JOURNAL OF THEORETICAL AND APPLIED MECHANICS
61, 4, pp. 819-831, Warsaw 2023
https://doi.org/10.15632/jtam-pl/172966

APPLICATION OF THE FOREST CLASSIFIER METHOD FOR DESCRIPTION OF MOVEMENTS OF AN OSCILLATOR FORCED BY A STOCHASTIC SERIES OF IMPULSES

Marek Sulewski, Agnieszka Ozga
AGH University of Krakow, Faculty of Mechanic Engineering and Robotics, Krakow, Poland
e-mail: sulewski@agh.edu.pl; aozga@agh.edu.pl

The article discusses the analysis of motion of an oscillator forced by a sequence of stochastic impulses with the use of decision tree algorithms and a random forest classifier. The aim of this paper is to verify the accuracy of distinguishing distributions in the desired time period and to check whether the length of the time interval affects the accuracy of data classification. Moreover, the statistical parameters directly influencing classification of distributions are presented. The analysis has been performed in Python environment, the data were obtained in computer simulation. The results of classification for two classification algorithms with regard to two divisions of the test and training set sizes are presented. In case of the decision tree classifier, it has been observed that for each time interval this algorithm classifies the data achieving a high level of accuracy, but for the purpose of data classification for each time period it selects different statistics, which makes it impossible to unequivocally determine which statistic influences the recognition of distribution. In case of the random forest classification algorithm, the importance and influence of the parameters on the distribution between the three distributions are the same both in 5-minute and 10-minute intervals. The differences between significance of the parameters depending on length of the interval are not significant.

Keywords: machine learning, random forest classifier, stochastic series of impulses

1. Introduction

The theory of random dynamical systems is an interdisciplinary domain (Sobczyk, 1983; Iwankiewicz and Kotulski, 2009; Zembaty, 2009; Socha and Soong, 1991; Liu *et al.*, 2023; Banks *et al.*, 2023). It is applied in mechanics (Litak *et al.*, 2008; Bozzoni *et al.*, 2011; Hračov and Náprstek, 2017; Weber *et al.*, 2021; Smolnicki *et al.*, 2013; Rączka *et al.*, 2013) where, like in the present paper, the idea of research refers to non-deterministic mechanics. The state of a system observed at any given moment does not univocally determine the states of the system at consecutive moments, which issues from the random character of stimulaton. It is worth mentioning that in Poland the studies on random dynamic systems flourished in the third quarter of the 20th century (Piszczek, 1982; Skalmierski and Tylikowski, 1972) when, by means of a complex mathematical apparatus, a series of stochastic equations describing various kinds of dynamic mechanical systems have been determined. Having reached a certain level of knowledge, the studies have not been continued.

The research on systems forced by a random series of impulses also started in the second half of the 20th century (Roberts, 1966, 1972; Roberts and Spanos, 2003) and have been continued ever since. The first attempts at verification of properties of a stochastic model by means of simulation methods were presented by Professor Iwankiewicz (Iwankiewicz, 1993). The research was also conducted by (Mazur-Śniady and Śniady, 1986; Jabłoński and Ozga, 2008) in Poland.

By introducing machine learning into the analysis of dynamic mechanical systems, we are starting the next level of research. In this original approach, thousands of samples have been analyzed for different parameters of random forcing in order to examine usefulness of developed algorithms.

In this study, the developed algorithms are aimed at solving an inverse problem, namely recognition of the distribution of size of impulses forcing vibrations of an oscillator. Although there exists a mathematical model that allows one to calculate the impulse distributions, its applicability, however, is limited. Thus, other solutions of this problem are searched for, and this stage of research is described in this paper.

2. Mathematical model of an oscillator forced by a random series of impulses

The random forest classifier discussed in this paper was carried out for a one-dimensional physical system (Jabłoński and Ozga, 2013) the state of which was described by random variable $x(t)$. The equation of vibrations of an oscillator with damping was presented in a dimensionless form

$$\frac{d^2x}{dt^2} + 2b\frac{dx}{dt} + a^2x = f(t) \tag{2.1}$$

At this stage of research into the possibility of recognizing random distributions of excitation impulses, it is difficult to assess whether the oscillator will be in the form of a mechanical or electronic system. Since the value of $x(t)$ is accepted as dimensionless, consequently, the units occurring with coefficients a, b, and $f(t)$ are referred to time only.

The $f(t)$ is a series of random impulses with random strength η_i occurring at random instants of time t_i

$$f(t) = \sum_{t_i<t} \eta_i \delta(t - t_i) \tag{2.2}$$

where $\delta(t - t_i)$ – Dirac distribution.

The time intervals between impulses $\tau_i = (t_i - t_{i-1})$ are independent continuous random variables for which the function of probability density assumes the form of an exponential function

$$t(\tau) = \begin{cases} \lambda \mathrm{e}^{-\lambda\tau} & \text{for} \quad \tau \geqslant 0 \\ 0 & \text{for} \quad \tau < 0 \end{cases} \tag{2.3}$$

where the constant λ is the impulse occurrence frequency.

For the random variable described by Eq. (2.3), the mean distance between impulses is $1/\lambda$ s, standard deviation also amounts to $1/\lambda$ s, while the median interval between impulses equals $\ln 2/\lambda$.

