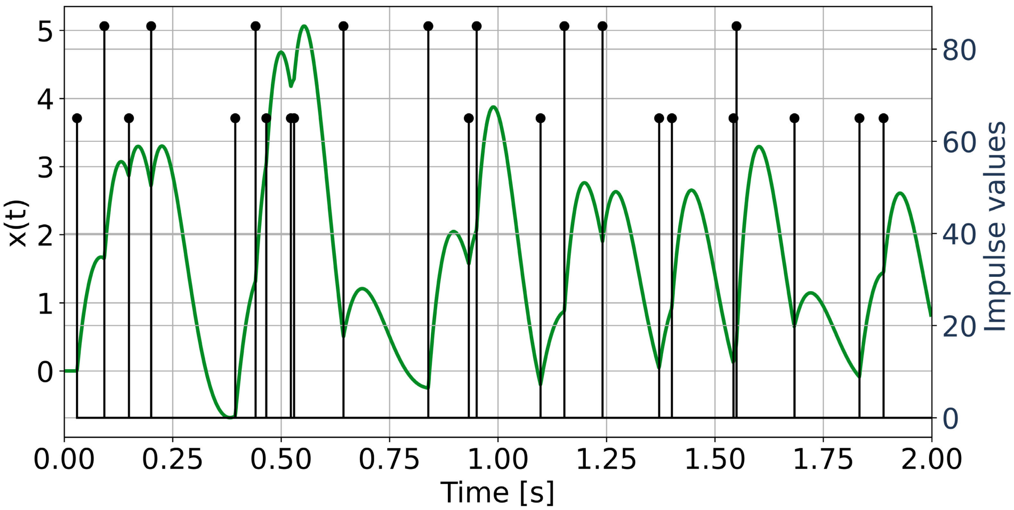

FIGURE 1. Sample movement of vibrations of an oscillator in a correct process.

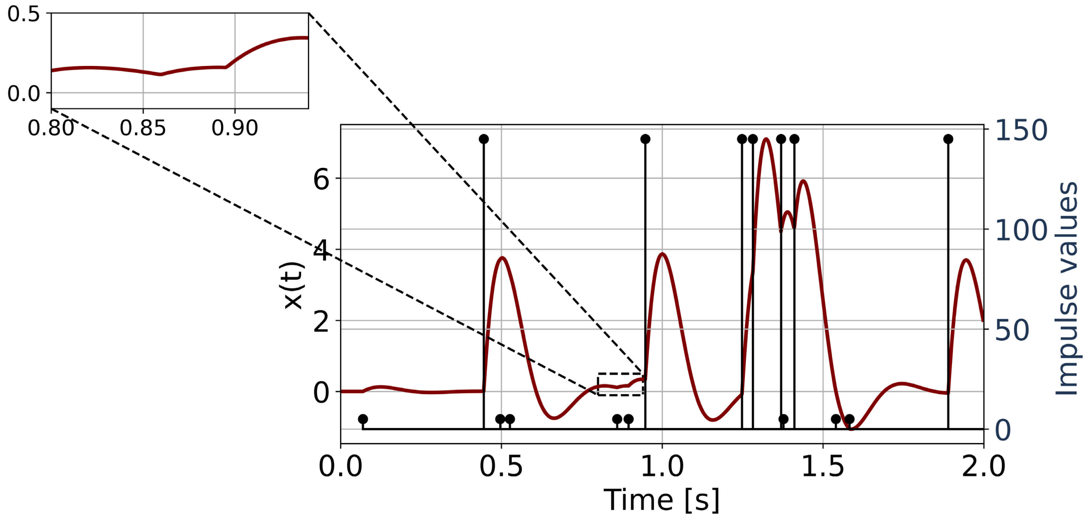

FIGURE 2. Sample courses of vibrations of an oscillator in a flawed process.

value and different statistical parameters shown in Tables 4, 5, and 6 [32], [33]. Distributions including errors in the process are characterized by high standard deviation and high changeability coefficient, correct distributions are marked by low coefficients of variation and low standard deviation.

Sample movement of vibrations of an oscillator in a correct process and a flawed one were shown in Figs. 1 and 2 [34]. Parameters of the simulations were presented in Table 7.

The smaller the impulses occur in a random series, the harder it is to notice changes in the movement of vibrations. Hence, further analyses will be focused on estimators of moments (5) and not courses.

B. SELECTING OF A FRAGMENT OF TIME SERIES FOR CLASSIFICATION

In the case when two random variables occur: a continuous one T describing the intervals between impulses, and that of distribution of values $\phi(\eta)$, recognition of the distribution of values requires two parameters connected with time: the length of time interval that is necessary for the analysis, and the time moment when the analysis begins.

