# Intelligent information systems

UDC 004.9 **doi:** https://doi.org/10.20998/2522-9052.2024.2.09

Oleh Pihnastyi[1], Georgii Kozhevnikov[1], Anna Burduk[2]

[1] National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine
[2] Wroclaw University of Science and Technology, Wroclaw, Poland

## METHOD FOR GENERATING A DATA SET FOR TRAINING A NEURAL NETWORK IN A TRANSPORT CONVEYOR MODEL

**Abstract. The object** of research is a stochastic input flow of material coming in the input of a conveyor-type transport system. Subject of research is the development of a method for generating values of the stochastic input material flow of transport conveyor to form a training data set for neural network models of the transport conveyor. **The goal** of the research is to develop a method for generating random values to construct implementations of the input material flow of a transport conveyor that have specified statistical characteristics calculated based on the results of previously performed experimental measurements. The article proposes a method for generating a data set for training a neural network for a model of a branched, extended transport conveyor. **A method** has been developed for constructing implementations of the stochastic input flow of material of a transport conveyor. Dimensionless parameters are introduced to determine similarity criteria for input material flows. The stochastic input material flow is presented as a series expansion in coordinate functions. To form statistical characteristics, a material flow implementation based on the results of experimental measurements is used. As a zero approximation for expansion coefficients, that are random variables, the normal distribution law of a random variable is used. **Conclusion.** It is shown that with an increase in the time interval for the implementation of the input material flow, the correlation function of the generated implementation steadily tends to the theoretically determined correlation function. The length of the time interval for the generated implementation of the input material flow was estimated.

**Keywords:** transport conveyor; neural network; stochastic process; dataset generator; material flow.

### Introduction

Conveyor-type transport systems service a very important role in reducing the cost of material extraction in the mining industry [1,2]. The conveyor is not only the most economical way to move bulk materials, but practically the only means for moving material over difficult terrain [3]. Reducing the unit cost of transporting material is achieved by increasing the material loading factor of the transport system. For the mining industry, the value of the load factor for the standard operating mode of the transport system is in the acceptable range of 0.5-0.7 [4]. At the same time, the share of costs for moving material in transport systems several kilometers long reaches 20% [5, 6]. As the uneven loading of the transport system with material increases, the load factor decreases, which leads to a nonlinear multiple increase in transport costs [7, 8]. To reduce transport costs, systems are developed to control the speed of the conveyor belt [9, 10], to control the flow of material from the accumulating bunker [11–13], combined control systems [14]. Cost reduction methods using energy management methodology are also proposed [15]. At the early stage of designing systems for controlling flow parameters, models were mainly used, the calculation of parameters of which is based on numerical methods: finite element method (FEM) [16], finite difference method (FDM), system dynamics equations [17]. Models based on numerical methods made it possible to synthesize algorithms for controlling the flow parameters of a single conveyor.

The emergence of an analytical model of a conveyor [18] made it possible to design control systems for branched transport systems consisting of a dozen conveyors [13]. The next stage in the development of models, as well as control systems for the flow parameters of a transport conveyor, was the construction of transport system models based on the use of regression equations [19, 20] and neural networks [21–23]. Models using neural networks opened up the possibility of modelling extended and highly branched transport systems containing hundreds of individual conveyors. The characteristics of extended conveyor systems, the length of which reaches one hundred kilometres, are given in [24]. An analysis of the functioning of highly branched conveyor-type transport systems is presented in [25, 26]. In parallel with the development of models based on neural networks, analytical models for extended transport systems consisting of a large number of sequentially located conveyors were improved [27]. Representing the state of flow parameters along the transport route in the form of a series using Heaviside functions made it possible to obtain compact analytical expressions for modelling the state of flow parameters of conveyors sequentially located along the transport route. This has opened up new opportunities for aggregating extensive conveyor-type transport systems.

However, a promising direction for modeling highly branched transport systems remains the direction based on the use of neural networks. A model of a branched transport system based on a neural network is presented in [28]. When training a neural network, a data set was traditionally used that was built for a deterministic input material flow. This is due to the difficulty of generating data sets for training neural networks in transport conveyor models, which is the main obstacle to the use of these models for the design of highly efficient control systems. To solve the problem of generating a training data set, the results of experimental measurements of the input material flow can be used [29, 30]. Currently, there are already publications containing an analysis of the implementation of a stochastic flow of material coming in the input of a transport conveyor [31, 32]. Typically, a single study contains only one realization of stochastic material flow, that is not sufficient to generate a training dataset. On the other hand, the duration of experimental measurements should exceed the average time of material transportation along the route in order to exclude the influence of the initial distribution of the material flow along the route on the value of the output material flow. For extended transport systems, transportation time reaches several hours. This makes it difficult to directly use experimental measurements of the input material flow to construct a training dataset. In this situation, designing a generator of values for the stochastic flow of material coming in the input of the transport system will solve this problem. Such a generator, based on the values of the statistical characteristics of the experimental flow of a material, will ensure that the neural network is trained to generate the required number of realizations of the stochastic material flow with a given time duration. This will make it possible to build an effective neural network training process and improve both the modeling accuracy and the accuracy of material flow control in an industrial environment. This paper explores the prospects for developing a generator values of stochastic material flow based on functional relationships between the statistical characteristics of implementations input material flow obtained from experimental measurements.

## Problem statement

With a limited set of sample data specified by a single implementation of a random stationary process on time interval $t \in [t_{\min}, t_{\max}]$, time averaging for stationary process $\lambda(t)$ can be replaced by averaging over the statistical population

$$m_\lambda = \frac{1}{t_{\max} - t_{\min}} \int_{t_{\min}}^{t_{\max}} \lambda(t) dt; \tag{1}$$

$$\sigma_\lambda{}^2 = \frac{1}{t_{\max} - t_{\min}} \int_{t_{\min}}^{t_{\max}} (\lambda(t) - m_\lambda)^2 dt; \tag{2}$$

$$\begin{gathered} k_\lambda(\eta) = 1/(t_{\max} - t_{\min}) \times \\ \times \int_{t_{\min}}^{t_{\max}} (\lambda(t) - m_\lambda)(\lambda(t+\eta) - m_\lambda) dt; \\ k_\lambda(\eta) = k_\lambda(-\eta), \end{gathered} \tag{3}$$

$$\begin{gathered} \lambda_{\max} = \max(\lambda(t)); \\ \lambda_{\min} = \min(\lambda(t)); \\ t_{\max} = \max(t);\ t_{\min} = \min(t);\ t \in [t_{\min}, t_{\max}], \end{gathered} \tag{4}$$

where $m_\lambda$, $\sigma_\lambda$ is the mathematical expectation and standard deviation for random process $\lambda(t)$, calculated based on its implementation; $\lambda_{\max}$, $\lambda_{\min}$ are maximum and minimum value of the implementation of a random process $\lambda(t)$; $k_\lambda(\eta)$ is correlation function of a stationary process; $\eta$ is correlation time. A sufficient condition for the fulfillment of equalities (1) – (3) is the limit equality:

$$\lim_{\eta \to \infty} k_\lambda(\eta) \to 0. \tag{5}$$

Let us define the input flow of material in the form of a random stationary process $\gamma(\tau)$, that determines the value of the input flow of material $\gamma$ at an arbitrary moment of dimensionless time $\tau$. Let us introduce dimensionless parameters to describe the material flow arriving on the entrance of the transport conveyor:

$$\begin{gathered} \gamma(\tau) = \frac{\lambda(t)}{m_\lambda}; \quad \vartheta = \frac{2\eta}{t_{\max} - t_{\min}}; \\ \tau = 2\frac{t - t_{\min}}{t_{\max} - t_{\min}} - 1; \quad \tau \in [-1, 1]; \end{gathered} \tag{6}$$

$$m = 1; \quad \sigma = \frac{\sigma_\lambda}{m_\lambda}; \quad k_s(\vartheta) = k_s(-\vartheta). \tag{7}$$

and present the statistical characteristics of the parameters of the material flow in dimensionless form:

$$\tau_n = \frac{2n}{N} - 1; \quad n = 0, N; \quad i = 0, \frac{N}{2}; \tag{8}$$

$$\begin{gathered} m = \frac{1}{2}\int_{-1}^{1} \gamma(\tau) d\tau = \frac{1}{N+1}\sum_{n=0}^{N} \gamma(\tau_n); \\ \sigma^2 = \frac{1}{2}\int_{-1}^{1} (\gamma(\tau) - m)^2 d\tau = \frac{1}{N+1}\sum_{n=0}^{N} (\gamma(\tau_n) - m)^2; \end{gathered} \tag{9}$$

$$\begin{gathered} k_s(\vartheta_i) = \frac{1}{2}\int_{-1}^{1} (\gamma(\tau) - m)(\gamma(\tau + \vartheta_i) - m) d\tau = \frac{2}{N+1} \times \\ \times \sum_{n=N/2}^{N} (\gamma(\tau_n) - m)(\gamma(\tau_n - \vartheta_i) - m), \quad \vartheta_i = 2\frac{i}{N}. \end{gathered} \tag{10}$$

Let us replace the implementation of stationary process $\gamma(\tau)$ with the implementation of a random process, represented by a sequence of random values of the material flow $\Theta_n$, constant in the value during random time intervals $T_n$. Random variables $\Theta$ and $T$ have a given distribution law. For the presented approximation of random process $\gamma(\tau)$, it is necessary to construct a random flow generator with statistical characteristics (1)–(3) that define random process $\gamma(\tau)$.

## Construction of a generator of input material flow values to form a training data set

Let us define the input flow of material $\gamma(\tau)$ in the form of decomposition:

$$\gamma(\tau) = m(\tau) + \sum_{n=0}^{\infty} \Theta_n \rho_n(\tau); \quad m(\tau) = m = const; \quad (11)$$

$$\rho_n(\tau) = H(\tau_n - \tau) - H(\tau_{n-1} - \tau);$$

$$H(S) = \begin{cases} 0, & S < 0; \\ 1, & S \ge 0; \end{cases} \quad \tau_n = \sum_{k=0}^{n} T_k, \quad (12)$$

where $\Theta_n$ is independent centered random variables with standard deviation $\sigma_{\Theta n}$; $m(\tau)$ is mathematical expectation of random process $\gamma(\tau)$; $T_n$ is independent random variables with standard deviation $\sigma_{Tn}$ and mathematical expectation $m_{Tn}$; $H(x)$ is Heaviside function.

Functions $\rho_n(\tau)$ are orthogonal functions on interval $\tau \in [\tau_{n-1}, \tau_n]$

$$\frac{1}{\tau_n - \tau_{n-1}} \int_{\tau_{n-1}}^{\tau_n} \rho_n(\tau)\rho_k(\tau) d\tau = \begin{cases} 1 & if \ n = k; \\ 0 & if \ n \ne k. \end{cases} \quad (13)$$

The sequence of values of random variables $(\Theta_n, T_n)$ form the implementation of the random process $\gamma(\tau)$:

$$\gamma(\tau) = m + H(\tau_0 - \tau)\Theta_0 + ... = \Theta_0; \quad 0 \le \tau < \tau_0;$$

$$\gamma(\tau) = m + H(\tau_0 - \tau)\Theta_0 + \\ +\left(H(\tau_1 - \tau) - H(\tau_0 - \tau)\right)\Theta_1 + ... = \Theta_1; \quad (14) \\ \tau_0 \le \tau < \tau_1;$$

$$\gamma(\tau) = m + H(\tau_0 - \tau)\Theta_0 + ... + \\ +\left(H(\tau_n - \tau) - H(\tau_{n-1} - \tau)\right)\Theta_n + ... = \Theta_n; \\ \tau_{n-1} \le \tau < \tau_n.$$

Along the dimensionless time axis $[0, \tau]$ there is a simple flow of events $T(\tau)$ with intensity $1/m_{Tn}$. At the moment of occurrence of event $T_n$, random process $\Theta(\tau)$ takes on a random value $\Theta_n$, keeping it constant until the occurrence of the next event $T_{n+1}$ (12). The centered random variables $\Theta_0$, $\Theta_1$, $\Theta_2, \ldots, \Theta_n$ are independent and distributed with the same density $f_\Theta(\Theta)$ and standard deviation $\sigma_{\Theta n} = \sigma_\Theta$:

$$1 = \int_{\lambda_{\min} - m}^{\lambda_{\max} - m} f_\Theta(\Theta) d\Theta; \quad M\left[\Theta_n^2\right] = \sigma_{\Theta n}^2 = \sigma_\Theta^2. \quad (15)$$

Likewise, random variables $T_0$, $T_1$, $T_2, \ldots T_n$ are independent and distributed with equal density. Let us assume in the zero approximation that the probability of occurrence of event $T_n$ has an exponential law:

$$f_T(T) = (1/m_T) \cdot \exp(-T/m_T);$$

$$1 = \int_0^{\infty} (1/m_T) \cdot \exp(-T/m_T) dT; \quad m_T = M[T]. \quad (16)$$

Let us determine the characteristics of the stochastic material flow $\gamma(\tau)$. Since the time dependence is concentrated in the deterministic function $\rho_n(\tau)$, and the random behavior in the random variable $\Theta_n$, it follows:

$$M\left[\gamma(\tau)\right] = M\left[m(\tau) + \sum_{n=0}^{\infty} \Theta_n \rho_n(\tau)\right] = m + \\ + \sum_{n=0}^{\infty} M\left[\Theta_n \rho_n(\tau)\right] = m + \sum_{n=0}^{\infty} \rho_n(\tau) M\left[\Theta_n\right] = m; \quad (17)$$

$$D\left[\gamma(\tau)\right] = \\ = M\left[\left(\sum_{n=0}^{\infty} \Theta_n \rho_n(\tau)\right)^2\right] = \sum_{n=0}^{\infty} \rho_n^2(\tau) M\left[\Theta_n^2\right] = \quad (18) \\ = \sum_{n=0}^{\infty} \rho_n^2(\tau) \sigma_{\Theta n}^2 = \sum_{n=0}^{\infty} \sigma_{\Theta n}^2 = \sigma_\Theta^2.$$