The values of the η_i impulse are independent discrete random variables with a finite expected value. When intervals between the impulses and strength of the impulses are independent random variables, then the solution of this problem for zero initial conditions is

$$x(t) = \frac{1}{c} \sum_{0<t_i<t} \eta_i \mathrm{e}^{-b(t-t_i)} \sin[c(t - t_i)] \tag{2.4}$$

where $c = \sqrt{a^2 - b^2}$.

The article examines $\overline{m}_i(t)$ – estimators of the k-th stochastic raw moments of the random variable $x(t)$ calculated using the equation

$$\overline{m}_k = \frac{1}{t/h} \sum_{n<t/h} x^k(nh) \tag{2.5}$$

where h is the period of sampling, t is time.

3. Designing of experimental studies using the qualitative method of analysis

Studies on the stochastic model described by Eqs. (2.1)-(2.5) should be appropriately designed so that step changes do not occur in the computed ordinary moments some time after the start. Earlier analyses have shown (Jabłoński and Ozga, 2010, 2012; Ozga, 2019) that oscillators with strong damping should be used, and that impulses should occur frequently enough for values of estimators of ordinary moments calculated from Eq. (2.5) to change to the least extent. The rate of the oscillator own vibrations should be selected appropriately for the impulse occurrence frequency λ. The last step consists in checking whether the value of random impulses is sufficient to force vibrations of the oscillator.

The research is carried out in order to solve the inverse problem, namely to discern the distribution of impulses forcing vibrations of an oscillator at a shortest possible period of time. There is an infinite number of possible cases for which simulation or experimental studies could be developed. Taking into account previous experiences as well as the time necessary to generate one trial, and applying the principle that simple systems will allow for clear presentation of the solution of the research problem, the authors selected an oscillator with damping $b = 10$ and the frequency of vibrations $c = 20$ for $\lambda = 10$ and $h = 10^{-3}$ (Fig. 1). In the presented simulation investigations, vibrations of the oscillator evoke three distributions Φ_i of the pseudo-random variable η_i

1. Φ_1 : $p(\eta_1 = 80) = 0.5$ $p(\eta_2 = 70) = 0.5$
2. Φ_2 : $p(\eta_1 = 130) = 0.5$ $p(\eta_2 = 20) = 0.5$
3. Φ_3 : $p(\eta_1 = 140) = 0.5$ $p(\eta_2 = 10) = 0.5$

The parameters of the distributions were selected so that:

- the expected value of the distributions forcing vibrations in all three cases was the same and amounted to 75,
- the distribution Φ_1 was characterized by two impulses η_1 and η_2 of a similar force influencing the oscillator,
- the distributions Φ_2 and Φ_3 were characterized by two events of different forces of the impact. The value of η_1 symbolizes an impulse of a great force of impact while the value of η_2 an impulse of a little force of impact on the oscillator,
- the differences in statistics (Table 1) of the three discrete random variables are distributed as follows: between Φ_1 and the remaining distributions, the differences are significant. The parameters of the distributions Φ_2 and Φ_3 are similar.

Table 1. Parameters of distributions of the random variables Φ_i

Distribution	Parameter		
	Φ_1	Φ_2	Φ_3
Expected value	75	75	75
Standard deviation	5	55	65
Changeability coefficient	0.0667	0.7383	0.8667
Second raw moment	5 650	8 650	9 850

Figure 1 represents execution of a single movement $x(t)$ and the raw moments computed on the basis of this single execution. Thanks to strong damping occurring in the system at a certain time after the start (see Fig. 1 after 500 s), subsequent impulses do not cause step changes at the computed moments. Distributions were selected so that the mean value was the same in all three cases, hence the estimators of the first ordinary raw moment have similar values after

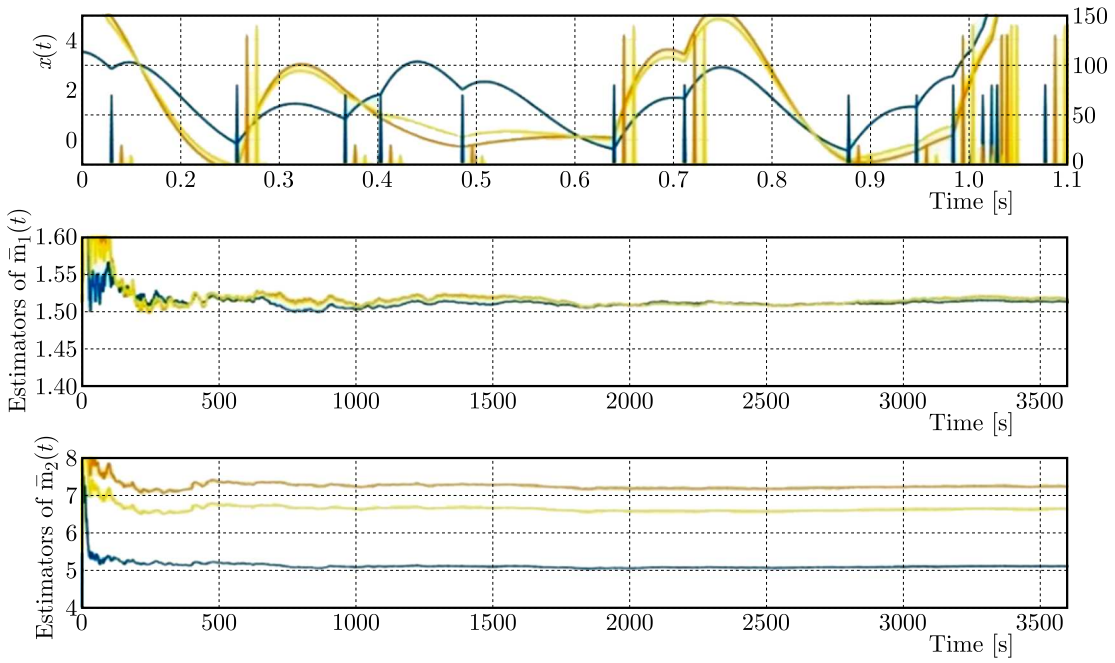

Fig. 1. The movements $x(t)$ of vibrations of an oscillator forced by a random series of impulses for one second; below the first stochastic raw moment and the second one computed on the basis of $x(t)$ for 900 s

