MATHEMATICAL MODELING OF THE DEVICE FOR TRANSPORTING SMALL-SIZED CARGO

Oleksandr ZARIVNYI[1*], Yuriy ROMASEVYCH[2]

The paper presents a method for synthesizing the mathematical model of a two-wheel device designed for transporting small-sized cargo. It is characterized as an unstable dynamic system and modeled by using a recurrent neural network (RNN) with long short-term memory layers (LSTM). It is trained on experimental data in the form of time series. The latter includes the tilt angle of the device and the slew velocity of the balance mechanism's shaft obtained through experimental research. The processing and analysis of the obtained forecasting results of the system's states one and ten time-steps ahead are conducted with the help of RNN. Mathematical model based on an LSTM RNN describes the device quite well for one-time step ahead. For the ten time-steps ahead the overall forecasting error is slightly shifted to negative values. However, the trained RNN may be used to develop of the device motion optimal control.

Keywords: two-wheel device; unstable dynamic system; LSTM; mathematical model, dynamic system.

1. Introduction

In any human activity at all times, cargo transporting has consistently been one of the main problems and its relevance does not decrease. Methods and means for solving the stated problems develop constantly. The technical means employed in this context are extremely diverse, ranging from delivery robots to gigantic tankers.

This study is dedicated to the question of transporting small-sized cargo that arises in areas of courier delivery in the city, within a single enterprise, or within storage facilities and workshops. In order to perform a task of such scale, a considerable amount of vehicle constructions, ranging from different trolleys to automatized robots, for example, humanoid Boston Dynamics "Atlas" robot [1], Amazon Prime delivery robot "Scout" [2] has been developed. However, two-wheel scooter-like type vehicles have not yet been examined due to the complexity of ensuring their dynamic and static balance. Nevertheless, such two-wheel vehicles offer several advantages over other types: they occupy minimal space during parking or cargo transporting, provide high longitudinal stability, and maintain a

[1*] PhD student, Faculty of Construction and Design, National University of Life and Environmental Sciences of Ukraine, e-mail: alex-zar@ukr.net (corresponding author)
[2] Prof., Dr. of Technical Sciences, National University of Life and Environmental Sciences of Ukraine, e-mail: romasevichyuriy@ukr.net

low center of gravity, allowing for higher speed and movement across public roadways or within storage facilities, etc.

To ensure the stability of the device's motion for transporting small-scale cargo, safe exploitation conditions, optimize energy consumption, and enhance the device's capacity, it is necessary to synthesize a stabilization system for its position. These calculations must be based on an adequate mathematical model of the device's position which later allows for obtaining a control system that ensures the device's position stability. Thus, adequate mathematical model of the device is the goal of the current study. In order to achieve it, the following objectives should be investigated: 1) to analyze the current stage of the researches in the area of mathematical modelling of electromechanical underactuated systems and select the most proper for the further study; 2) to carry out experiments and record the data of device control and movement; 3) develop ANN and train it based on the obtained data; 5) estimate the ANNs errors for device movement and make conclusion about adequateness of ANN application as a device mathematical model.

2. Analysis of relevant scientific articles

The device for which the synthesis of the mathematical model is accomplished refers to the domain of electromechanical systems. Without taking into consideration the electrical part of the device (it is characterized by significantly quicker transient processes in comparison to the mechanical part of the system), the system may be examined as solely mechanical one. In general, there are three approaches to building mathematical models of mechanical systems: a "white-box", a "black box" and a "gray box" - a combination of the first two [3].

The first one is based on physical laws and usually the equation of motion for the system's elements is derived either from Newton's equations or through Euler-Lagrange methods. Such an approach to building mathematical models and constructions of two-wheel vehicles is presented in the work [4], where a structure of a two-wheel scooter-like type vehicle with a flywheel balancing mechanism is proposed. The mathematical model of this system is created in the form of equations of motion with their further linearization. Another two-wheel vehicle, the construction of which is similar to that of a bicycle, is described in the work [5]. Two gyroscopes are used to stabilize its position. The mathematical model of such a device is presented as differential equations of motion. A mathematical model of a two-wheel robot described by an inverted pendulum model is the basis for the work [6]. In the work [7], a motion model of a two-wheel self-balancing robot was obtained, which later allowed for the synthesis of optimal control of its motion using Lagrangian [8] and Kane's [9] methods. A model of a two-wheel device (the Essboard) that has freely rotating wheels and a stepper motor is examined in the

work [10]. Geometric, kinematic, and dynamic characteristics of the device's structure are determined when building its mathematical model.

The mentioned works are characterized by extremely comprehensive (in terms of corresponding measurements and their processing) research. Thus, before building a mathematical model in the form of equations of motion (a "white box") assumptions that simplify the process of equations formulation are made [6, 10]. This allows for the reduction of the number of calculations, however, when neglecting certain parameters (deformation of the device's elements, deflection of the vehicle's wheel's tire, the shape of the wheel contact patch with the ground and its slippage, etc.) a model that can be used only in the first approximation is obtained. Increasing the accuracy of the model demands consideration of other significant effects, which requires a substantially larger amount of calculations.

Geometrical, kinematic, or dynamic characteristics of the device are not expected to be considered when building the "black-box" model. For this, ample is determining the input and output characteristics that describe the state of the system, whereas a mathematical model is determined through numerical methods of function approximation. This approach is applied when building a low-drive system of an inverted pendulum motion model. A least squares estimation is used for this purpose in the work [11]. The advantages of this method lie in its consideration of any relevant factors that influence the functioning of the system. One shortcoming of such an approach is the absence of general recommendations regarding the selection of the functions based on which the approximation of the dynamic system response takes place. This shortcoming may partially be eliminated by using artificial neural networks (ANNs). Particularly, recurrent neural networks (RNNs) with long short-term memory layers (LSTM) which enable high-quality forecasting of time series are used for this purpose [12]. The latter, as is known, are the input and output characteristics of a dynamic system, collected over a period of time. The employment of LSTM has allowed for building a mathematical model of the motion of the mechanical second-order oscillatory system and determining optimal coefficients for a PID controller [13]. Mathematical modeling of the nonlinear Lorenz system is accomplished with the help of LSTM [14]. Data used for the training of the ANN is the time series that is obtained based on the numerical integration of a differential equations system. A similar approach is used in the work [15], where a dynamic nonlinear system is also examined. The problem of forecasting states of nonlinear dynamic systems using an RNN adaptive learning method is examined in the work [16]. In the field of robotics, RNNs are also utilized for discovering mathematical models of motion or synthesizing optimal control, for example, for the two-legged robot [17] or the robot manipulator [18].