An analysis should start at the moment when the subsequent impulses exert the least influence on the changes at raw moments computed from the movement of vibrations. Additionally, the differences between the imposed distribution and that executed by the pseudorandom generator of random numbers should be as little as possible both for the T variable and the ϕ distributions. Estimators of the moments as well as estimators of executed distributions are computed from the moment $t = 0$ till a certain time moment. Both the time moment from which the investigations can be started and the duration of the time interval should be initially decided upon on the basis of exploration research and then verified

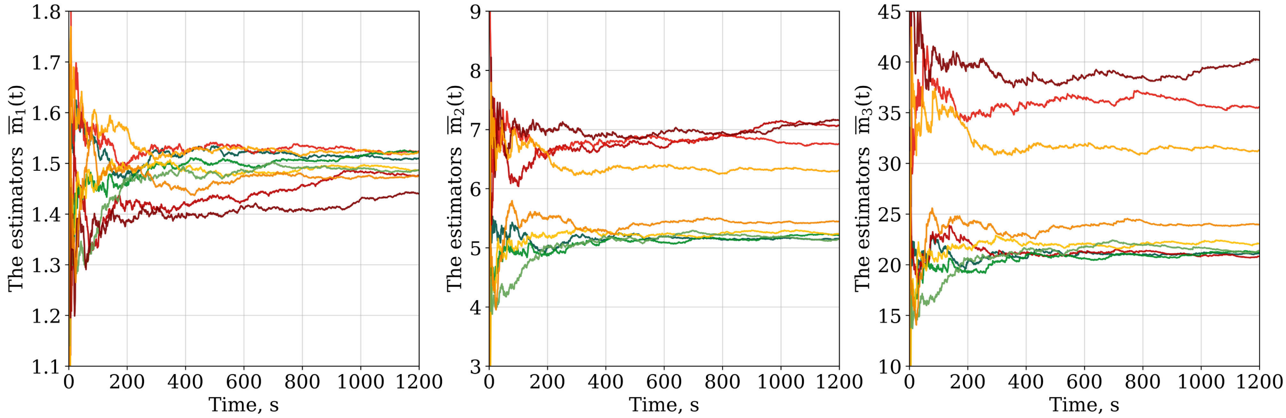

FIGURE 3. The first, the second and the third stochastic raw moments $\overline{m}$ calculated from the location *x(t)* for a selected samples.

by algorithms of Artificial Intelligence with the use of feature engineering.

In the investigations described in the current paper, three initial stochastic raw moments were computed with the use of the equation (5) for the nine generated distributions ϕ. It must be remembered, however, that they are individual trials, and in these investigations a one thousand trials were generated. Fig. 3 represents stochastic moments for two trials of each of the nine distributions. In accord with the assumption of the simulation, in the case of the initial moment, time series of the moments approach one value, which is why they seem to be indistinguishable. Subsequent stochastic moments differ in mean values, which are approached by particular signals generated in a given class.

C. EXPLORATORY INVESTIGATIONS

In order to better visualize all analysed trials, channels were designated by way of determining the maximum value, minimum one and the mean value (Fig. 4) from all trials in particular classes, every second.

Exploratory research indicated that in the case of the second and third stochastic moments, the time series of moments become, in a way, distinguishable after 300 seconds. This observation was confirmed by the preliminary studies connected with classification. Additionally, exploratory research showed that further division of particular tunnels takes place at certain time intervals that are difficult to describe precisely. In the case of the second moment, after 600 seconds we are able to distinguish fairly precisely only those groups that differ significantly from one another. Distributions ϕ_A, ϕ_G, ϕ_I, ϕ_K, and ϕ_O, ϕ_R are indistinguishable. This means that the distributions characterizing a correct process ϕ_A, ϕ_G, ϕ_K get mixed with the distribution characterizing a transitory process ϕ_I. In the case of the second and third moments, after 600 seconds most classes become possible to distinguish, and the groups characterizing different types of processes seem distinguishable as well. Hence it was decided that the studies connected with Artificial Intelligence algorithms covered 300 second intervals starting at the 300th, 600th and 900th second.

D. STATIONARY CHARACTER OF AN INDIVIDUAL TRIAL

Stationary character of the examined time series was checked in three time windows for all three stochastic moments. On the level of significance equal to 0.05 examination of stationary character was conducted with the use of two tests: Augmented Dickey-Fuller (ADF test) and Phillips-Perron (PP test) [35]. Both tests showed that in the II, III and IV time window stationary signals constitute approximately 1% of all examined series. The number of stationary signals of subsequent models in the examined time windows according to ADF test are presented in the Fig. 5.

E. ARTIFICIAL INTELLIGENCE ALGORITHMS

After the initial attempts at application of various kinds of algorithms of AI and executing a comparison of results, the authors decided to apply the algorithm of random forest [36]. A random forest consists of a definite number of decision trees, and the result of classification is decided by the highest number of votes received from individual trees. Each tree consists of a root, branches and nodes. Branches lead from the root to subsequent nodes, and at each node one condition is tested, which allows for the choice of branches leading to further nodes. If no branch spreads from a node, it is called a leaf. Each leaf is marked with a class, which is assigned an observation. The main advantage of random forest is that they allow for determining which features of a dataset are the most significant. It is also possible to visualize the division that has been executed.