1800 s. Stochastic raw moments of the second order (and subsequent orders) are different. It should be emphasized, however, that the presented results were obtained in the simulation that was organized in a specific way. In all three cases, the impulses worked at the same random time, and, what is more, if the strongest impulse was randomly chosen in the first distribution, the impulses selected in the second and third distributions were also the strongest ones. Similarly, the same pertains to the weakest impulses.

It should be taken into account that there exists an infinite number of possible movements, and the possibility of differentiating between the distributions when impulses occur at different random times and have different random values should be checked. In order to visualize this problem, thousand samples were generated for each of the discussed distributions. As it has already been mentioned, it is the estimators of the second raw moments computed on the basis of moments (2.5) that are used in the analysis. Calculations include all values of $x(t)$ since the very beginning till the moment when the values are recorded in the file. Movement (2.4) and moments (2.5) were determined with the sampling frequency of 10^3 s, and the values of moments were saved in the file every second. The analyzed time series of the second order raw moments were presented as tunnels (Fig. 2) covering all recorded samples. The tunnel was created by determining the maximum and minimum values of each of the thousand samples separately for each second. The mean value of all the calculated estimators was also computed.

Visualization of the research in the form of tunnels shows that for the second moment the samples generated for the distribution Φ_1 differ significantly from the others. In further analysis, this distribution will act as a control group, and the research questions will concern the Φ_2 and Φ_3 distributions.

Based on the second diagram, we can also state that we have to do with three time-dependent phenomena. During the start, up to the 600th second, the distributions Φ_2 and Φ_3 are distinguishable in approximately 50% of cases. Between the 600th and 1800th second, the number of samples in which time series have similar statistical parameters decreases. After the 1800th second all three distributions are distinguishable. The time at which the tunnels formed from

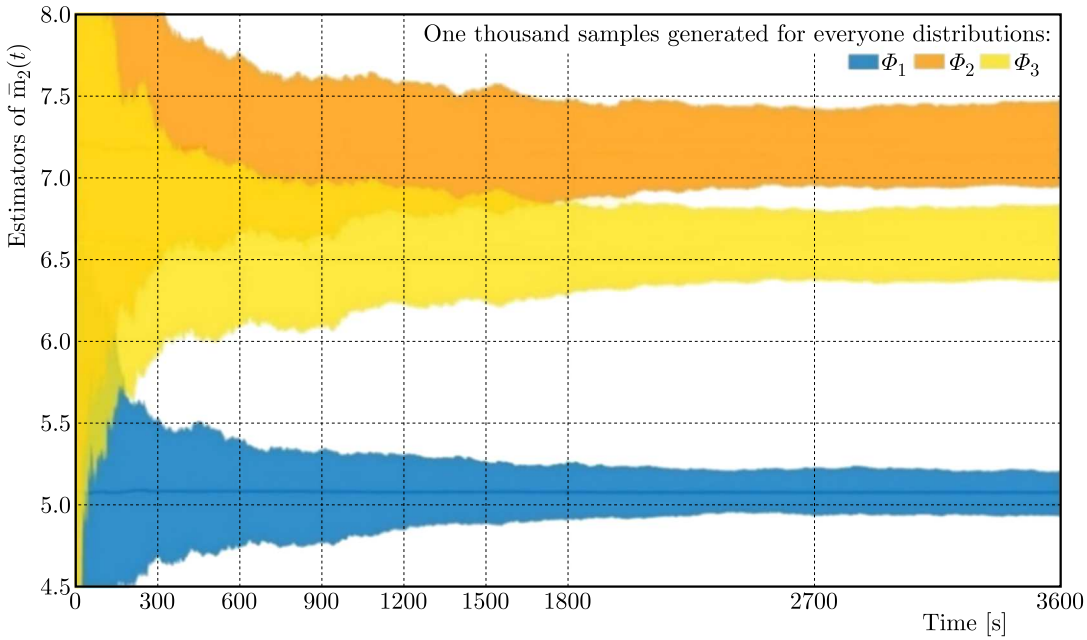

Fig. 2. The second stochastic raw moment calculated from the location $x(t)$ for a thousand different samples presented as tunnels

the generated samples for two similar distributions split up is an approximated value – this is how the samples were randomly distributed. It should be expected, however, that the time series of the subsequent samples will split up within an approximated time interval. These initial exploratory investigations allow for posing two research questions:

1. How precise differentiation of the distributions Φ_2 and Φ_3 before the 1800th second is possible?
2. Does the duration of the analyzed time interval influence the accuracy of classification using the decision tree algorithm and random forest classifier?