In the "gray-box" approach a specific case of both the previous approaches is considered, where the general form of equations of motion is known, but the coefficients or physical characteristics are unknown. The "gray-box" approach is

used for building a mathematical model of a hoverboard [3]; based on the study of the device's kinematics a general equation of motion is determined, and the unknown coefficients are determined based on the experimental data of the system through methods of mathematical approximation of the function. For another case mentioned in the work [19], when the general equation of motion of the device is known, but the coefficients of the equations are unknown, the latter are determined using an RNN.

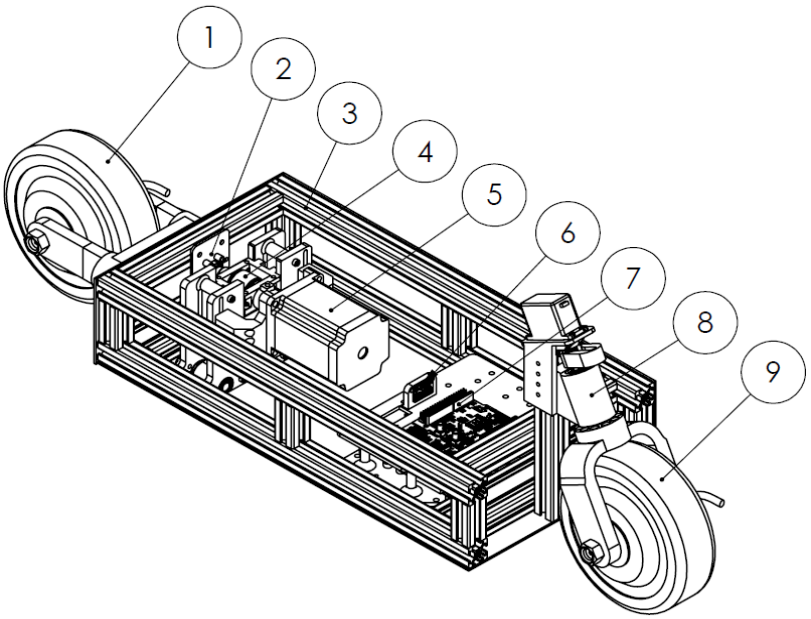
Employment of the "white-box" and "gray-box" approaches requires substantial numerical calculations. Not all dynamic parameters of the system can be measured with sufficient accuracy, including the effects of dry and viscous friction, gaps, clearances, and other nonlinearities. Therefore, when building mathematical models of mechanical systems, the "black-box" approach is often utilized, which involves the further use of LSTM.

3. Materials and methods

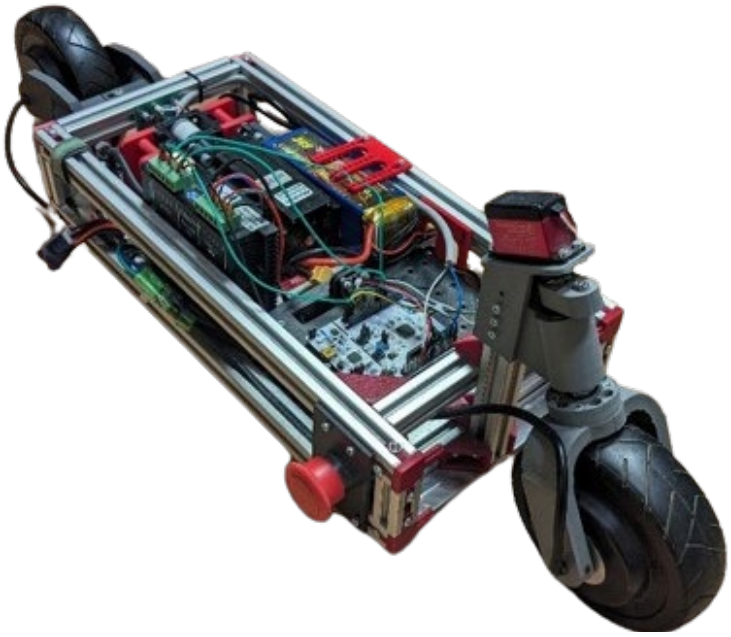
A device for transporting small-sized cargo (Fig. 1, a) with a mass of 10 kg, can transport cargo up to 12 kg. It is a prototype and the studied issue is connected with its mathematical modelling. A developed mathematical model will allow to design of a proper control algorithm, which is the issue for further investigations.

The scale of the device remains unchangeable for the cargoes in the range of approximately 5…12 kg. Increasing the transported masses needs increasing of device scale, and vice versa. However, this issue lies out of the scope of the current study. It allows for the accomplishment of the transporting of the cargo that is placed on its platform across public roadways. The device's balancing is secured by means of a special mechanism (item 4 in Fig. 1, a). Its working principle lies in shifting the device's center of mass when tilted from the equilibrium position. The shift occurs due to the deviation of the rear wheel from the device's frame, driven by a stepper motor (position 5 in Fig. 1, a) NEMA 23 57BYGH2100-4004A-8 (step value of 1.8°, rated current of 4 A, torque value of 2.7N·m. The latter, through the belt drive, tilts crankshafts, which are hinge-mounted on the fork of the rear wheel (Fig. 2).