If at time $\tau$ event $\Theta_n$ is observed, this means that event $T_n \ge (\tau - \tau_n)$ has occurred. Let us find the probability that at time $\tau$ event $\Theta_n$ is observed and event $T_n \ge (\tau + \vartheta - \tau_n)$ will occur. Let's use conditional probability

$$P\left(T_n \ge (\tau + \vartheta - \tau_n) \mid T_n \ge (\tau - \tau_n)\right) = \\ = P\left(T_n \ge (\tau + \vartheta - \tau_n)\right) / P\left(T_n \ge (\tau - \tau_n)\right) = \quad (19) \\ = \exp\left(-\frac{\tau + \vartheta - \tau_n}{m_T}\right) \Big/ \exp\left(-\frac{\tau - \tau_n}{m_T}\right) = \exp\left(-\frac{\vartheta}{m_T}\right).$$

Thus, the probability that at time $\tau$ event $\Theta_n$ is observed and event $T_n \ge (\tau + \vartheta - \tau_n)$ occurs, provided that event $T_n \ge (\tau - \tau_n)$ occurs, is $\exp(-\vartheta / m_T)$. Let us use this result to calculate the correlation function for a random stationary process $\gamma(\tau)$:

$$k(\vartheta) = M\left[\left(\sum_{n=0}^{\infty} \Theta_n \rho_n(\tau)\right)\left(\sum_{i=0}^{\infty} \Theta_i \rho_i(\tau + \vartheta)\right)\right] = \\ = M\left[\sum_{n=0}^{\infty}\sum_{i=0}^{\infty} \Theta_i \Theta_n \rho_i(\tau + \vartheta)\rho_n(\tau)\right] = \\ = M\left[\sum_{n=0}^{\infty} \Theta_n^2 \rho_n(\tau + \vartheta)\rho_n(\tau)\right] = \quad (20) \\ = \sum_{n=0}^{\infty} M\left[\Theta_n^2 \rho_n(\tau + \vartheta)\rho_n(\tau)\right] = \\ = \sum_{n=0}^{\infty} M\left[\Theta_n^2\right] M\left[\rho_n(\tau + \vartheta)\rho_n(\tau)\right] = \sigma_\Theta^2 \exp\left(-\frac{\vartheta}{m_T}\right),$$

where $M[\Theta_i \Theta_n] = 0$, due to the fact that the centered random variables $\Theta_n$ are independent centered random variables; $M\left[\Theta_n^2\right] = \sigma_\Theta^2$ by definition (15); $\rho_n(\tau) = 0$ if $\tau \notin [\tau_{n-1}, \tau_n]$ by definition of function $\rho_n(\tau)$ (12).

The correlation function is proportional to $\sigma_\Theta^2$, since for an arbitrary moment in time $\tau \in [\tau_{n-1}, \tau_n]$ the value of the material flow is constant, and the dispersion for the distribution law for each random variable $\Theta_0, \Theta_1, \Theta_2, \ldots, \Theta_n$ is equal to $\sigma_\Theta^2$, $M[\Theta_n^2] = \sigma_\Theta^2$.

Expression (11) is the canonical expansion of the stochastic process $\gamma(\tau)$ into coordinate functions $\rho_n(\tau)$. The expansion coefficients are centered random variables $\Theta_n$. The coordinate functions $\rho_n(\tau)$ are orthogonal on the interval $\tau \in [\tau_{n-1}, \tau_n]$. The type of coordinate functions is chosen from the condition of optimal approximation of the implementation of the input material flow. This choice of a coordinate function can correspond to the method of extracting material at different speeds of moving of equipment at different time intervals. The canonical expansion can have different laws of distribution of random variables $\Theta_n$. Thus, a stochastic flow of material can be represented by different versions of the law of distribution of random variables $\Theta_n$. The equality of the mathematical expectation, standard deviation and correlation function for two processes does not mean the equality of the distribution laws of two random processes. As a zero approximation, we will assume that random variables $\Theta_n$ have a normal distribution law.

Taking into account that $3\sigma_\Theta < m$ [8], we rewrite expression (15) in the form

$$\int_{\lambda_{\min}-m}^{\lambda_{\max}-m} f_\Theta(\Theta) d\Theta \approx$$

$$\approx \int_{\lambda_{\min}-m}^{\lambda_{\max}-m} \frac{1}{\sigma_\Theta\sqrt{2\pi}} \exp\left(-\frac{\Theta}{2\sigma_\Theta^2}\right) d\Theta \approx . \qquad (21)$$

$$\approx \int_{-\infty}^{\infty} \frac{1}{\sigma_\Theta\sqrt{2\pi}} \exp\left(-\frac{\Theta}{2\sigma_\Theta^2}\right) d\Theta = 1.$$

The probabilistic characteristics of a stationary stochastic process do not depend on time. Thus, the one-dimensional distribution density of the values of the stochastic input material flow (4) does not depend from time, and the mathematical expectation and dispersion of the stochastic material flow are constant values. Despite the fact that random process $\gamma(\tau)$ is represented by a composition of a set of random variables 2, the law of distribution of values of random process $\gamma(\tau)$ with selected coordinate functions $\rho_n(\tau)$ is not a normal distribution law. This is explained by the fact that for different sections of the random process $\gamma(\tau)$ and $\gamma(\tau + \vartheta)$ the probability of their coincidence is not zero, but is determined by the expression:

$$P\{\gamma(\tau) = \gamma(\tau + \vartheta)\} = \exp\left(-\frac{\vartheta}{m_T}\right). \qquad (22)$$

On the other hand, for two random variables distributed according to the normal distribution law, the probability of matching the values is zero. It follows that when approximated in the form of canonical expansion (11), the distribution law of material flow values is not a normal distribution law.

## Analysis of results

As an example, consider the construction of a generator of input material flow values to simulate the material flow incoming to the input of a transport conveyor (NCC Industry, Uddevalli, Sweden), Fig. 1 [29].

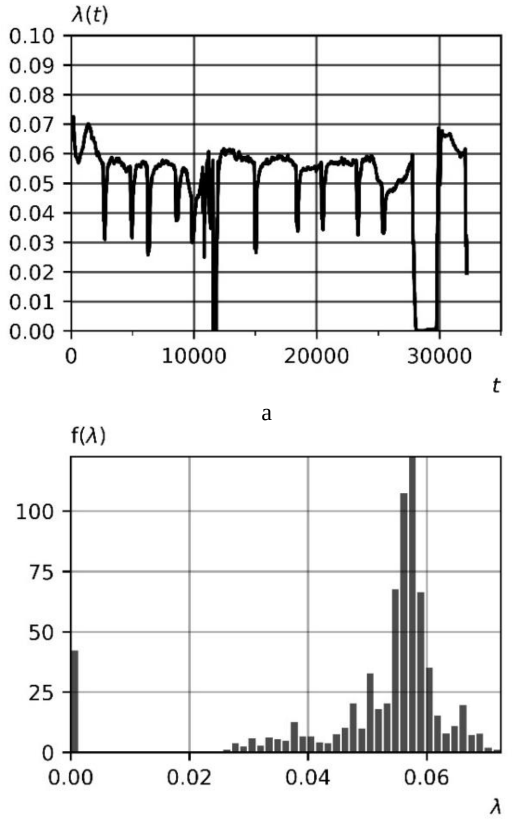
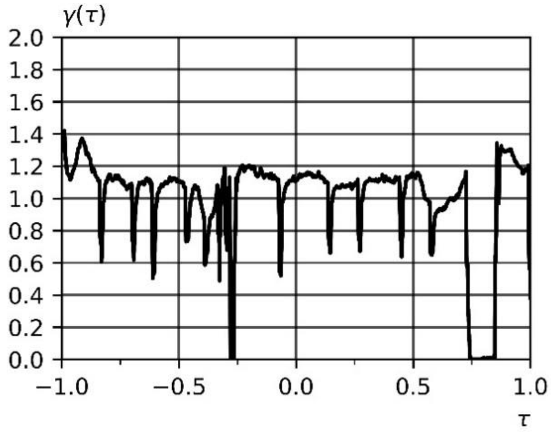