In order to answer the research questions posed above, further exploratory investigations should be carried out. It is necessary to check the distributions for single samples in two intervals from 600th second to 2000th one and after 2000th second, presenting the values that occur in the time series in the form of a frequency distribution. The conducted analysis shows that depending on the time interval, statistical parameters describing discrete distributions are different – they differ in the mean value, dominant, variance, etc. All distributions represented in Fig. 3 are multimodal.

On the basis of Fig. 3 it can be assumed that for any several minutes, the long interval starting after the 1800th second, no matter whether we take its mean value, median or dominant, all three distributions will be distinguishable. We can take into consideration either all three statistical parameters or just one of them. Before the 1800th second, the situation is more complicated, hence it seems necessary to use supervised learning algorithms to discern the distributions which force vibrations on the basis of analysis of one sample. Moreover, neither the parameters that could be used for classification of the distribution nor the length of the time interval that should be taken into consideration are known. From the point of view of its application in engineering designs, the sooner we know what distribution we are to deal with, the better. Therefore, nine time intervals presented in Fig. 4 were assigned for further analysis. The intervals are five minutes long while the length of the time intervals with even indices is ten minutes.

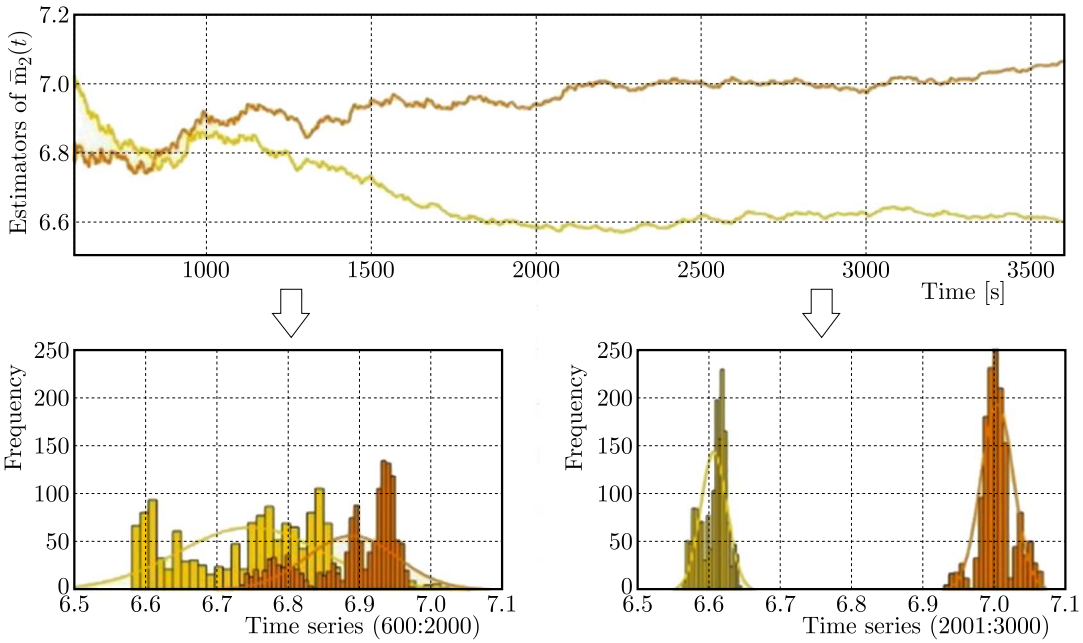

Fig. 3. Distribution of the second raw stochastic moment calculated from the location $x(t)$ for two samples generated from Φ_2 and Φ_3 distributions

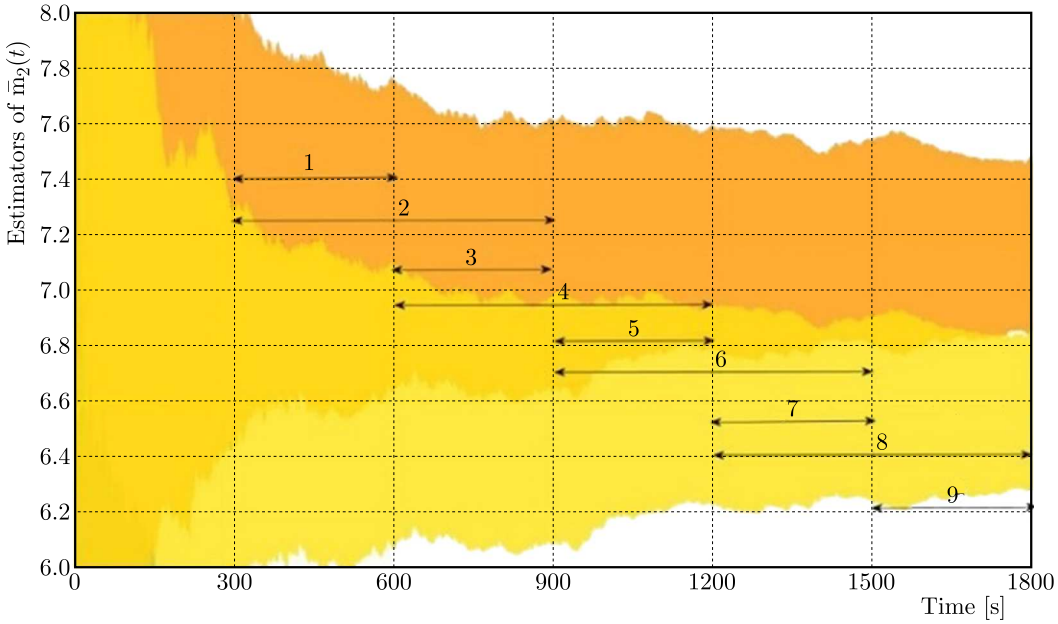

Fig. 4. Time intervals selected for further analysis, represented as tunnels composed from 1000 samples