Arduino Nucleo F446RE development board (item 7 in Fig. 1, a) is used for controlling the process of the device stabilization and kinematic characteristics experimental data record. 9-axis gyro-accelerometer MPU9250 (item 6 in Fig. 1, a) is connected to the board.



a)



b)

Fig. 1. Device for transporting small-sized cargo: a) 3D-model; b) lab installation

In the Fig. 1, a) the following notation are used: 1 - rear drive motor wheel, 2 - sensor of the rear wheel's slew, 3 - frame of the device, 4 - balancing mechanism,

5 - stepper motor, 6 - 9-axis gyro-accelerometer, 7 - control board, 8 - rotary support with the wheel slew mechanism, 9 - front drive motor wheel.

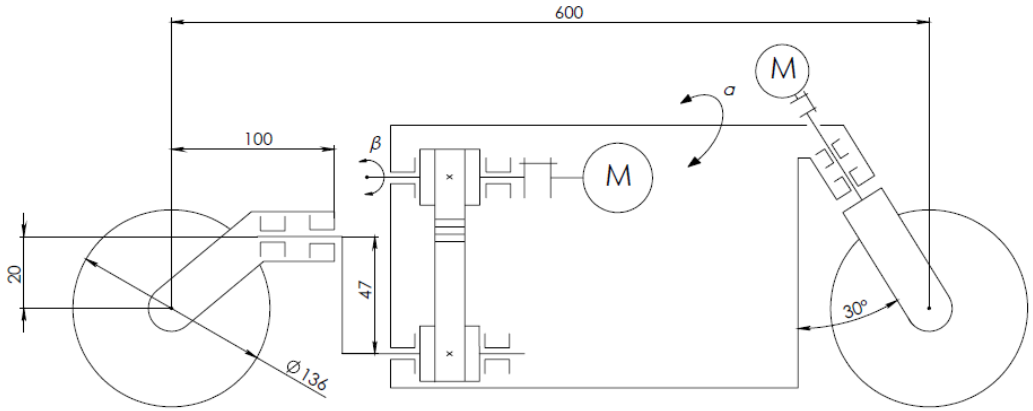


Fig. 2. The kinematic scheme of the device for transporting small-sized cargo augmented with the main dimensions of the device

In order to build a mathematical model of a considered dynamic system, experimental data at its balancing was collected, specifically: the tilt angle α of the device and the angular velocity of the stabilization mechanism $\dot{\beta}$ (Fig. 3).

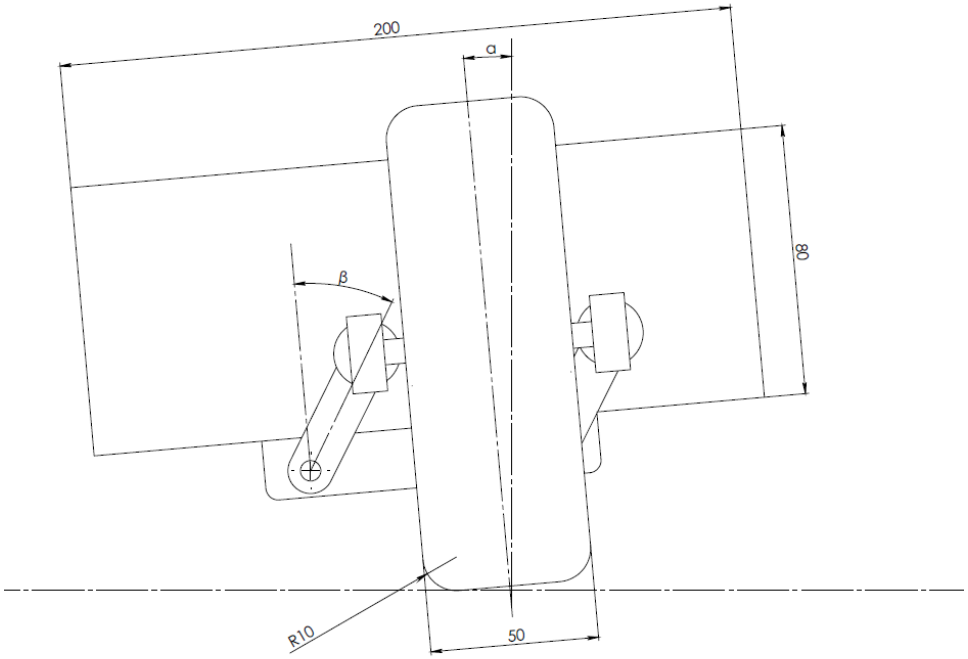


Fig. 3. Indication of angles α and β

Due to the fact that the accelerometer and gyroscope data obtained from the MPU9250 module is affected by noise, a complementary method of determining the tilt angle of the device is employed [20]. The following expression is used for this purpose:

$$\alpha_i = K_{gyro}\alpha_{gyro.i} - (1 - K_{gyro})\alpha_{accel.i}, \quad (1)$$

where i – the subscript that indicates the number of the signal measurement cycle; K_{gyro} – the subscript that determines the weight of influence of the gyroscope data (K_{gyro}=0,985) when determining the angle α_i; $\alpha_{accel.i}$ – the angle that is obtained using the accelerometer, which is determined using the following expression:

$$\alpha_{accel.i} = 57{,}29 \cdot \arcsin(u_{y.i} / g), \quad (2)$$

where $u_{y.i}$ – the acceleration of the device along the horizontal axis; g – the acceleration of gravity; $\alpha_{gyro.i}$ – the angle that is obtained using the gyroscope, which is determined using the following expression:

$$\alpha_{gyro.i} = \alpha_{i\text{-}1} + (\varphi_{.i\text{-}1} + \varphi_{z.i})/2 \cdot \Delta t, \quad (3)$$