In the case of random forest it is very important to select hyperparameters properly, because thanks to this the risk of overfitting of the model is reduced.In the current study the hyperparameters of the random forest were selected with the help of grid search method, which consists in searching of the grid of parameters with respect to a given criterion; in this case it was the highest precision. The

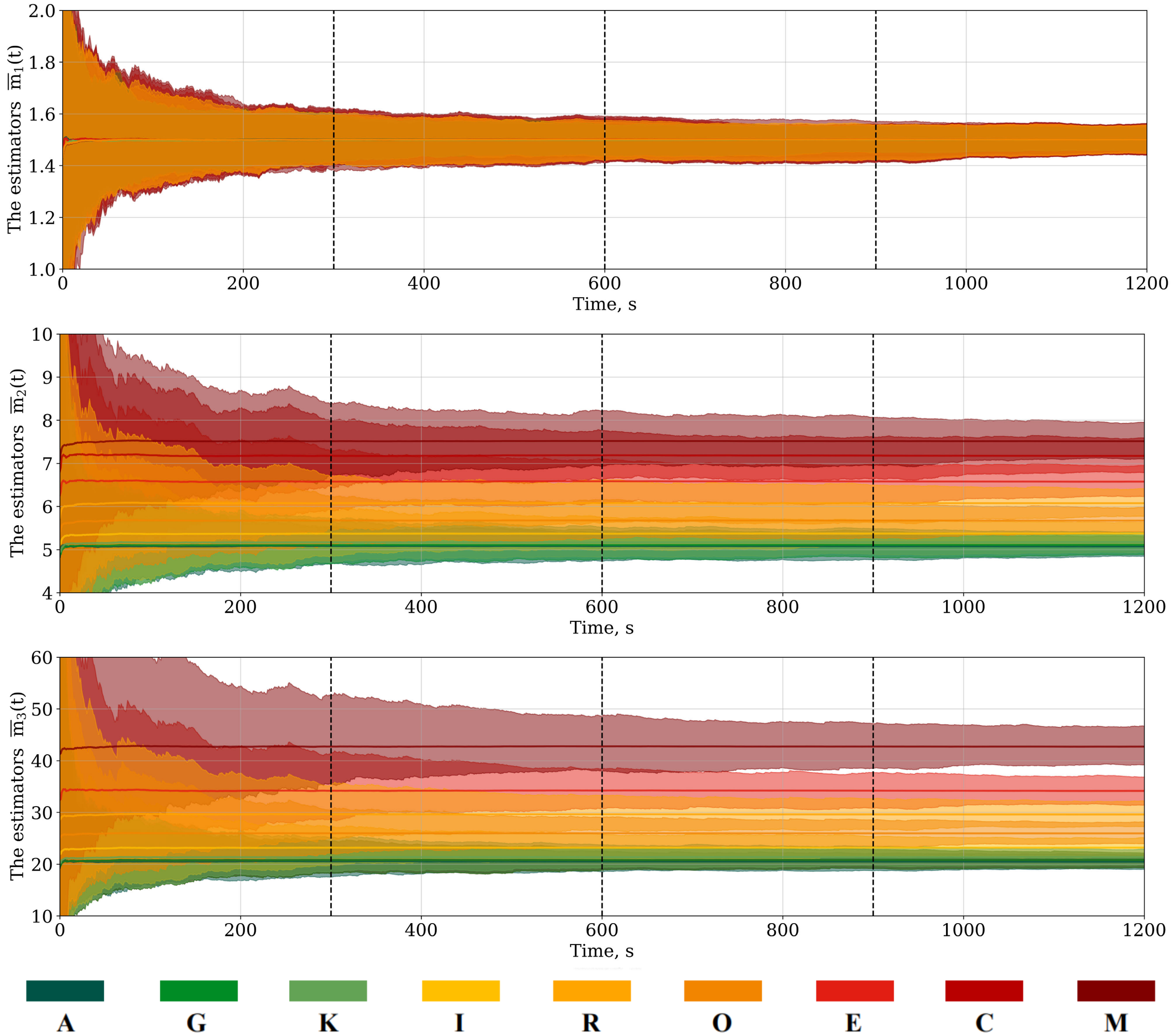

FIGURE 4. The first, the second and the third stochastic raw moments $\overline{m}$ calculated from the location *X(t)* for a thousand different movements.

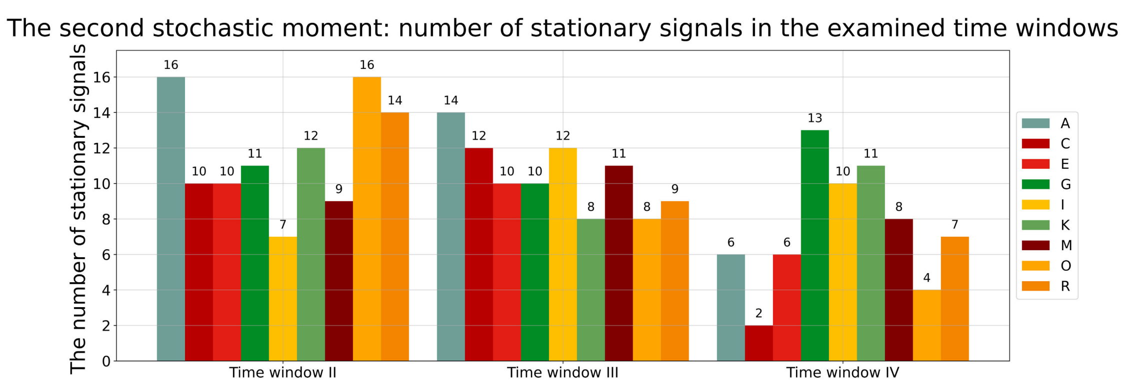

FIGURE 5. The number of stationary signals for the second stochastic moment in the examined time windows according to ADF test.