We will answer the research questions using the following supervised machine learning algorithms. They are the decision tree classifier and the random forest classifier, later described as DTC and RFC, respectively. Along with the indicated time intervals, an analysis was performed using the previously prepared code. At least about 300 impulses were randomly selected during five-minute time intervals, which for the exponential distribution given by Eq. (2.3) allowed one to obtain the λ imposed in the simulations with an accuracy of 0.1%. The tests carried out for shorter values of the time interval indicated the obtaining of less accurate classification results than those presented in the next Section.

4. Data analysis using the decision tree classifier and random forest classifier algorithms

Machine learning (ML) algorithms find patterns, information links in data sets, and then help making decisions and forecasts based on shared data. Out of the logistic regression, elements like naive Bayes, k-nearest neighbors, decision tree, random forest and support vector machine algorithms have been considered. The decision tree classifier and random forest tree algorithms (Szeliga, 2019) have been chosen.

The first one is a decision tree classifier which is a supervised learning algorithm that is dedicated for classification problems. It is a solution which works as a flow chart. It divides the data points into a finite number of categories. This algorithm automatically selects the variables that differentiate the variable the most.

The second one is a random forest algorithm. It is an expansion of the decision tree classifier. First, it builds multiple decision trees based on training data. Then it matches the new data from the test set to one of the trees as a random forest. It averages the data to combine it with the closest random tree on the data scale. The random forest models are useful because they solve the decision tree algorithm problem of unnecessarily forcing data points within a category.

4.1. Determination of the most important statistical parameters describing the time series of the second raw moment estimators

The analysis started with the calculation of basic statistical parameters (Bąk *et al.*, 2020) such as amplitude, one percent above the standard deviation, minimum, mean, and maximum value, maximum slope, percent close to the median, median, median absolute deviation, skewness, standard deviation and weighted mean. Statistical parameters have been defined for the second raw moment, separately for each of the considered samples. Using the same algorithm, the influence of a given parameter on the classification of distributions for two-time intervals was determined (Fig. 5).

Using machine learning methods, it was possible to determine basic statistics that make it possible to recognize the distribution. Figure 5 shows these statistics for nine time periods. In addition, a parallel classification was devised for divisions of the training and test sets in ratios 30/70 and 50/50. The proposed division of the test and training sets results from the typical division in the 30/70 ratio (Szeliga, 2019). One thousand trials is used in the classification. To obtain as many test cases as possible in the area belonging to both tunnels (Fig. 4), the division of 50/50 was also made.

Focusing on Fig. 5, it should be noted that there is a problem with determining the basic statistics that allow one to distinguish the distributions. For odd (300 s) and even (600 s) intervals using the DTC algorithm, each time interval shown in Fig. 5 (marked as TP_x, where x is the next time interval) is described by different statistics. Moreover, it should be observed that the significance of the statistics changes for each time period. For this reason, this algorithm is not suitable for this type of analysis.

A similar verification was conducted for the same intervals with the use of a random tree classifier (Fig. 6).

From this analysis (Fig. 6), it can be seen that in the case of the intervals with a duration of 300 s, the standard deviation (24%), skewness (20%), amplitude (15%) and median absolute deviation (11%) have the highest impact on classification. The minimum value (1%) and the maximum slope (1%) have the lowest impact.

DTC for 30/70									
	TP_1	TP_2	TP_3	TP_4	TP_5	TP_6	TP_7	TP_8	TP_9
Amplitude		0.03		0.01					
Maximum slope		0.45				0.48			
Maximum									
Mean					0.37				
Median	0.34	0.09							
Median absolute deviation	0.65	0.01		0.50		0.02		0.51	
Minimum									
Percent beyond 1 std					0.63				
Percent close to median									
Skew			0.49	0.49		0.48		0.49	
Std		0.42	0.51			0.01	1		1
Weighted average						0.01			

DTC for 50/50									
	TP_1	TP_2	TP_3	TP_4	TP_5	TP_6	TP_7	TP_8	TP_9
Amplitude		0.27	0.51				0.51		0.51
Maximum slope									
Maximum	0.04			0.57					
Mean	0.30								
Median	0.10				0.75				
Median absolute deviation					0.05	0.50			
Minimum	0.03	0.65				0,01			
Percent beyond 1 std		0.07						1	
Percent close to median	0.02								
Skew 0.42		0.49	0.43	0.20	0.49				
Std	0.10						0.49		0.49
Weighted average									

Fig. 5. Representation of basic statistics determined for the second raw moments using DTC algorithms for the indicated time intervals considering division of the harvest in the proportion of 30/70 and 50/50

RFC for 30/70									
	TP_1	TP_2	TP_3	TP_4	TP_5	TP_6	TP_7	TP_8	TP_9
Amplitude	0.15	0.11	0.15	0.11	0.15	0.11	0.15	0.11	0.15
Maximum slope	0.01	0.07	0.01	0.074	0.01	0.07	0.01	0.07	0.01
Maximum	0.09	0.04	0.09	0.04	0.09	0.04	0.09	0.04	0.09
Mean									
Median	0.05	0.07	0.05	0.07	0.05	0.07	0.05	0.07	0.05
Median absolute deviation	0.11	0.15	0.11	0.15	0.11	0.15	0.11	0.15	0.11
Minimum	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
Percent beyond 1 std	0.10	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
Percent close to median									
Skew	0.20	0.29	0.20	0.29	0.20	0.29	0.20	0.29	0.20
Std	0.24	0.19	0.24	0.19	0.24	0.19	0.24	0.19	0.24
Weighted average	0.03		0.03		0.03		0.03		0.03