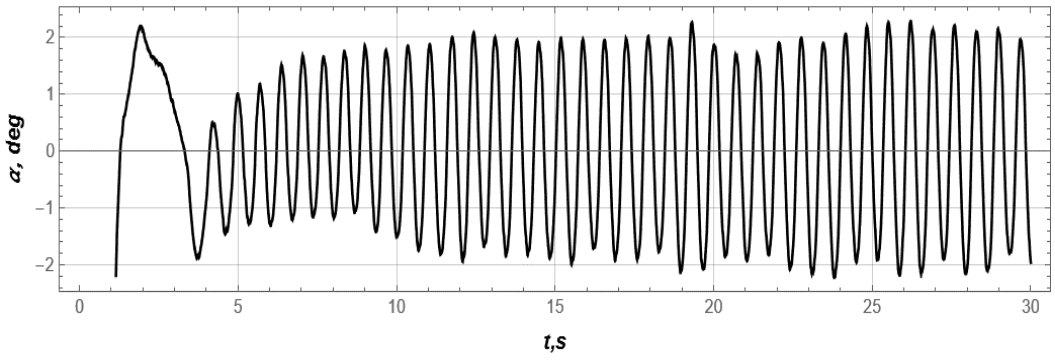
where $\varphi_{z.i\text{-}1}$ and $\varphi_{z.i}$ – the previous and current value of angular velocity obtained using the gyroscope; Δt – the time between the measurements (Δt=0.0036 s). For the first measurement cycle, α_0=0 is set.

The gyroscope data is processed using a digital moving average filter [21]. The filter's window size comprises three measurements.

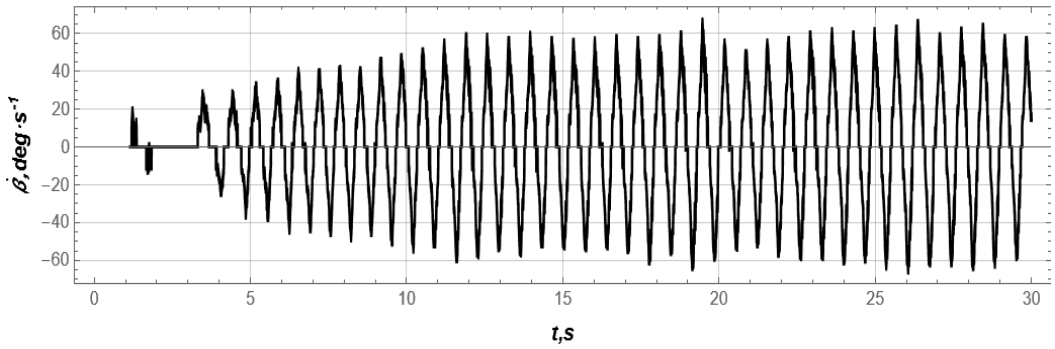
A PD controller is used to collect data of the device's position relative to the vertical axis stabilization. The controller receives the tilt angle α of the device as input, producing the angular velocity of the stabilization mechanism's tilt $\dot{\beta}$ as output. For the PD-controller, values of the proportional and differential component coefficients are selected empirically. Both are equal to three. These specific values allow for stabilizing the device's position relative to the vertical axis. The work of this controller may only be considered as a temporary measure that allows for collecting an experimental data array in the process of stabilizing the device's position.

The device for transporting small-sized cargo requires position stabilization only in one plane. Therefore, characteristics that describe the motion of the device in the plane that is perpendicular to its functional motion direction are chosen for building a mathematical model. One characteristic is the tilt angle α of the device (Fig. 4a) which controls the tilt of the system relative to the vertical axis, and the other characteristic is the angular velocity of the stabilization mechanism's tilt $\dot{\beta}$ (Fig. 4, b) that changes the control influence in the process of the device's position stabilization.

The experimental data record is organized in batches. An array of data comprising 8405 elements is assembled in the form of a time series. The experiment lasted approximately 30 seconds.



a)



b)

Fig. 4. Experimental data: a) the tilt angle α of the device; b) angular velocity of the stabilization mechanism $\dot{\beta}$

Neural network technologies are employed to build the mathematical model of the dynamic system of the device for transporting small-sized cargo. ANNs, created by analogy with the functioning of condensations of biological neurons, are mostly presented as feedforward neural networks. The usage of feedforward ANNs is typically limited to statistical problems that do not consider the time coordinate. Thus, an ANN of another type, specifically an RNN, should be used for problems of time series forecasting. They also have certain constraints, for example, they almost do not consider states of the system that are obtained more than 10 time steps back. This is due to the vanishing (or "explosion") of the error gradient when training an ANN. LSTM which can consider up to 1000 time steps depending on the complexity of the ANN structure is devoid of this shortcoming [22]. This is why specifically the LSTM was used for the mathematical modeling of the stabilization dynamics of the device for transporting small-sized cargo.

The implementation of the structure, calculations, and training of the ANN are accomplished in the Wolfram Cloud environment [23]. It consists of the input layer, hidden layers, and output layer (Fig. 5). Neurons of the input layer (Fig. 5)

receive a 10x2 matrix. It corresponds to ten time measurements of the system's state, each of which includes two characteristics, namely the tilt angle α of the device and the angular velocity of the stabilization mechanism $\dot{\beta}$. This number of the system's states is chosen because at lower values the LSTM has not allowed for obtaining a high-quality forecast of the time series. The hidden layer consists of 3 layers each composed of 20 LSTM elements (elements 1-3 in Fig. 5). This number was selected empirically. A *SequenceLastLayer* (element 4 in Fig. 5) which returns the last value of the succession (a 10x20 matrix) of outputs of the previous layer of the LSTM is used after the LSTM layers. The linear layer (element 5 in Fig. 5) is the last one in this structure. Thus, the developed ANN structure receives 10 system states as input and returns a one time-step ahead forecast.