**Fig. 1**. The input material flow $\lambda(t)$ (NCC Industry, Uddevalli, Sweden, [29]): a – implementation of the input material flow; b – histogram of the distribution of values $\lambda$ of the input material flow

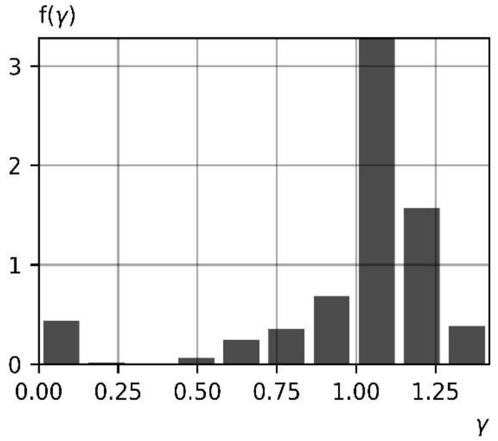
Taking into account dimensionless parameters (6), (7), let's present the input flow of material $\lambda(t)$ in dimensionless form $\gamma(\tau)$, Fig. 2.

The implementation of a dimensionless material flow $\gamma(\tau)$ contains minimums, which characterize the process of supplying material in batches with a quasi-constant value of the input material flow. Let us define the input flow of material within time interval $\tau \in [\tau_{n-1}, \tau_n]$ in the form of the following approximation:

$$\gamma_n = \frac{1}{\tau_n - \tau_{n-1}} \int_{\tau_{n-1}}^{\tau_n} \gamma(\tau) d\tau, \qquad (23)$$




a

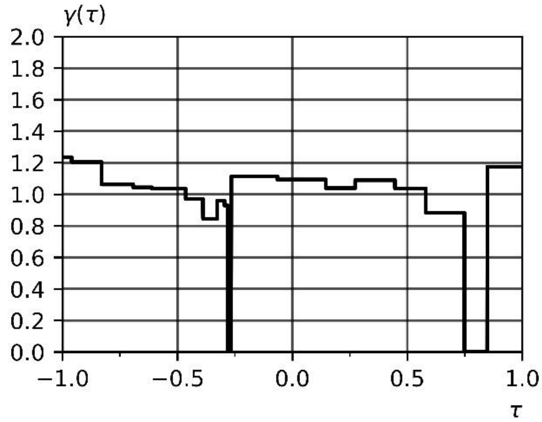



b

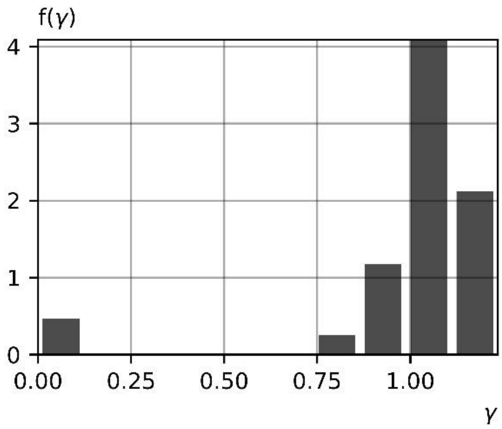
**Fig. 2**. Dimensionless input material flow $\gamma(\tau)$:
a– implementation of the input material flow;
b – histogram of the distribution of values $\gamma$ of the input material flow

where $\gamma_n$ is the average value of the material flow in interval $\tau\in[\tau_{n-1}, \tau_n]$.

The input material flow in accordance with the approximate expression (23) is presented in Fig. 3. For the original implementation of the material flow (Fig. 2) and the approximated implementation (Fig. 3) the correlation functions are shown in Fig. 4.



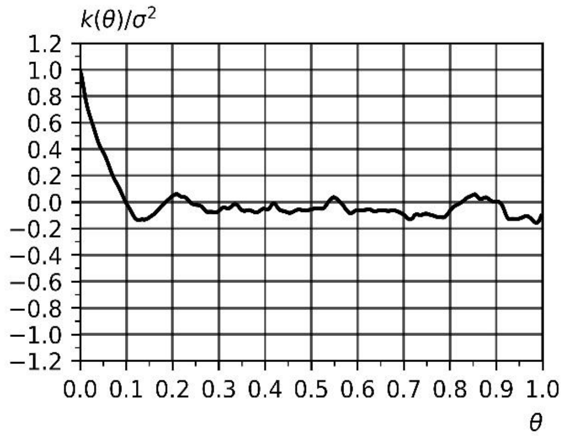

a



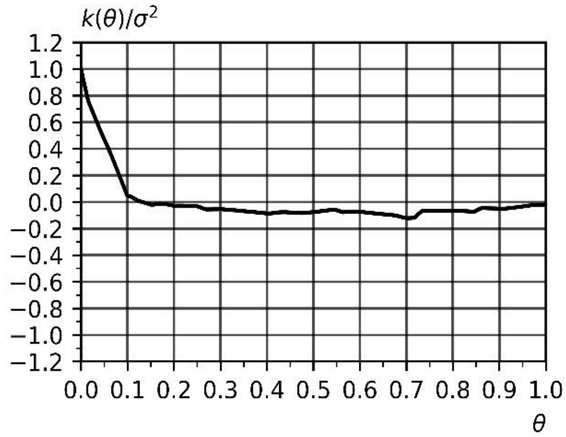

b

Fig**. 3**. Dimensionless approximated input material flow $\gamma(\tau)$:
a – implementation of the input material flow;
b – histogram of the distribution of the input material flow values $\gamma$

Statistical characteristics calculated based on the implementations of the input material flow are presented in Table 1. The mathematical expectation, standard deviation, and even the maximum value differs slightly for the implementations of the input material flow, which indicates a fairly accurate approximation of the input material flow.




a




b

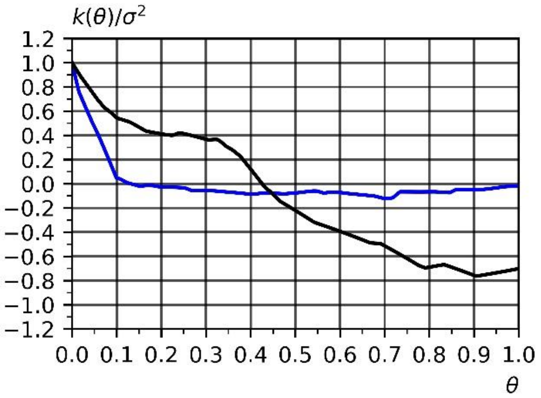
**Fig. 4**: Correlation function $k_s(\vartheta)$ for the implementation of the input material flow:
a – the initial implementation of the input material flow;
b – an approximated implementation of the input material flow