Fig. 6. Representation of basic statistics for the second raw moments determined using RFC algorithms for the indicated time intervals considering division of the harvest in the proportion of 30/70

In terms of the 600 s time intervals, the largest percentage was for skewness (29%), standard deviation (19%) and median absolute deviation (15%). The smallest share was that of the minimum value (2%). Additionally, in the case of RFCs, the percent close to median and mean

value is not included in the classification. The statistics presented in Fig. 6 are the same for both divisions. The importance and influence of the parameters on the distribution between the three distributions is the same both in the 5-minute and 10-minute intervals. The differences between the significance of the parameters depending on the length of the interval are not significant.

Since in particular time intervals, the statistical parameters responsible for the classification do not change only in the case of random forest algorithms, only these algorithms will be used for the research presented in the following Sections of this article.

4.2. Definition of the best RFC parameters

For the time series generated during simulation tests, the selection of hyperparameters was performed using the *grid_search* function.

Hyperparameters are the parameters that cannot be learned directly from estimators. It is recommended to search the hyperparameter space for the indicated time intervals to obtain the best cross-validation result. It allows one to find the best parameter values for a given estimator and the parameters which can be used to define the classification algorithm. Definition of these parameters reduces the time of the classification process.

In the case of DTC algorithms from the attributes like splitter, max depth, min samples split, min samples leaf, max features, max leaf nodes, class weight, we decided to define max depth, min samples leaf, min samples split and splitter. The max depth attribute defines the maximum depth of the tree. The min samples leaf attribute signifies the minimum number of the samples required to be at a leaf node. This parameter may affect the smoothing of the model. The min samples split determines the minimum number of samples that is required for internal division in the case of classification.

In the case of RFC algorithms, we need to mention the n estimators, criterion, max depth, min samples split, min samples leaf, max features, max leaf nodes, random state, verbose, max samples parameters. For the analysis we choose the following parameters: n estimators, that is the number of trees in the random forest classifier algorithm, and max depth attribute.

4.3. Evaluation of the classification algorithms

In the first step, classification of the downloaded data was conducted and classifier evaluation measures, such as precision, recall, f1-score and support, were determined.

The precision measure is responsible for recognizing a class, for example which part positively predicted elements are of all those marked as a part of this register.

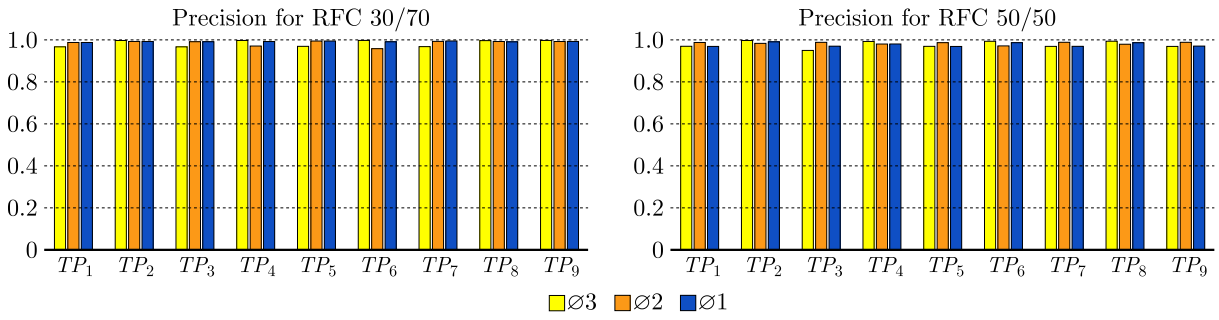

Fig. 7. Precision parameter for RFC for two splits

Figure 7 shows the precision parameter for nine-time intervals for two divisions 30/70 and 50/50 (test set/training set). It should be noted how this parameter changes depending on the recognized distribution. In the case of division, where there is 70% of the training set, we can see that for all time intervals the precision is 100% for the first distribution, which in the case