TABLE 8. Selected features.

	The second raw moment	The third raw moment
II time window	Wavelet energy Wavelet variance Sum absolute difference	Median Wavelet absolute mean -
III time window	Wavelet absolute mean Autocorrelation	Wavelet absolute mean Spectral distance
IV time window	Wavelet standard deviation Wavelet energy	Wavelet energy Maximum

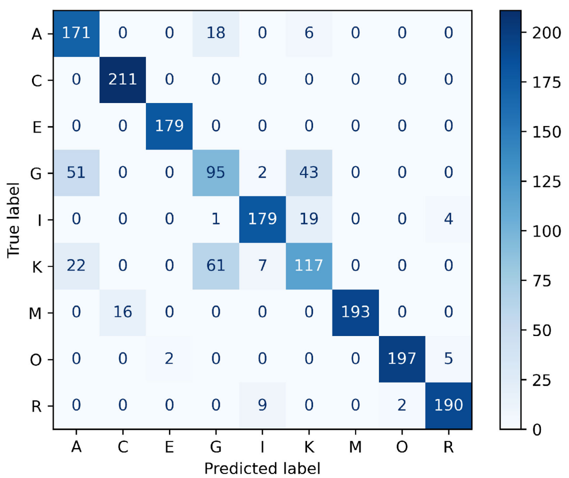

FIGURE 6. Confusion matrix for 30 features.

analysis was focused on the most important hyperparameters of the random forest, like the number of trees, the function measuring quality of division, maximum depth of the tree, minimum number of samples needed for node division and a minimum number of samples on a leaf.

Optimization of the selected hyperparameters of the forest allowed us to obtain reproducible results. In order to determine whether a division of a data set influenced the results, cross-validation was conducted – the dataset was divided into five equal parts, each of which became in turn a test set and the remaining parts was a training set. There were no significant differences between the obtained results, which means that random character of the division of the data set into the training set and the test one had no significant impact on the obtained result.

F. EVALUATION OF THE CLASSIFICATION

In evaluation of the classification, three metrics were applied:

- **accuracy** - the ratio of all correctly classified cases to all classified cases (6).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

where TP is true positive, FP is false positive, TN is true negative and FN is false negative.

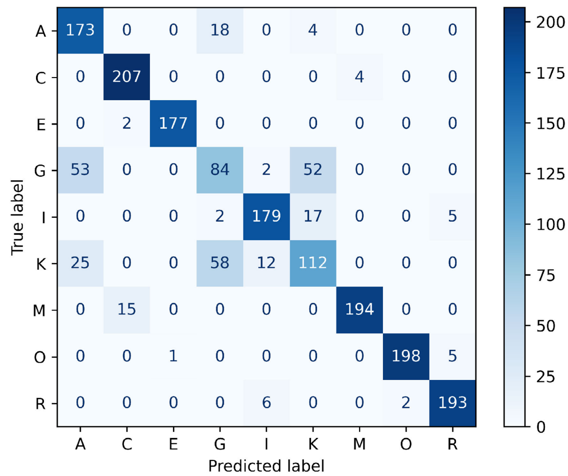

FIGURE 7. Confusion matrix for selected features.

TABLE 9. Metrics.

	case I		case II	
class	**precision**	**recall**	**precision**	**recall**
A	0,69	0,89	0,77	0,91
C	0,92	0,98	1,00	1,00
E	0,99	0,99	0,99	1,00
G	0,52	0,44	0,57	0,53
I	0,90	0,88	0,96	0,93
K	0,61	0,54	0,68	0,62
M	0,98	0,93	1,00	1,00
O	0,99	0,97	1,00	0,98
R	0,95	0,96	0,97	1,00
accuracy	0,84		**0,89**	

- **precision** - informing how many cases of all those expected to be positive (TP + FP) are truly positive (TP) (7).

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

- **recall** - the ratio of cases that were expected to be true positives (TP) to those that should have been assigned to a given class (TP + FN) (8).