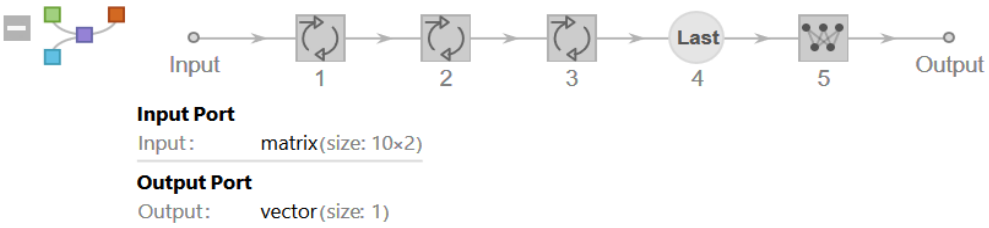


Fig. 5. The ANN structure used in the research

After the ANN structure development, its training is conducted. Training, validation, and test datasets must be prepared before the beginning of the training. Each of the datasets must include normalized data which is necessary to avoid the ANN "paralysis" phenomenon. For example, a normalized value of the tilt angle of the device is calculated using the following expression:

$$\hat{\alpha}_k = \frac{(\alpha_k - \alpha_{\min})}{\alpha_{\max} - \alpha_{\min}}, \tag{4}$$

where α_k – is the k-th value of the angle α in the time series; $\alpha_{\min}$ and $\alpha_{\max}$ are the minimum and maximum values of the angle in the time series. Here and below the upper symbol «^» means normalized values. A similar procedure is carried out for the $\dot{\beta}$ data. All three datasets contain normalized data organized into training pairs:

$$\hat{\mathrm{X}} \to \hat{\mathrm{Y}}, \tag{5}$$

where $\hat{\mathrm{Y}}$ – the ANN output ($\hat{\alpha}_{k+1}$ scalar) that corresponds to the forecast of the tilt angle $\hat{\alpha}$ in the subsequent time interval; k – the number of the measurement that takes values from 10 to 5379 (the volume of the training data); $\hat{\mathrm{X}}$ – the 10x2 matrix that is received at the ANN input and which is presented as follows:

$$\hat{\mathrm{X}} = \begin{bmatrix} \hat{\alpha}_{k-9} & \hat{\alpha}_{k-8} & \hat{\alpha}_{k-7} & \hat{\alpha}_{k-6} & \hat{\alpha}_{k-5} & \hat{\alpha}_{k-4} & \hat{\alpha}_{k-3} & \hat{\alpha}_{k-2} & \hat{\alpha}_{k-1} & \hat{\alpha}_{k} \\ \hat{\beta}_{k-9} & \hat{\beta}_{k-8} & \hat{\beta}_{k-7} & \hat{\beta}_{k-6} & \hat{\beta}_{k-5} & \hat{\beta}_{k-4} & \hat{\beta}_{k-3} & \hat{\beta}_{k-2} & \hat{\beta}_{k-1} & \hat{\beta}_{k} \end{bmatrix}^{T} \tag{6}$$

The validation data includes 1345 training pairs, and the test data includes 1681 training pairs. The ANN training is carried out using the ADAM method [24] with a batch size of 100 and a total of 1000 training cycles (Fig. 6). The size of the batch and the number of the training cycles *R* are selected empirically (Fig. 6). Upon 800 cycles of the ANN training the reduction of the error magnitude almost does not occur, so the ANN may be considered trained. Apart from that, the analysis of the error plot based on the validation dataset (Fig. 6) indicates the absence of the ANN overfitting.

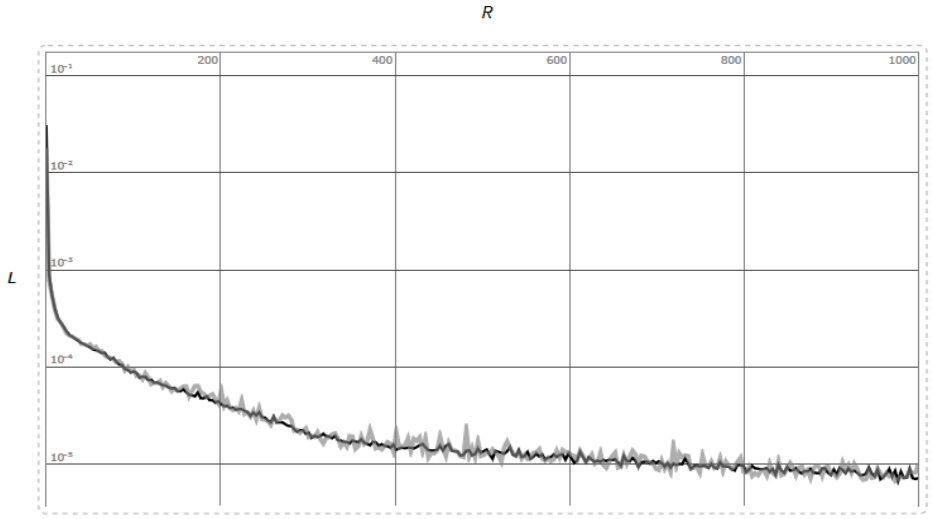


Fig. 6. The error magnitude *L* during the ANN training on the training (black plot) and the validation (gray plot) datasets

Upon training the ANN, a quality assessment of its forecasts is conducted based on the test dataset. The determination of forecast errors occurs in several stages. At first, an ANN forecast array is formed based on the test dataset:

$$\hat{\mathrm{Y}}_{ANN} = f\left(\hat{\mathrm{X}}_{test}\right), \tag{7}$$

where $\hat{\mathrm{Y}}_{ANN}$ – is the ANN forecast array based on the test dataset; f – the nonlinear operator of the trained ANN; $\hat{\mathrm{X}}_{test}$ – is the input 10x2 matrix of the ANN from the test dataset.