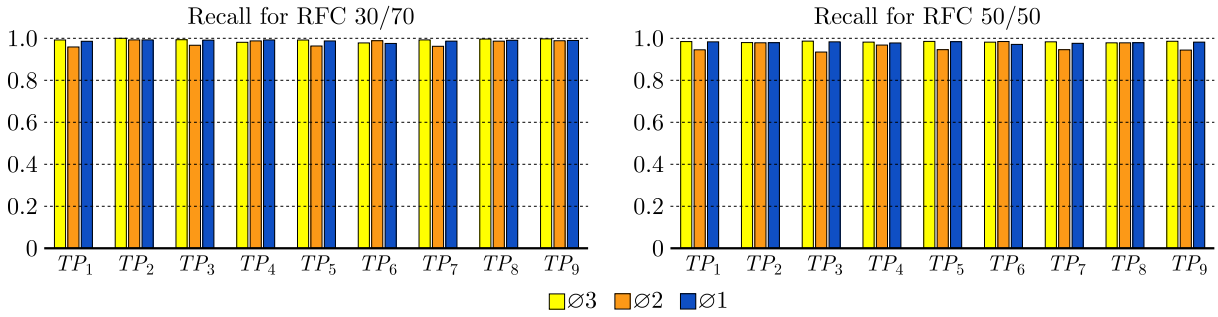

Fig. 8. Recall parameter for RFC for two splits

of these specifically analyzed data may be a determinant of proper operation of the algorithm (Fig. 8). There are slight differences when the test set is 50%. The level is between 98% and 100%, which means that the test set in the case of these variables should have more than five hundred samples. Expected difficulties exist with the second and third distributions. The precision in the case of the 30/70 division only for the 8th and 9th time interval is flawless for all three distributions. In the case of the 50/50 division, it was not possible to obtain 100% precision in the rare of the analyzed time intervals.

The next parameter was recall. This parameter called in some publications sensitivity or true positive rate informs how many elements from a given class have been correctly recognized.

The recall on the tested splits was also on high level. We can see that in the case of the 30/70 split, the first and third to seventh time periods did not reach the 100% recall level for the three distributions. Only for the eighth and ninth time periods was the 100% recall level reached.

For the 50/50 split, no time intervals except the eighth one, achieved the 100% recall level.

The last parameter is an F_1-score determined according to equation

$$F_1 = 2\frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{4.1}$$

In statistical rating of classification, it is a harmonic mean of the precision and recall. Figure 9 presents the F_1-score for nine time periods for 30/70 and 50/50 splits.

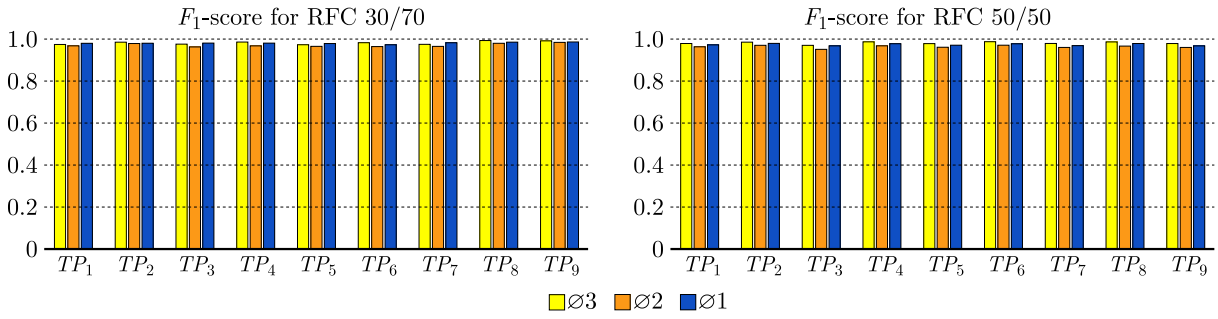

Fig. 9. F_1-score parameter for RFC for two splits

All these tree parameters allow us to evaluate the RFC for the mentioned time periods.

4.4. Summary of the classification algorithms in nine time periods

Figure 10 shows the accuracy of the RFCs on the test and training sets when their split is 30/70 and 50/50, respectively, for the three distributions. It should be emphasized that for this analysis, the search for the best hyper parameters was conducted in order to classify the appropriate class in the best possible way. The accuracy of classification using RFCs on training sets for all time intervals and both division is 100%.

RFC for 30/70									
	TP_1	TP_2	TP_3	TP_4	TP_5	TP_6	TP_7	TP_8	TP_9
Accuracy using DTC on training set	1	1	1	1	1	1	1	1	1
Accuracy using DTC on test set	0.989	1	0.989	0.994	0.989	0.989	0.989	1	1

RFC for 50/50									
	TP_1	TP_2	TP_3	TP_4	TP_5	TP_6	TP_7	TP_8	TP_9
Accuracy using DTC on training set	1	1	1	1	1	1	1	1	1
Accuracy using DTC on test set	0.987	0.987	0.980	0.999	0.987	0.993	0.987	0.997	0.987

Fig. 10. Representation of accuracy on training and the test set for second raw moments using RFC algorithms for the indicated time intervals considering division of the harvest in the proportion of 30/70 and 50/50

Accuracy for 30% of the test set ranged from 98.9% to 100%. For a 50% of the test set, the accuracy ranged from 98% to 100%.

5. Conclusion

After the analysis of the distributions of the second raw stochastic moments for the indicated nine time intervals, using the algorithm of random decision trees and the random forest classifier, the second algorithm was selected for further work based on the achieved results. It can be concluded that the algorithm RFC allows for recognition of distributions and determination of statistical parameters, which have the highest impact on distinguishing distributions.

When answering the question how it is possible to precisely differentiate the distributions Φ_2 and Φ_3 before 1800 s, it should be noted that high precision was achieved in fitting the data to a given distribution. For the RFC classifier, for the case where a 30% test set was considered, for 5-minute intervals an average of 100% for Φ_1, 100% for Φ_2 and 97% for Φ_3 was achieved. In the case of 10-minute intervals, 100% was achieved for Φ_1 and Φ_3 and 98.5% for Φ_2. Thus, the differences in the precision of classification depending on the length of the intervals are not significant.

Based on the results obtained for the decision trees algorithm, it should be stated that it is not suitable for the analysis of this type of technical problems because despite the positive results of the classifier operation, the share of individual statistics and their significance changes with each time interval.