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

G. SELECTION OF FEATURES

Time series were imported to Python environment where, with the help of Time Series Feature Extraction Library (TSFEL [37]) software statistical and spectra features of subsequent stochastic moments were extracted for each of time windows under consideration. In the first stage, the method of selects features from the tsfresh library [38] was used. The features whose impact on the classification of the nine distribution presented in the paragraph was insignificant, were rejected. This allowed for initial reduction of the size of

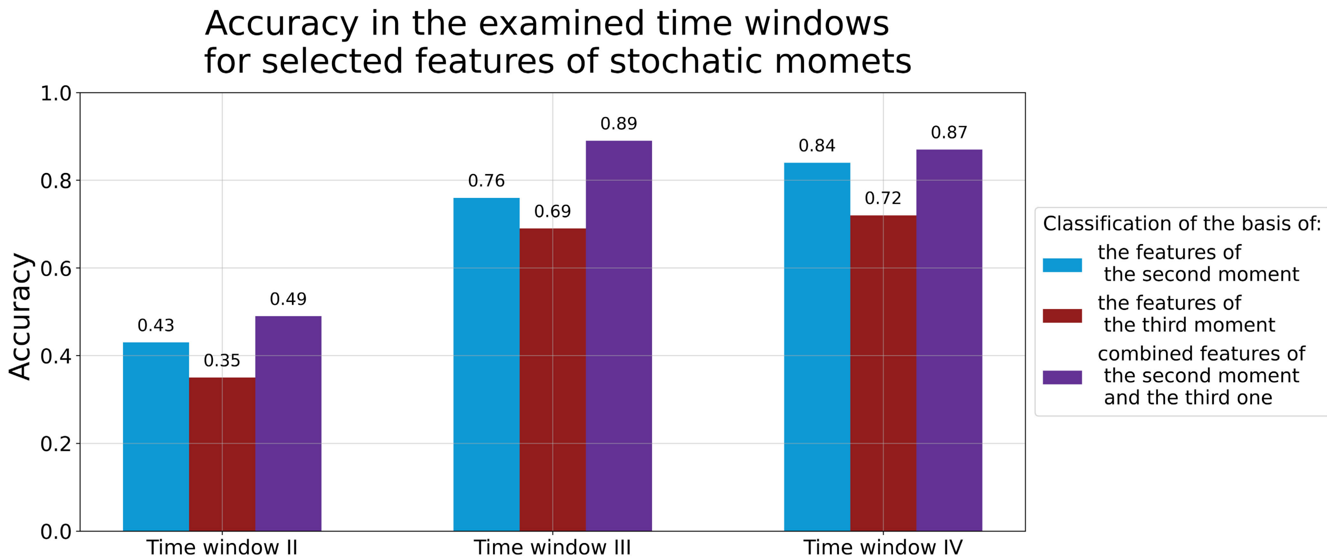

FIGURE 8. Accuracy in the examined time windows for selected feature of stochastic moments.

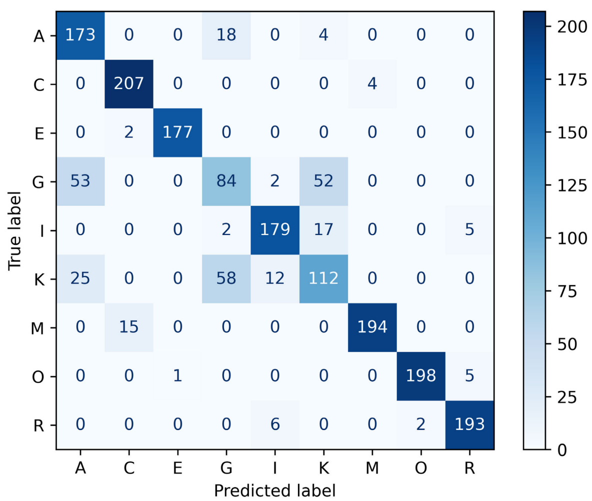

FIGURE 9. Confusion matrix for case I.

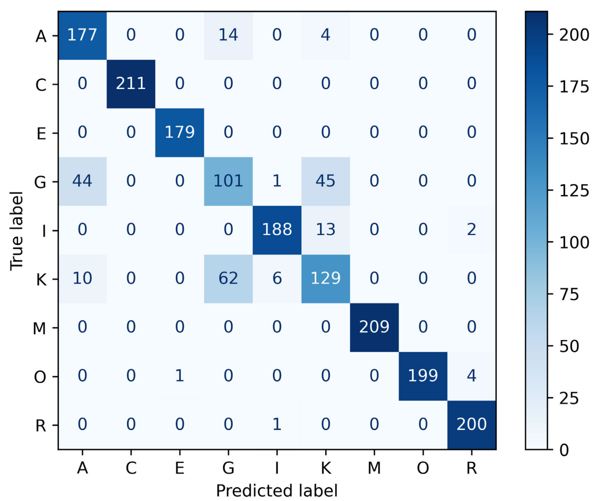

FIGURE 10. Confusion matrix for case II.

the data describing the properties of the time series. Initial selection was not possible for the first time window, since the analysis showed that none of the features had a significant impact on classification of the presented distributions. Later, the value of Spearman correlation coefficient between the values of features determined from particular moments and the class of distribution was examined and on this basis a decision was made to limit each of the sets to 30 most important features. The final decision concerning selection of features for each of the examined cases was made on the basis of SHAP (SHapley Additive exPlanations [39]) analysis. SHAP analysis allowed for examination of the influence of each of the 30 features on classification. In each examined case from 2 to 4 most significant features were selected (depending on the window and moment under consideration. In the case of the first moment in each of the examined windows, the greatest influence was exerted by the statistical features of absolute differences of the signal: the mean, sum and median - as well as the signal distance. As regards the remaining moments in each window other features were selected (Table 8).