It is worth mentioning the originality of ANN testing for ten time-steps ahead, which is applied in the current study. Since the dynamical controlled system

is under consideration the forecast error (only for the case of ten time-steps ahead) is carried out in the following manner. The ANN is fed by the matrix $\hat{\mathrm{X}}_1 = \begin{bmatrix} \hat{\alpha}_{k-9} & \hat{\alpha}_{k-8} & \hat{\alpha}_{k-7} & \hat{\alpha}_{k-6} & \hat{\alpha}_{k-5} & \hat{\alpha}_{k-4} & \hat{\alpha}_{k-3} & \hat{\alpha}_{k-2} & \hat{\alpha}_{k-1} & \hat{\alpha}_k \\ \hat{\dot{\beta}}_{k-9} & \hat{\dot{\beta}}_{k-8} & \hat{\dot{\beta}}_{k-7} & \hat{\dot{\beta}}_{k-6} & \hat{\dot{\beta}}_{k-5} & \hat{\dot{\beta}}_{k-4} & \hat{\dot{\beta}}_{k-3} & \hat{\dot{\beta}}_{k-2} & \hat{\dot{\beta}}_{k-1} & \hat{\dot{\beta}}_k \end{bmatrix}^T$ and returns the scalar $\hat{\alpha}_{k+1}$. The next step: ANN is fed by the matrix $\hat{\mathrm{X}}_2 = \begin{bmatrix} \hat{\alpha}_{k-8} & \hat{\alpha}_{k-7} & \hat{\alpha}_{k-6} & \hat{\alpha}_{k-5} & \hat{\alpha}_{k-4} & \hat{\alpha}_{k-3} & \hat{\alpha}_{k-2} & \hat{\alpha}_{k-2} & \hat{\alpha}_k & \hat{\alpha}_{k+1} \\ \hat{\dot{\beta}}_{k-8} & \hat{\dot{\beta}}_{k-7} & \hat{\dot{\beta}}_{k-6} & \hat{\dot{\beta}}_{k-5} & \hat{\dot{\beta}}_{k-4} & \hat{\dot{\beta}}_{k-3} & \hat{\dot{\beta}}_{k-2} & \hat{\dot{\beta}}_{k-1} & \hat{\dot{\beta}}_k & \hat{\dot{\beta}}_{k+1} \end{bmatrix}^T$ and returns the scalar $\hat{\alpha}_{k+2}$. Note, that in the matrix $\hat{\mathrm{X}}_2$ the forecast $\hat{\alpha}_{k+1}$ of previous step and the control value of the current step are used. The procedure must be continued until $\hat{\alpha}_{k+10}$ is obtained. Thus, for the ANN forecasting from the second to tenth steps the data „fusion" is exploited: the forecasted angles $\hat{\alpha}$ and measured angular velocities $\hat{\dot{\beta}}$ form the ANN inputs.

Further, the ANN forecast error array is calculated:

$$\hat{\mathrm{E}} = \hat{\mathrm{Y}}_{ANN} - \hat{\mathrm{Y}}_{test}, \tag{8}$$

where $\hat{\mathrm{E}}$ – is the ANN forecast error array; $\hat{\mathrm{Y}}_{test}$ – is the ANN output ($\hat{\alpha}_{i+1}$ scalar), that corresponds to the forecast of the tilt angle α in the subsequent time interval in the test dataset. Upon that, a denormalization of the obtained data is conducted:

$$\mathrm{E} = \frac{\hat{\mathrm{E}}}{\alpha_{\max} - \alpha_{\min}} + \alpha_{\min}. \tag{9}$$

Quality assessment of the ANN forecasting is conducted using four indicators, two of which are absolute and two of which are relative ones. Absolute indicators are obtained using the following formulas:

$$e_{RMS} = \sqrt{\frac{1}{1681} \cdot \sum_{j=1}^{1681} e_j^2}, \tag{10}$$

$$e_{\max} = \max\left(\left|e_j\right|\right), \tag{11}$$

where e_j − is the j-th value of the forecast error in array E; j − the subscript that takes values from 1 to 1681 (the volume of the test dataset).

Relative indicators are obtained using the following formulas:

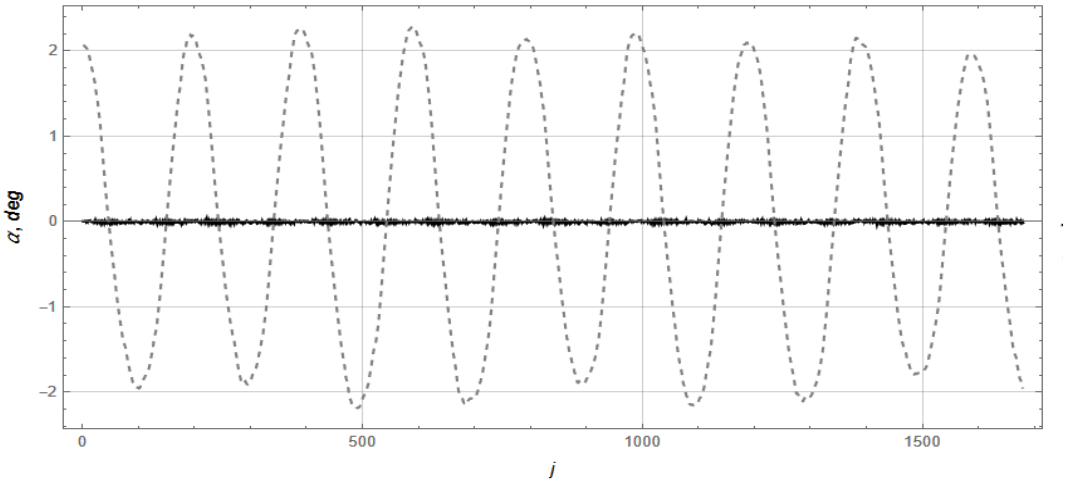
$$e_{RMS.\%} = \frac{e_{RMS}}{\sqrt{\frac{1}{1681}\cdot\sum_{j=1}^{1681}\alpha_j^2}}\cdot 100\%, \tag{12}$$

$$e_{\max.\%} = \frac{e_{\max}}{\max\left(\left|\alpha_j\right|\right)}\cdot 100\%. \tag{13}$$