*Table 1* – **Characteristics of the implementation of the input material flow**

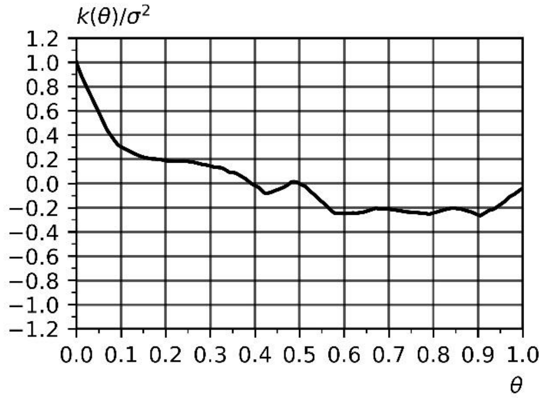
| Parameter | Original implementation, $x_1$ | Approximated implementation, $x_2$ | Relative deviation $\lvert x_1 - x_2 \rvert / x_1$ |
|---|---|---|---|
| Mathematical expectation | 0.0510781 | 0.0514527 | 0.0073338 |
| Standard deviation | 0.0153045 | 0.0145025 | 0.0524028 |
| Maximum value | 0.0725200 | 0.0691900 | 0.0459183 |
| Minimum value | 0.0000000 | 0.0000000 | - |

For small values of $\vartheta \in [0; 0.1]$ the correlation function for the approximated implementation of the input material flow is a fairly accurate approximation of the original implementation. For value $\vartheta > 0.1$, the correlation function for the approximated implementation can be represented as averaging the values of the correlation function for the original material flow implementation. This behavior can be explained by the fact that the average length of the approximation interval $\Delta\tau_n = \tau_n - \tau_{n-1}$ is approximately equal to value $M[\Delta\tau_n] \approx 0.1$.

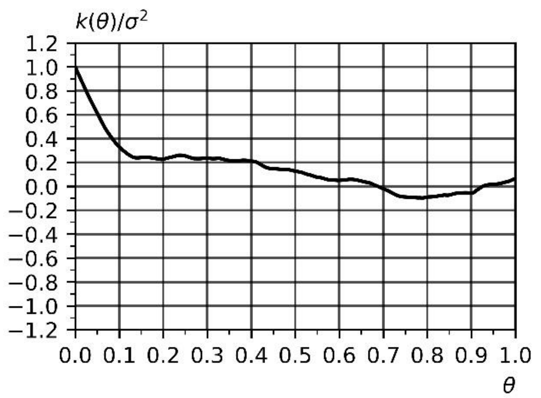
To implement the input material flow, constructed using a random process value generator for the input material flow (11), the correlation function is presented in Fig. 5.



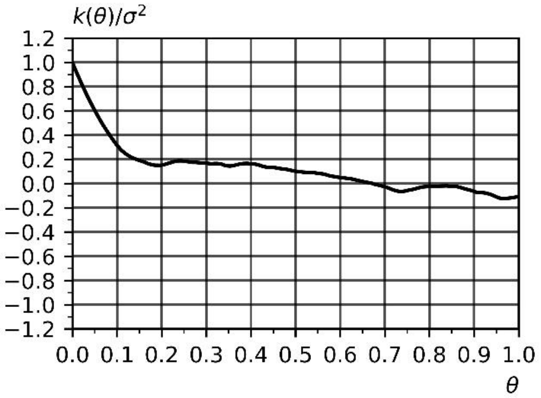

a – $\tau_{\min} = -1$; $\tau_{\max} = 1$



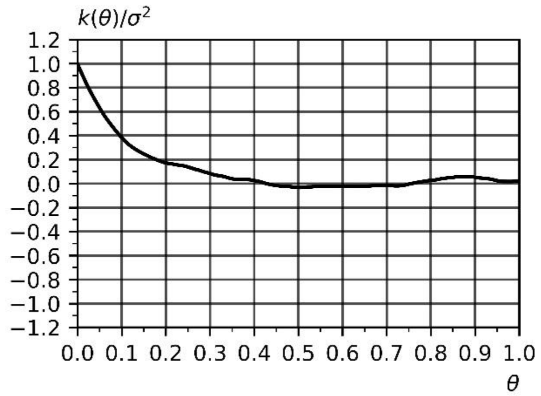

b – $\tau_{\min} = -1$; $\tau_{\max} = 3$



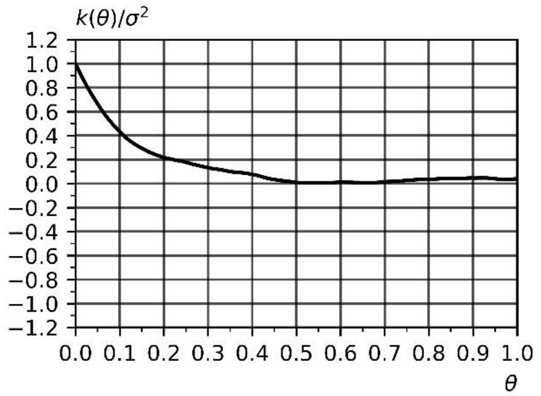

c – $\tau_{\min} = -1$; $\tau_{\max} = 7$




d – $\tau_{\min} = -1$; $\tau_{\max} = 15$




e – $\tau_{\min} = -1$; $\tau_{\max} = 31$



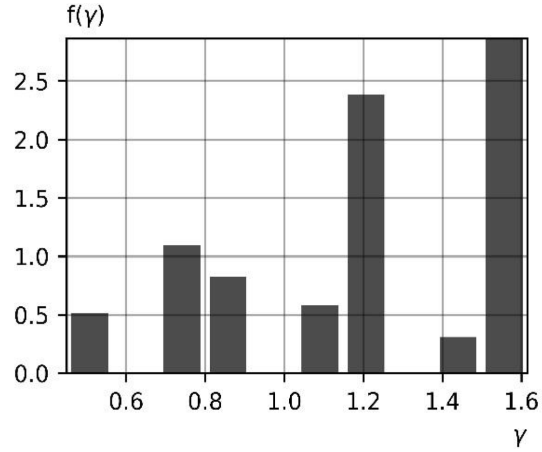

f – $\tau_{\min} = -1$; $\tau_{\max} = 63$

**Fig. 5**: Correlation function $k_s(\vartheta)$for implementing the generated values of the input material flow at time interval $[\tau_{\min}, \tau_{\max}]$

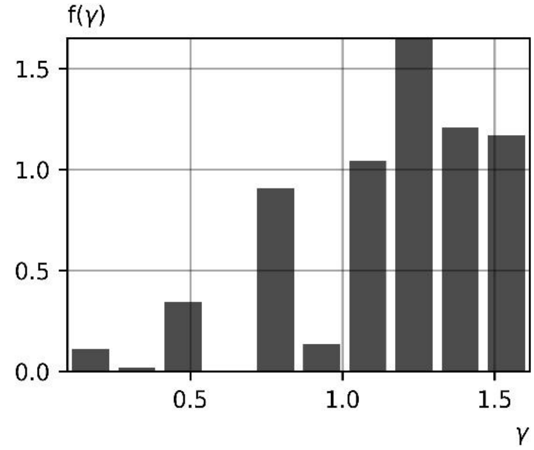
For generation time interval $\tau \in [-1; 1]$, corresponding to the initial implementation for dimensionless material flow $\gamma(\tau)$, the correlation functions for the original and generated implementations are quite different (Fig. 5, a).