Analyzing Table 8, we can say that the signal characteristics obtained after wavelet transformation that are of greatest importance. Selection of merely a few features was sufficient, and the obtained results were close to those acquired in the analysis using all generated features (Fig. 6, 7).

V. RESULTS

The investigations involved an attempt at classification distribution of values $\phi(\eta)$ on the basis of features of each moment as well as combined features of two moments (the second and the third one) and the features of all three moments in each window.

In the case of the first moment, in each of the time windows under study we are able to distinguish only those groups, which differ from one another significantly. Distributions characterizing the correct process are impossible to distinguish from the distribution characterizing a transitory one.

In the case of the second moment, after 600 s most classes become distinguishable for the groups characterizing different types of processes and poorly distinguishable for the processes of the same type, e.g, transitory ones.

In the case of the third moment, after 600 s most classes are classified correctly. What is classified incorrectly includes the samples representing, first of all, two distributions: distribution ϕ_C, characterizing a flawed process and the distribution ϕ_G, characterizing a correct process.

After combining the features of the moments we can classify fairly precisely the groups after the passage of 600 s. This means that using the features of at least two moments we are able to achieve higher precision in a shorter time (Fig. 8). Classification on the basis of combined features of the second moment and the third one brings slightly better results than those conducted on the basis of all three moments.

The results of the investigations will be discusses for two cases:

- **case I**: classification of the basis of the features of the second moment in the fourth time window (between 900 and 1200 s),
- **case II**: classification on the basis of combined features of the second moment and the third one in the third time window (between 600 and 900 s).

In the first case classification used two features of the second moment: wavelet standard deviation and wavelet energy. In the second case two features of the second moment: wavelet absolute mean as well as autocorrelation, and two features of the third moment: wavelet absolute mean and spectral distance. The data were imported to PYTHON environment and rescaled to the range 0 – 1. Then, for each vase hyperparameters were selected and classification was executed.

In accord with the formulae: (6), (7), (8) metrics were computed for each class (Table 9). Fig. 9 and 10 represent confusion matrix for the analysed cases.

Random forest copes well with the prediction of signals belonging to distribution characteristic of various processes. Both stationary signals and non-stationary ones are sometimes classified incorrectly. In both examined cases there are three groups that are the most difficult to recognize, since they are highly similar to one another: A (80/70), G (85/65) and K (90/60). It is worth noting, however, that each of these distributions is characterized by a correct process. Application of the second moment and the third one improved the classification of these groups by several percent (increase in precision and sensitivity). Precision of the model was also improved. The main advantage of application of features of two moments is obtaining better results in a shorter time. However, if we wanted to distinguish between groups A, G and K, it seems necessary to wait for new solutions in the Artificial Intelligence methods. At the present moment we demonstrate the best results that could be obtained. Even with the use of deep learning methods a better classification could not be obtained. An analysis of incorrectly classifies samples showed that the stationary character of time series does not influence classification.

VI. CONCLUSION

Due to the fact that size of particles in the dust pipe should be controlled as early as during the coal grinding process while diagnosing of the dust particle size is an open problem, the authors decided to develop a model of responses of the system to impulse forcing. Using the same model for analyses with the help of simulation investigations, the authors define different cases of systems, forcing distributions and intensity of forcings. In this way they identify difficulties encountered while solving the problem. The authors use the achievements of Artificial Intelligence methods, applying feature engineering as well as various types of algorithms, gradually approaching the solution of the reverse task, namely, identification of distributions of impulses forcing the vibrations of the system, based on the analysis of the course of vibrations.

The work discusses the problem of classification of vibrations forced by a random series of impulses. The goal was to achieve a correct classification of three groups of distributions characteristic of three different processes: correct, transitory and flawed. Parameters of the oscillator were selected so that random impulses occurred before the vibrations caused by previous random impulses died out. In the analysed problem it was important to minimise the impact of the random variable T defining the intervals between impulses. Time intervals of 300 s were adopted for the analysis. On the basis of exploratory studies confirmed with the help of feature engineering, the beginnings of the examined time series were established. The extracted features of stationary time series and non-stationary ones were used in classification. The features were selected with the use of some methods of feature engineering, including SHAP method, which allows for simultaneous interpretation of the results of the model.

The studies showed that:

- Application of machine learning in the problem under consideration allows for an analysis of vibrating systems in which subsequent random impulses occur before the vibrations forced by previous impulses die out. For a finite time interval such an analysis would not be possible with the use of a mathematical model.
- The stationary character of time series does not influence classification.

- Using the features of the second stochastic moment only we can classify the distributions with the precision of 0.84 after 900 s.
- Using the features of both the second stochastic moment and the third one simultaneously we can classify the distributions with the precision of 0.89 after 600 s.
- In the remaining cases under study the quality of classification is significantly lower.