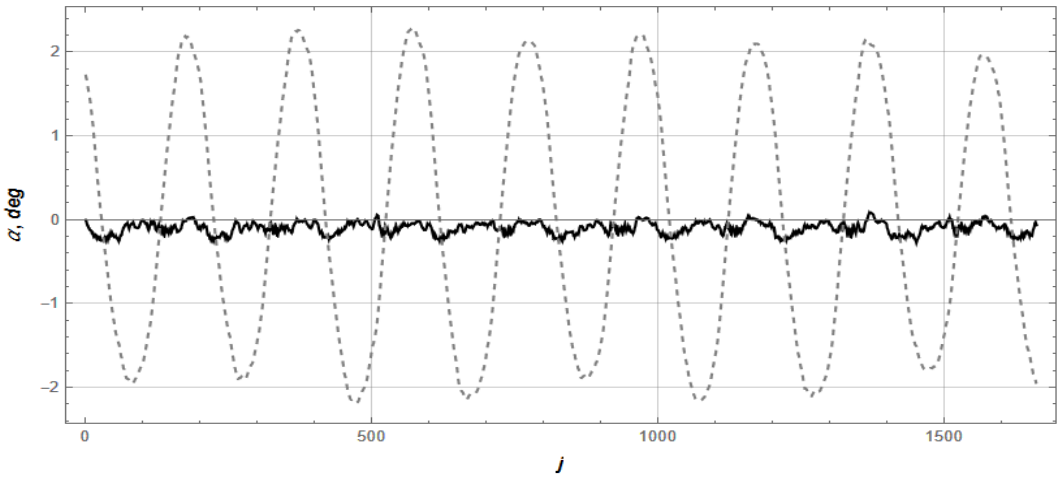
We further conduct the analysis of the obtained data.

4. Results and discussion

Verification of the performance of the ANN in forecasting the tilt angle of the device for transporting small-sized cargo is conducted in two stages. In the first stage, one time-step ahead forecasting is conducted, while in the second stage, ten time-steps ahead are conducted. The obtained results are presented in Fig. 7. Fig. 7 demonstrates that the magnitude of the error is significantly smaller than the tilt angle value. Additionally, values of numerical estimates (10) - (13) were calculated and are provided in Table 1.



a)



b)

Fig. 7. The tilt angle of the device plot (dotted) and the ANN performance error plot (solid) based on the test data when verifying the forecast 1 step ahead (a) and 10 steps ahead (b)

Table 1.

Estimates of denormalized forecast errors

Forecast horizon	Estimates			
	e_{RMS}, deg	$e_{RMS.\%}$, %	e_{max}, deg	$e_{max.\%}$, %
1 step ahead	$1.24 \cdot 10^{-2}$	$8.11 \cdot 10^{-1}$	$3.84 \cdot 10^{-2}$	$1.68 \cdot 10^{0}$
10 step ahead	$5.08 \cdot 10^{-2}$	$3.34 \cdot 10^{0}$	$1.65 \cdot 10^{-1}$	$7.24 \cdot 10^{0}$

In Fig. 7, a, the plot of the tilt angle error when forecasting 1 step ahead does not show clearly defined patterns, while cyclic error changes can be observed when forecasting 10 steps ahead (Fig. 7, b). Moreover, the average value of the forecast error of the tilt angle of the device is displaced downwards relative to the O*j* axis. Thus, we can conclude that when conducting forecasts 10-time steps ahead, the developed ANN on average returns values of the angle α that are underestimated. Apart from that, in Fig. 8 shows an increase in the error magnitude at low values of the angle α.

In general, graphic dependencies (Fig. 7) and the calculated estimates (Table 1) give grounds for asserting that the developed ANN allows for the accomplishment of high-quality forecasting of the future states of the dynamic system (device), which allows for treating it as the mathematical model of the device for small-sized cargo transporting.

The applied here approach makes it possible to indicate the negative shift of the forecast error while the forecasting horizon is increased. Analysis of this fact brings two original outputs: 1) there is a reasonable limit of forecasting when the error becomes significant and further increasing of horizon forecasting is connected

with unreliable data appearance; 2) the trend of the forecasting error shift may be taken into account to correct the ANN outputs and to improve the quality of ANN forecasting. The latter is an issue for further investigations, which requires a much deeper understanding of the correlations „time horizon – ANN forecast error".

In general, the proposed approach gives the technique for estimation of ANNs forecasting errors, especially of middle- and long-terms forecasts. The technique application helps to understand the evolution of errors forecasting with relation to the forecasting term (for how many steps ahead the forecast is made).

5. Conclusions

1) In order to build a mathematical model of the device for small-sized cargo transporting position stabilization dynamics an analysis of the up-to-date scientific research on this subject is conducted. As a result of the analysis, the "black-box" approach is chosen for achieving the stated goals. It is implemented by using an ANN with LSTM layers.
2) In order to train the ANN an experiment is conducted and an array of experimental data in the form of a time series corresponds to the succession of the system's states (tilt angle α) and the applied controls (angular velocity of the stabilization mechanism's slew $\dot{\beta}$) is collected. The obtained data array is organized into training pairs and separated into the training, validation, and test datasets. The ANN training and the verification based on the validation dataset have shown a high quality of forecasts and the absence of overfitting.
3) When validating the trained ANN based on the test data it has been determined that it returns a sufficiently high quality of the subsequent values of the tilt angle α 1 and 10 time steps ahead forecast. Moreover, the quality of the ANN forecasts decreases with the forecast horizon increasing, however, such a decrease in the forecasting quality is not significant. For example, for the $e_{RMS.\%}$ error, the quality of the forecast decreased by a factor of 4 when increasing the forecast horizon by a factor of 10.
4) The obtained results allow for recommending the trained ANN as the mathematical model of the device for transporting a small-sized cargo position stabilization system, which will further be used for the synthesis of an optimal system of the device position stabilization.