A double (Fig. 5, b) and quadruple (Fig. 5, c) increase in the generation time interval led to the convergence of the two correlation functions.

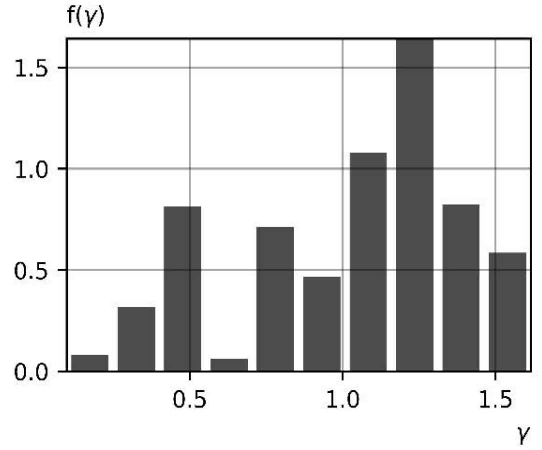
A further increase in the time interval of generating values of the input material flow leads to the fact that the correlation function of the generated material flow takes the form in accordance with the expression (20) obtained above. The distribution densities of random process $\gamma(\tau)$ values for each variant of the correlation function in accordance with the length of the generation time interval are presented in Fig. 6.



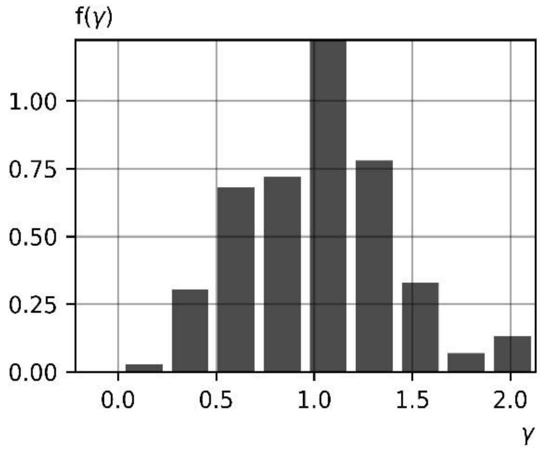

a – $\tau_{min} = -1$; $\tau_{max} = 1$



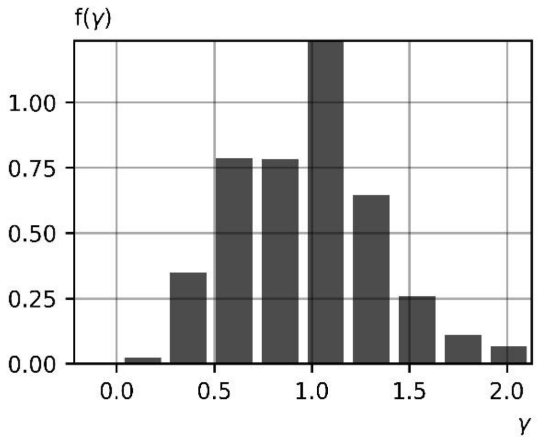

b – $\tau_{min} = -1$; $\tau_{max} = 3$



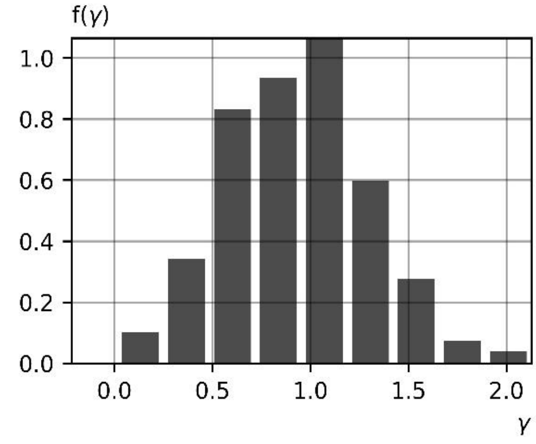

c – $\tau_{min} = -1$; $\tau_{max} = 7$




d – $\tau_{min} = -1$; $\tau_{max} = 15$




e – $\tau_{min} = -1$; $\tau_{max} = 31$




f – $\tau_{min} = -1$; $\tau_{max} = 63$

**Fig. 6**: The histogram distribution of generated values for the input flow of material $k_s(\vartheta_i)$ on time interval $[\tau_{min}, \tau_{max}]$

Comparative analysis of statistical characteristics for the initial, approximated and generated implementation of the input material flow is given in Table 2. The characteristics of the generated implementation of the input material flow are given for interval $\tau \in [-1; 63]$. The implementations of the input material flow, which are given in Table 2, have similar values of the mathematical expectation and standard deviation.

Let us consider the implementation of the input material flow generated in accordance with the canonical expansion of the random process (11) on interval $\tau \in [-1; 1]$ (Fig. 7). The histogram of the distribution of dimensionless values of the input material flow in this case, as well as the correlation function, are quite different from the histogram of the distribution of dimensionless values of the initial implementation of the input material flow.

*Table 2* – **Comparative analysis of the characteristics of the implementation of the input material flow**

| Parameter | Original implementation | Approximated implementation | The generated implementation |
|---|---|---|---|
| Mathematical expectation | 1.0000000 | 1.0073338 | 0.9766413 |
| Standard deviation | 0.2996297 | 0.2839282 | 0.2857810 |
| Maximum value | 1.4197871 | 1.3545927 | 1.9190534 |
| Minimum value | 0.0000000 | 0.0000000 | 0.0843052 |

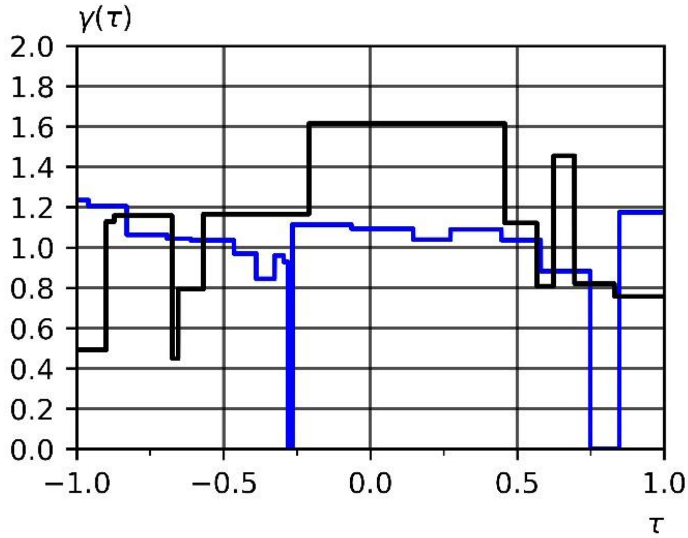



**Fig. 7**: An input material flow $\gamma(\tau)$ :
blue line is implementation of the input material flow based on experimental data;
black line is generated implementation of the input material flow

Increasing the interval for generating random values of the input material flow leads to the fact that the distribution law of the values of the generated implementation of the input material flow approaches the normal distribution law, but is not a normal distribution law. Thus, the used canonical expansion (11) with expansion coefficients of $\Theta_n$ can be accepted for the analysis of input material flows only as a zero approximation. Further research is required to clarify the choice of the type of distribution law for random variables $\Theta_n$, which are used as coefficients of the canonical expansion (11). However, despite similar correlation functions (Fig. 4) and similar values of statistical characteristics (Table 2), the presented material flows differ significantly. The type of distribution law for the values of the input material flow has a significant impact on the generation of values the implementation of the input material flow.