The conducted analysis indicates that the best results will be obtained when the features of the second stochastic moment and those of the third one are used. On the basis of the features of these two moments we are able fastest and most precisely to distinguish those distributions which are characteristic of different processes. Nevertheless, it is the hardest to distinguish the distributions characterized by the lowest changeability, low standard deviation and the same mean.

The next stage of the discussed investigations will be defining the minimum length of the time window and minimum number of trials, and further - executing of experimental studies.

REFERENCES

[1] P. D. Spanos and B. A. Zeldin, "Generalized random decrement method for analysis of vibration data," *J. Vibrat. Acoust.*, vol. 120, no. 3, pp. 806–813, Jul. 1998.

[2] I. Gullo, M. Di Paola, and P. Spanos, "Spectral approximation for wind induced structural vibration studies," *Meccanica*, vol. 33, pp. 291–298, Jun. 1998.

[3] S. Benfratello, M. Di Paola, and P. D. Spanos, "Stochastic response of MDOF wind-excited structures by means of Volterra series approach," *J. Wind Eng. Ind. Aerodynamics*, vols. 74–76, pp. 1135–1145, Apr. 1998.

[4] K. Sobczyk, *Stochastic Wave Propagation*. Amsterdam, The Netherlands: Elsevier, 1985.

[5] W. Kurnik and A. Tylikowski, "Stochastic stability and nonstability of a linear cylindrical shell," *Ingenieur-Archiv*, vol. 53, no. 6, pp. 363–369, 1983.

[6] J. G. Gilson, "On stochastic theories of quantum mechanics," *Math. Proc. Cambridge Phil. Soc.*, vol. 64, no. 4, pp. 1061–1070, Oct. 1968.

[7] L. de la Peña-Auerbach, "A new formulation of stochastic theory and quantum mechanics," *Phys. Lett. A*, vol. 27, no. 9, pp. 594–595, Sep. 1968.

[8] N. Arley and V. Borchsenius, "On the theory of infinite systems of differential equations and their application to the theory of stochastic processes and the perturbation theory of quantum mechanics," *Acta Mathematica*, vol. 76, nos. 3–4, pp. 261–322, 1945.

[9] A. Tylikowski and W. Marowski, "Vibration of a non-linear single degree of freedom system due to poissonian impulse excitation," *Int. J. Non-Linear Mech.*, vol. 21, no. 3, pp. 229–238, Jan. 1986.

[10] J. B. Roberts, "The response of linear vibratory systems to random impulses," *J. Sound Vibrat.*, vol. 2, no. 4, pp. 375–390, Oct. 1965. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022460X65901161

[11] J. B. Roberts, "On the response of a simple oscillator to random impulses," *J. Sound Vibrat.*, vol. 4, no. 1, pp. 51–61, Jul. 1966. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022460X66901532

[12] J. B. Roberts, "System response to random impulses," *J. Sound Vibrat.*, vol. 24, no. 1, pp. 23–34, Sep. 1972. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022460X72901198

[13] R. Iwankiewicz and K. Sobczyk, "Dynamic response of linear structures to correlated random impulses," *J. Sound Vibrat.*, vol. 86, no. 3, pp. 303–317, Feb. 1983. [Online]. Available: <Go to ISI>://WOS: A1983QB37500001

[14] R. Iwankiewicz, "Dynamic-response of linear structures to random trains of impulses with random spatial shapes," *J. De Mecanique Theorique Et Appliquee*, vol. 4, no. 4, pp. 485–504, 1985. [Online]. Available: <Go to ISI>://WOS: A1985ASP9900004

[15] M. Jabłoński and A. Ozga, "Distribution of stochastic impulses acting on an oscillator as a function of its motion," *Acta Phys. Polonica A*, vol. 118, no. 1, pp. 74–77, Jul. 2010.

[16] R. Iwankiewicz, "Dynamical systems with multiplicative random impulse process excitation," in *Proc. IUTAM Symp. Nonlinear Stochastic Dyn.*, N. S. Namachchivaya and Y. K. Lin, Eds., Dordrecht, The Netherlands: Springer, 2003, pp. 343–352.

[17] M. Jabłoński and A. Ozga, "Determining the distribution of values of stochastic impulses acting on a discrete system in relation to their intensity," *Acta Phys. Polonica A*, vol. 121, no. 1A, pp. A-174–A-178, Jan. 2012.

[18] A. Ozga, "The effect of pulse amplitudes on quality of determining distribution of pulses forcing vibration of an damped oscillator," *Acta Phys. Polonica A*, vol. 128, no. 1A, pp. A-67–A-71, Jul. 2015.

[19] M. Sulewski and A. Ozga, "Application of the forest classifier method for description of movements of an oscillator forced by a stochastic series of impulses," *J. Theor. Appl. Mech.*, vol. 61, pp. 819–831, Oct. 2023.

[20] J. Cao and G. Fan, "Signal classification using random forest with kernels," in *Proc. 6th Adv. Int. Conf. Telecommun.*, May 2010, pp. 191–195.

[21] T. Radivilova, L. Kirichenko, and B. Vitalii, "Comparative analysis of machine learning classification of time series with fractal properties," in *Proc. IEEE 8th Int. Conf. Adv. Optoelectronics Lasers (CAOL)*, Sep. 2019, pp. 557–560.