Acknowledgement

The article was financed by a research subsidy number: 16.16.130.942 and project No 10.16.130.79990.

References

1. Banks D.M., Bekker J.C., Vermeulen H.J., 2023, Parameter estimation of a distribution transformer model using pseudo-random impulse sequence perturbation, *31st Southern African Universities Power Engineering Conference, SAUPEC*, 1-6
2. Bąk I., Markowicz I., Mojsiewicz M. Wawrzyniak K., 2020, Mathematical statistics. Examples and problems (in Polish), *Statystyka matematyczna. Przykłady i zadania*, CeDeWu Sp. z o.o., Warszawa
3. Bozzoni F., Corigliano M., Lai C.G., Salazar W., Scandella L., Zuccolo E., Latchman J., Lynch L., Robertson R., 2011, Probabilistic seismic hazard assessment at the Eastern Caribbean Islands, *Bulletin of the Seismological Society of America*, **111**, 5

4. Fenner M.E., 2020, *Machine Learning with Python for Everyone*, Addison Wisley Data

5. Hračov S., Náprstek J., 2017, Approximate complex eigensolution of proportionally damped linear systems supplemented with a passive damper, *Procedia Engineering*, **199**, 1677-1682

6. Iwankiewicz R., 1993, Stochastic methods in problems of dynamical systems subjected to a random series of impulses (in Polish), [In:] *Metody Stochastyczne w Zagadnieniach Układów Dynamicznych Poddanych Losowym Seriom Impulsów*, Prace Naukowe Instytutu Materiałoznawstwa i Mechaniki Technicznej Politechniki Wrocławskiej, Wrocław

7. Iwankiewicz R., Kotulski Z., 2009, *Stochastic Methods in Mechanics: Status and Challenges*, IPPT PAN, Warszawa

8. Jabłoński M., Ozga A., 2008, Statistical characteristics of vibrations of a string forced by stochastic forces, *Mechanics*, **27**, 1, 1-7

9. Jabłoński M., Ozga A., 2010, Distribution of stochastic impulses acting on an oscillator as a function of its motion, *Acta Physica Polonica A*, **118**, 74-77

10. Jabłoński M., Ozga A., 2012, Determining the distribution of values of stochastic impulses acting on a discrete system in relation to their intensity, *Acta Physica Polonica A*, **121**, A174-A178

11. Jabłoński M., Ozga A., 2013, *Distribution of Random Pulses Acting on a Vibrating System as a Function of its Motion*, AGH University Press, Krakow

12. Litak G., Borowiec M., Friswell M. I., Szabelski K., 2008, Chaotic vibration of a quarter-car model excited by the road surface profile, *Communications in Nonlinear Science and Numerical Simulation*, **13**, 7, 1373-1383

13. Liu Y., Xu J., Lu J., Gui W., 2023, Stability of stochastic time-delay systems involving delayed impulses, *Automatica*, **152**, 110955

14. Mazur-Śniady K., Śniady P., 1986, Dynamic response of linear structures to random streams of arbitrary impulses in time and space, *Journal of Sound and Vibration*, **110**, 1, 59-68

15. Ozga A., 2019, Determining the distribution of stochastic impulses in linear discrete dynamical systems (in Polish), AGH University Press, 130-136, ISBN: 978-83-7464-997-1

16. Piszczek K., 1982, *Stochastic Methods in the Theory of Mechanical Vibrations* (in Polish), PWN, Warszawa

17. Rączka W., Konieczny J., Sibielak M., 2013, Laboratory tests of shape memory alloy wires, *Solid State Phenomena*, **199**, 365-370

18. Roberts J.B., 1966, On the response of a simple oscillator to random impulses, *Journal of Sound and Vibration*, **4**, 1, 51-61

19. Roberts J.B., 1972, System response to random impulses, *Journal of Sound and Vibration*, **24**, 1, 23-34

20. Roberts J.B., Spanos P.D., 2003, *Random Vibration and Statistical Linearization*, New York Dover Publication

21. Skalmierski B., Tylikowski A., 1972, *Stochastic Processes in Dynamics* (in Polish), PWN, Warszawa

22. Smolnicki T., Stańco M., Pietrusiak D., 2013, Distribution of loads in the large size bearing – problems of identification, *Tehnički Vjesnik*, **20**, 5, 831-836

23. Sobczyk K., 1983, Stochastic methods in mechanics: state and development trends (in Polish), *Mechanika Teoretyczna i Stosowana*, **21**, 4, 557-564

24. Socha L., Soong T.T., 1991, Linearization in analysis of nonlinear stochastic systems, *Applied Mechanics Reviews*, **44**, 10

25. Szeliga M., 2019, *Practical Machine Learning* (in Polish), PWN SA, Warszawa

26. Tylikowski A., 1982, Vibrations of a harmonic oscillator caused by a series of random collisions (in Polish), *Prace Instytutu Podstaw Budowy Maszyn PW*, **13**, 101-112

27. Weber H., Kaczmarczyk S., Iwankiewicz R., 2021, Non-linear response of cable-mass-spring system in high-rise buildings under stochastic seismic cxcitation, *Materials*, **14**, 22, 6858

28. Zembaty Z., 2009, Tutorial on surface rotations from wave passage effects: stochastic spectral approach, *Bulletin of the Seismological Society of America*, **99**, 2B, 1040-1049

Manuscript received October 27, 2022; accepted for print August 25, 2023