## Conclusions

The current problem of generating a data set for training a neural network in a model of a branched, extended transport conveyor is considered. To generate a training data set, a method is proposed for constructing implementations of the input material flow with given statistical characteristics using a generator of random values of the input material flow.

To solve the problem of generating a data set for training a neural network, a two-stage method for constructing an implementation of the input material flow is proposed.

An approximation of the experimental implementation of the input material flow by a canonical decomposition with given coordinate functions in order to determine its statistical characteristics is performed at the first stage. At the second stage, the statistical parameters of the decomposition coefficients used to generate random values of the input material flow are calculated.

An estimate is given of the minimum length of the time interval required to generate random values. It is shown that the length of the time interval has a significant impact on the statistical characteristics of the constructed implementation of the input material flow.

REFERENCES


1. Tintaya, H. (2002), *DIN 22101:2002-08. Continous conveyors. Belt conveyors for loose bulk materials. Basics for calculation and dimensioning*. GERMAN STANDARD, Normenausschuss Maschinenbau (NAM) im DIN, 51 p., available at: https://www.academia.edu/42835007/DIN_22101_Belt_conveyors_for_loose_bulk_materials
2. SIEMENS (2023), "SIMINE portfolio for mining transportation", *Siemens Xcelerator Marketplace*, available at: https://www.siemens.com/global/en/industries/mining-industry/transport.html
3. Alspaugh, M. (2004), "Latest developments in belt conveyor technology", *2004 Int. Conf. MINExpo*, Las Vegas, USA, pp. 1-5, https://www.academia.edu/36592758/37726652-Latest-Developments-in-Belt-Conveyor-Technology.pdf
4. Mathaba, T. and Xia, X. (2015), "A parametric energy model for energy management of long belt conveyors", *Energies*, vol. 12, no. 8, pp. 13590–13608, doi: https://doi.org/10.3390/en81212375
5. Stadnik, M., Semenchenko, D., Semenchenko, A., Belytsky, P., Virych, S. and Tkachov, V. (2019), "Improving energy efficiency of coal transportation by adjusting the speeds of a combine and a mine face conveyor", *Eastern-European Journal of Enterprise Technologies*, vol. 1, no. 8(97), pp. 60–70, doi: https://doi.org/10.15587/1729-4061.2019.156121
6. Pihnastyi, O. and Khodusov, V. (2020), "Development of the controlling speed algorithm of the conveyor belt based on TOU-tariffs", *2020 Int. Conf. Information-Communication Technologies & Embedded Systems* (ICT&ES), Mykolaiv, Ukraine, pp. 1–14, available at: http://ceur-ws.org/Vol-2762/paper4.pdf
7. Semenchenko, A., Stadnik, M., Belitsky, P., Semenchenko, D. and Stepanenko, O. (2016), "The impact of an uneven loading of a belt conveyor on the loading of drive motors and energy consumption in transportation", *Eastern-European Journal of Enterprise Technologies*, vol. 4, no. 1(82), pp. 42–51, doi: https://doi.org/10.15587/1729-4061.2016.75936
8. Pihnastyi, O., Khodusov, V. and Kotova, A. (2022), "The problem of combined optimal load flow control of main conveyor line", *Acta Montanistica Slovaca*, vol. 27, no. 1, pp. 216–229, doi: https://doi.org/10.46544/AMS.v27i1.16
9. Ristić, L., Bebić M. and Jeftenić, B. (2013), "Development of the Algorithm for Energy Efficiency Improvement of Bulk Material Transport System", *Electronics*, no. 8, pp. 30–39, doi: https://doi.org/10.7251/els1317030r
10. Halepoto, I., Shaikh, M. and Chowdhry B. (2016), "Design and Implementation of Intelligent Energy Efficient Conveyor System Model Based on Variable Speed Drive Control and Physical Modeling", *Journal of Control and Automation*, vol. 6, no. 9, pp. 379–388, doi: https://doi.org/10.14257/ijca.2016.9.6.36
11. Bardzinsk, P., Walker, P. and Kawalec, W. (2018), "Simulation of random tagged ore flow through the bunker in a belt conveying system", *International Journal of Simulation Modelling*, no. 4, pp. 597–608, doi: doi.org/10.2507/IJSIMM17(4)445
12. Kawalec, W. and Król, R. (2016), "Generating of Electric Energy by a Declined Overburden Conveyor in a Continuous Surface Mine", *Energies*, vol. 13, no.14, doi: https://doi.org/10.3390/en14134030
13. Pihnastyi, O. and Cherniavska, S. (2022), "Improvement of methods for description of a three-bunker collection conveyor", *Eastern-European Journal of Enterprise Technologies*, vol. 4, no.5(119), pp.33–41, doi: https://doi.org/10.15587/1729-4061.2022.265770
14. Pihnastyi, O, Kozhevnikov, G. and Polozhyi, D. (2023), "Synthesis of an Optimal Control Algorithm for a Transport Conveyor with a Reversing Section", *2023 Int. Conf. IEEE 4th KhPI Week on Advanced Technology*, Kharkiv, Ukraine, pp. 187–191, doi: https://doi.org/10.1109/KhPIWeek61412.2023.10312939
15. Marais, J. (2007), "Analysing DSM opportunities on mine conveyor systems", Thesis (Electronical Engineering) - North-West University, Potchefstroom Campus, Cape Town, South Africa, 145, p., available at: https://repository.nwu.ac.za/bitstream/handle/10394/1825/marais_johannesh%281%29.pdf?sequence=1
16. He, D., Pang, Y., Lodewijks, G. and Liu, X. (2016), "Determination of Acceleration for Belt Conveyor Speed Control in Transient Operation", *International Journal of Engineering and Technology*, vol. 3, no. 8, pp. 206–211, doi: https://doi.org/10.7763/IJET.2016.V8.886
17. Zhang, S. and Xia, X. (2010), "Optimal control of operation efficiency of belt conveyor systems", *Applied Energy*, vol. 87, no 6, pp. 1929–1937, doi: https://doi.org/10.1016/j.apenergy.2010.01.006
18. Pihnastyi, O. and Khodusov, V. (2020), "Hydrodynamic Model of Transport System", *East European Journal of Physics*, no. 1, pp. 121–136, doi: https://doi.org/10.26565/2312-4334-2020-1-11
19. Andrejiova, M. and Marasova, D. (2013), "Using the classical linear regression model in analysis of the dependences of conveyor belt life", *Acta Montanistica Slovaca*, Technical University of Kosice, No. 18(2), pp. 77–84, available at: https://doaj.org/article/3ee2a4afeca94d8e8ce8468d48bbf2ed
20. Lu, Ya. and Li, Q. (2019), "A regression model for prediction of idler rotational resistance on belt conveyor", *Measurement and Control*, No. 52(5), pp. 441–448, doi: https://doi.org/10.1177/0020294019840723
21. Wiecek, D., Burduk, A. and Kuric, I. (2019), "The use of ANN in improving efficiency and ensuring the stability of the copper ore mining process", *Acta Montanistica Slovaca*, vol. 24, no 1, pp. 1–14, available at: https://www.researchgate.net/publication/335571480_The_use_of_ANN_in_improving_efficiency_and_ensuring_the_stability_of_the_copper_ore_mining_process
22. Xinglei, L. and Hongbin, Y. (2015), "The Design and Application of Control System Based on the BP Neural Network", *III International Conference on Mechanical Engineering and Intelligent Systems* (ICMEIS), Tianjin, China, pp. 789–793, available at: https://www.atlantis-press.com/article/25837377.pdf
23. Shareef, I.R. and Hussein, H.K. (2021), "Implementation of Artificial Neural Network to Achieve Speed Control and Power Saving of a Belt Conveyor System", *Eastern-European Journal of Enterprise Technologies*, no. 2(2 (110)), pp. 44–53, doi: https://doi.org/10.15587/1729-4061.2021.224137
24. Pihnastyi, O. and Khodusov, V. (2020), "Neural model of conveyor type transport system", *Third International Workshop on Computer Modeling and Intelligent Systems*, CEUR Workshop Proceedings, vol. 2608, pp. 804–818, available at: http://ceur-ws.org/Vol-2608/paper60.pdf