[22] V. Bulakh, L. Kirichenko, and T. Radivilova, "Time series classification based on fractal properties," in *Proc. IEEE 2nd Int. Conf. Data Stream Mining Process. (DSMP)*, Aug. 2018, pp. 198–201.

[23] M. Ozga and G. Borowski, "The use of granulation to reduce dusting and manage of fine coal," *J. Ecolog. Eng.*, vol. 19, no. 3, pp. 218–224, May 2018.

[24] A. Mikhailov, "Coal-peat compositions for co-combustion in local boilers," *J. Mining Inst.*, vol. 220, pp. 538–544, Apr. 2016.

[25] M. Y. Chernetskiy, A. A. Dekterev, A. P. Burdukov, and K. Hanjalić, "Computational modeling of autothermal combustion of mechanically-activated micronized coal," *Fuel*, vol. 135, pp. 443–458, Nov. 2014.

[26] E. Szatyłowicz and I. Skoczko, "Evaluation of the PAH content in soot from solid fuels combustion in low power boilers," *Energies*, vol. 12, no. 22, p. 4254, Nov. 2019.

[27] M. Jablonski and A. Ozga, "The role of ergodicity in the search for the stochastic distribution of impulses forcing the motion of a linear systems," presented at the 10th Conf. Active Noise Vib. Control Methods MARDiH, Wojanow, Poland, Jun. 2011.

[28] A. Ozga, "Technical note. optimization of parameters for a damped oscillator excited by a sequence of random pulses," *Arch. Acoust.*, vol. 39, no. 4, pp. 645–652, Mar. 2015. [Online]. Available: <Go to ISI>://WOS:000350293000024

[29] P. A. M. Dirac, "On the theory of quantum mechanics," *Proc. Roy. Soc. London. Ser. A, Containing Papers Math. Phys. Character*, vol. 112, no. 762, pp. 661–677, 1926.

[30] H. Nakao, K.-s. Arai, K. Nagai, Y. Tsubo, and Y. Kuramoto, "Synchrony of limit-cycle oscillators induced by random external impulses," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 72, p. 026220, Aug. 2005, doi: 10.1103/PhysRevE.72.026220.

[31] R. Iwankiewicz and S. R. K. Nielsen, "Dynamic response of non-linear systems to Poisson-distributed random impulses," *J. Sound Vibrat.*, vol. 156, no. 3, pp. 407–423, Aug. 1992. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022460X9290736H

[32] W. McKinney, "Data structures for statistical computing in Python," in *Proc. Python Sci. Conf.*, 2010, pp. 56–61, doi: 10.25080/Majora-92bf1922-00a.

[33] T. P. D. Team, "Pandas-dev/pandas: Pandas," Zenodo, Apr. 2024, doi: 10.5281/zenodo.10957263.

[34] J. D. Hunter, "Matplotlib: A 2D graphics environment," *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 90–95, 2007, doi: 10.1109/mcse.2007.55.

[35] S. Seabold and J. Perktold, "Statsmodels: Econometric and statistical modeling with Python," in *Proc. SciPy*, 2010, pp. 1–5.

[36] F. Pedregosa, S. Varoquaux, A. Gramfort, V. Michel, and B. Thirion, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Dec. 2011.

[37] M. Barandas, D. Folgado, L. Fernandes, S. Santos, M. Abreu, P. Bota, H. Liu, T. Schultz, and H. Gamboa, "TSFEL: Time series feature extraction library," *SoftwareX*, vol. 11, Jan. 2020, Art. no. 100456.
[38] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time series FeatuRe extraction on basis of scalable hypothesis tests (tsfresh—A Python package)," *Neurocomputing*, vol. 307, pp. 72–77, Sep. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231218304843
[39] S. M. Lundberg and S.-I. Lee, *A Unified Approach to Interpreting Model Predictions*. Red Hook, NY, USA: Curran Associates, Inc., 2017, pp. 4765–4774. [Online]. Available: http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf

NATALIA FRANKOWSKA received the master's degree from the Faculty of Mechanical Engineering, AGH University of Science and Technology, Kraków. She is currently pursuing the Ph.D. degree in mechanical engineering. Her research interests include machine learning, statistics, and stochastic mechanics.

PRZEMYSŁAW FRANKIEWICZ received the master's degree in mechanical engineering from the AGH University of Krakow. Currently, he is pursuing the Ph.D. degree in stochastic mechanics, with a focus on the application of deep learning. His main research interests include statistical learning, data science, and artificial intelligence. His recent research focuses on designing urban spaces using machine learning algorithms.

AGNIESZKA OZGA received the Master of Applied Science degree specializing in automatic control engineering and robotics from the AGH University of Science and Technology, the master's degree in mathematics from the Faculty of Mathematics and Physics, Jagiellonian University, specializing in information technology, and the Ph.D. degree in applied mechanics. Currently, she is a Researcher and a Professor with the Faculty of Mechanical Engineering and Robotics, AGH University of Krakow. Her research interests include applications of artificial intelligence to various fields of science, including stochastic mechanics.