REFERENCES

[1]. Robot Atlas. URL: https://bostondynamics.com/atlas/. Access data 21.01.2024.
[2]. Robot «Scout». URL: https://www.aboutamazon.com/news/transportation/meet-scout. Access data 21.01.2024.

[3]. *Komor D., Roman R., Precup R., David R., Pamfilii I*. Models of Two-Wheeled Mobile Robots with Experimental Validation. 14th International Symposium on Applied Computational Intelligence and Informatics. 2020, pp. 211-216. doi: 10.1109/SACI49304.2020.9118823.

[4]. *Ngoc K.V., Hong Q. N*. Design Low-Order Robust Controller for Self-Balancing Two-Wheel Vehicle. Hindawi, Mathematical Problems in Engineering. 2021, pp. 1-22. doi: 10.1155/2021/6693807.

[5]. *Tofigh M.A., Mahjoob M.J., Hanachi M.R., Ayati M*. Fractional sliding mode control for an autonomous two-wheeled vehicle equipped with an innovative gyroscopic actuator. Robotics and Autonomous Systems. 103756. doi: 10.1016/j.robot.2021.103756.

[6]. *Jeyed A., Ali J*. A nonlinear optimal control based on the SDRE technique for the two-wheeled self-balancing robot. Australian Journal of Mechanical Engineering. 2020, pp. 1–9. doi: 10.1080/14484846.2020.1745733.

[7]. *A. Ghaffari, A. Shariati & A. H. Shamekhi*. A modified dynamical formulation for two-wheeled self-balancing robots. Nonlinear Dynamics. **vol. 83**, 2016, pp. 217–230. doi: 10.1007/s11071-015-2321-9.

[8]. *D. Morin*. Introduction to Classical Mechanics. 2008, p. 719. doi: 10.1017/CBO9780511808951

[9]. *Z. Hussain*. KANE's method for dynamic modeling. 2016 IEEE International Conference on Automatic Control and Intelligent Systems. 2016, pp. 174-179. doi: 10.1109/I2CACIS.2016.7885310.

[10]. *Yisheng G., Zhaoheng Z., Daye C., Tao Z., Haifei Z., Li H*. Kinematic Modeling, Analysis, and Verification of an Essboard-Like Robot. Transactions on Mechatronics. 2021, pp. 864-875. doi: 10.1109/TMECH.2020.3009421

[11]. *K. Meidani, A. B. Farimani*. Identification of parametric dynamical systems using integer programming. Expert Systems with Applications. **vol. 219**. 2023, pp. 1-12 doi: 10.1016/j.eswa.2023.119622.

[12]. *Peter J. Brockwell, Richard A. Davis*. Introduction to Time Series and Forecasting. 2016, p. 425. doi: 10.1007/978-3-319-29854-2.

[13]. *Yong-Seok Lee, Dong-Won Jang*. Optimization of Neural Network-Based Self-Tuning PID Controllers for Second Order Mechanical Systems. Applied Sciences. **vol. 11**. 2021, pp. 1-12. doi: 10.3390/app11178002.

[14]. *S. Pawar, O. San, A. Rasheed, I. M. Navon*. A nonintrusive hybrid neural-physics modeling of incomplete dynamical systems: Lorenz equations. GEM - International Journal on Geomathematics. **vol. 12** 2021, pp. 1-24. doi: 10.48550/arXiv.2104.00114.

[15]. *F. Zhijun, Y. Peixin, L. Xiaoli, Y. Yuming*. Identification of model-free nonlinear system via parametric dynamic neural networks with improved learning. International Journal of Innovative Computing, Information and Control. **vol. 17**. 2021, pp. 1871-1885. doi: 10.24507/ijicic.17.06.1871.

[16]. *L. Shanwu, Y. Yongchao*. A recurrent neural network framework with an adaptive training strategy for long-time predictive modeling of nonlinear dynamical systems. Journal of Sound and Vibration. **vol. 506**. 2021, doi: 10.1016/j.jsv.2021.116167.

[17]. *B. Çatalbaş, Ö. Morgül*. Two-Legged Robot Motion Control with Recurrent Neural Networks. Journal of Intelligent & Robotic Systems. **vol. 104**., art. 59. 2022, pp. 1-30. doi: 10.1007/s10846-021-01553-5.

[18]. *T. Yuchuang, L. Jinguo, Z. Xin, J. Zhaojie*. Four-Criterion-Optimization-Based Coordination Motion Control of Dual-Arm Robots. Transactions on cognitive and developmental systems. **vol. 15**., № 2. 2022, pp. 794-807. doi: 10.1109/TCDS.2022.3182534.

[19]. *W. De Groote, E. Kikken, E. Hostens, S. Van Hoecke, G. Crevecoeur*. (2021). Neural Network Augmented Physics Models for Systems with Partially Unknown Dynamics: Application to Slider-Crank Mechanism. IEEE/ASME Transactions on Mechatronics. **vol. 27**. 2021, pp. 103-114. doi: 10.1109/TMECH.2021.3058536

[20]. *Flavius-Catalin P., Szeidert I., Ioan F., Vasar C*. Two-Wheeled Self-Balancing Robot. 2021. 15th International Symposium on Applied Computational Intelligence and Informatics. doi: 10.1109/saci51354.2021.9465568.

[21]. *Rob J Hyndman*. Moving averages. International Encyclopedia of Statistical Science. 2010, pp. 866-869. doi: 10.1007/978-3-642-04898-2_380.

[22]. *R. C. Staudemeyer, E. R. Morris*. Understanding LSTM. 2019, pp. 1-42. doi: 10.48550/arXiv.1909.09586.

[23]. Wolfram Cloud. URL: https://www.wolframcloud.com/. Access data 21.01.2024.

[24]. *D. P. Kingma, J. Ba*. ADAM: A Method for Stochastic Optimization. 2015, pp. 1-13. doi: 10.48550/arXiv.1412.6980