25. Król, R., Kawalec, W. and Gładysiewicz, L. (2017), "An effective belt conveyor for underground ore transportation systems", *IOP Conference Series: Earth and Environmental Science*, vol. 95, no. 4, pp. 1–9, doi: https://doi.org/10.1088%2F1755-1315%2F95%2F4%2F042047
26. Bajda, M., Błażej, R. and Jurdziak, L. (2019), "Analysis of changes in the length of belt sections and the number of splices in the belt loops on conveyors in an Alspaugh underground mine", *Engineering Failure Analysis*, no. 101, pp. 439–446. doi: https://doi.org/10.1016/j.engfailanal.2019.04.003
27. Pihnastyi, O., Khodusov, V. and Kotova, A. (2023), "Mathematical model of a long-distance conveyor", *Mining Science*, no. 30, pp. 27–43, doi: https://doi.org/10.37190/msc233002
28. Pihnastyi, O. and Ivanovska, O. (2021), "Improving the Prediction Quality for a Multi-Section Transport Conveyor Model Based on a Neural Network", *Int. conf. Information Technology and Implementation* (IT&I-2021), Kyiv, Ukraine, vol. 3132, pp. 24–38, available at: http://ceur-ws.org/Vol-3132/Paper_3.pdf
29. Bhadani, K., Asbjörnsson, G., Hulthén, E., Hofling, K. and Evertsson, M. (2021), "Application of Optimization Method for Calibration and Maintenance of Power-Based Belt Scale", *Minerals*, vol. 11, no 4, doi: https://doi.org/10.3390/min11040412
30. Zeng, F., Yan, C., Wu, Q. and Wang, T. (2020), "Dynamic behaviour of a conveyor belt considering non-uniform bulk material distribution for speed control", *Applied Sciences*, no. 10(13), pp. 1–19, doi: https://doi.org/10.3390/app10134436
31. Curtis, A. and Sarc, R. (2021), "Real-time monitoring volume flow, mass flow and shredder power consumption in mixed solid waste processing", *Waste Management*, no. 131, pp. 41–49, doi: https://doi.org/10.1016/j.wasman.2021.05.024
32. Carvalho, R. and Nascimento, Garcia L. (2020), "A UAV-based framework for semi-automated thermographic inspection of belt conveyors in the mining industry", *Sensors*, no. 22(8), pp. 1–19, doi: https://doi.org/10.3390/s20082243

Received (Надійшла) 15.01.2024
Accepted for publication (Прийнята до друку) 10.04.2024

ВІДОМОСТІ ПРО АВТОРІВ/ ABOUT THE AUTHORS

**Пігнастий Олег Михайлович** – доктор технічних наук, професор, професор кафедри мультимедійних та Інтернет технологій і систем, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
**Oleh Pihnastyi** – Doctor of Technical Sciences, Professor, Professor of Multimedia and Internet Technologies and Systems Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: pihnastyi@gmail.com; ORCID ID: http://orcid.org/0000-0002-5424-9843;
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=557192681564.

**Кожевніков Георгій Костянтинович** – кандидат технічних наук, доцент, професор кафедри мультимедійних та Інтернет технологій і систем, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;
**Georgii Kozhevnikov** – Candidate of Technical Sciences, Associate Professor, Professor of Multimedia and Internet Technologies and Systems Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: heorhii.kozhevnikov@khpi.edu.ua; ORCID ID: http://orcid.org/0000-0002-6586-6767;
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=6701803684.

**Бурдук Анна** – доктор технічних наук, професор кафедри лазерних технологій, автоматизації та організації виробництва, Університет науки і техніки, Вроцлав, Польща;
**Anna Burduk** – Doctor Habilitation Inżynier, Doctor of Technical Sciences, Professor of Laser Technologies, Automation and Production Organization Department, University of Science and Technology, Wrocław, Poland;
e-mail: anna.burduk@pwr.edu.pl; ORCID ID: http://orcid.org/0000-0003-2181-4380;
Scopus ID: https://www.scopus.com/authid/detail.uri?authorId=15031310600.

## Метод формування набору даних для навчання нейронної мережі у моделі транспортного конвеєра

О. М. Пігнастий, Г. К. Кожевніков, А. Бурдук

**Анотація. Об'єкт дослідження –** стохастичний вхідний потік матеріалу, що надходить на вхід транспортної системи конвеєрного типу. Предмет дослідження — розробка методу генерування значень стохастичного потоку вхідного матеріалу транспортного конвеєра для формування навчального набору даних в моделях транспортного конвеєра, заснованих на нейронній мережі. **Метою дослідження** є розробка методу генерування випадкових значень для побудови реалізацій вхідного потоку матеріалу транспортного конвеєра, що має задані статистичні характеристики, розраховані на основі результатів попередньо виконаних експериментальних вимірювань. Результати дослідження. У статті пропонується метод формування набору даних для навчання моделі нейромереж розгалуженого протяжного транспортного конвеєра. Розроблено метод побудови реалізацій стохастичного вхідного потоку матеріалу транспортного конвеєра. Введено безрозмірні параметри, що дозволяють визначити критерії подібності для вхідних потоків матеріалу. Стохастичний вхідний потік матеріалу представлений у вигляді розкладання в ряд координатних функцій. Для формування статистичних характеристик використано реалізацію потоку матеріалу, побудовану на результатах експериментальних вимірювань. Як нульове наближення для коефіцієнтів розкладання, що є випадковими величинами, використано нормальний закон розподілу випадкової величини. Показано, що зі збільшенням часового інтервалу для реалізації вхідного потоку матеріалу функція кореляції згенерованої реалізації наближається до теоретично визначеної функції кореляції. Виконано оцінку довжини часового інтервалу для згенерованої реалізації вхідного потоку матеріалу.

**Ключові слова:** транспортний конвеєр; нейронна мережа; стохастичний процес; генератор набору даних; потік матеріалу.