

Using machine learning algorithms to detect anomalies in the solar heating system

Murat Kunelbayev¹, Abdildayeva Assel², Taganova Guldana³

¹Al Farabi Kazakh National University, Institute of Information and Computer Technologies, Almaty, 05000, Kazakhstan

²Al-Farabi Kazakh National University, Almaty, 05000, Kazakhstan

³L.N.Gumilov Eurasian National University, Nur-Sultan, 05000, Kazakhstan

Received: August 5, 2021. Revised: November 23, 2021. Accepted: December 10, 2021. Published: December 31, 2021.

Abstract- This article explores the use of machine learning algorithms to identify anomalies in the solar heating system. A solar heating system that has been developed consists of several parts to simplify the description and modeling process. The authors propose a new architecture for neural networks based on ordinary differential equations. The idea is to apply the new architecture for practical problems of accident prediction (the problem of extrapolation of time series) and classification (classification of accidents based on historical data). The developed machine learning algorithms, artificial intelligence techniques, the theory of differential equations - these directions allow us to build a model for predicting the system's accident rate. The theory of database management (non-relational databases) - these systems allow you to establish the optimal storage of large time series.

Key words- flat solar collector, solar heating system, machine learning, algorithm

I. INTRODUCTION

Article [1] examined how to improve the performance and reliability of renewable energy systems to reduce the associated investment and operating costs. The work [2] investigated several machine learning (ML) applications in the Prognosis and Health Management (PHM) program aimed at improving the availability and performance of complex systems. The article uses deep learning (DL) methods to develop a system in lithium-ion batteries [3] and others.

Solar hot water (DHW) systems are commercially mature applications of solar thermal technology for residential and low temperature industrial applications. Paper [4] developed performance predictions for planar solar collectors to assess system health and assist in performing maintenance planning tasks.

The article [5] used various versions of artificial neural networks. Performance forecasting and efficiency analysis under various meteorological conditions are studied in SHW systems, solar heat pump and heat storage systems with various neural network models such as multilayer perceptron (MLP), radial basis functions (RBF) and adaptive fuzzy network based system output (ANFIS). In this work [6], a fault diagnosis system was developed, in

which the monitored data are compared with the predicted values, which makes it possible to detect faults in solar collectors with good results [6]. In article [7] with the help of memory (LSTM-RNN), implemented for the analysis of long-term time dependences. In this work [8], the study of LSTM neural networks for predicting global horizontal illumination for the day ahead was compared with satellite data. The article [9] developed controls with electrical grids for sizing a renewable energy storage system.

In this paper [10], the ANN, RNN and LSTM architectures were developed to predict the temperature in the SHW system; and compared under similar conditions to highlight their strengths and weaknesses. By comparing the prediction errors of a model generated using nominal data and data caused by errors, an error-based threshold is determined for classifying the health status of the system. The article [11] investigated synthetic data that is generated in TRNSYS under nominal conditions and conditions caused by anomalies. TRNSYS simulations used manufacturer's ratings, including design temperatures, equipment sizes and capacities, and control scheme for each subsystem. Paper [12] presents a study that combines a suitable artificial neural network (ANN) and TRNSYS to predict the performance of an Integrated Collector Storage (ICS) prototype. The dynamics of total consumption of electric and thermal energy in agriculture and in the whole country is considered in the study. The analysis of developments in the field of solar and bioenergetics their current state and possible directions of research on the use of RES in livestock production is given. A mathematical model of a hybrid energy supply system based on a photo-solar-bio energetic installation is developed which makes it possible to reduce the energy intensity of agricultural production. Experimental results of studies of the effectiveness of the proposed hydro breeding system are presented [13]. This article [14] focuses on the design and development of a knowledge-based decision support system (KI-DSS) for maintenance, repair and maintenance in industrial enterprises to support more efficient maintenance decisions and improve maintenance efficiency. This article [15] presents a theoretical study on reducing the starting torque of an unconventional reciprocating piston pump using new methods, for example, by changing the parameters of a wind turbine, such as the aerodynamic configuration of the rotor and blade elements, or by studying

the effect of wind speed on the starting torque.

Prediction of accidents and stable operation of technical devices is an urgent task in various fields. The forecast must be done in advance in order to ensure the high-quality operation of the devices. It is logical that this problem is cost-effective compared to the problem of complete replacement or overhaul of a specific technical device. Determination of accidents can be based on an understanding of how the main components of the device work. If all the components are working properly, the performance is better, and therefore there is less chance of an accident. Modern research has shown that continuous monitoring and analysis of components remains the most important task for many companies.

Currently, the only practical way to monitor the health of such components is to detect complex signals such as vibration, pressure and electrical current generated by technical devices and collected by industrial sensors. It is often difficult to extract useful information from such indirect, mixed and noisy signals, so there is a need for data analysis to extract specific characteristics from these complex signals. A machine learning algorithm is proposed for the regression and time series problem. When the model identified the deviation between the trained regression model and the actual value, the abnormal behavior of the technical device was correlated. LightGBM is used to build a regression model, the model results for the RSME metric are about 1%.

Predicting accidents and stable operation of a technical device is an important task in the problem of maintaining high-quality operation of various types of equipment. The solution to this problem in a logical manner is cost-effective in comparison with the problem of a complete replacement or deep repair of a technical device. Pre-failure detection of a technical device is based on predicting the failure rate of the main components of the device, which have a huge impact on performance. Modern research has shown that monitoring the characteristics of the components of a technical device is still the most important task for making informed decisions in the field of operation and technical device.

II. RESEARCH METHOD

The solar heating system consists of several parts, to simplify the description and modeling process, we describe each part of the system separately. Figure 15 shows a solar hot water heating system with a heat exchanger in a storage tank. The coil of the heat exchanger is installed inside the storage tank and allows you to heat water for consumption: for heating and hot water supply.

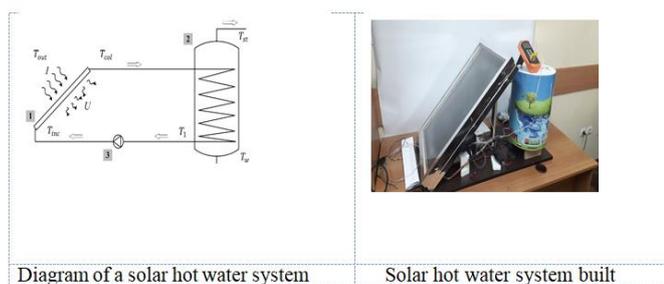


Figure 1. Solar hot water system. 1 - flat solar collector; 2 - storage tank with a heat exchanger; 3 - pump

In the central research object, this solar flat collector separately described the mathematical model of the solar flat collector (Figure 1) and performed energy and exergy analysis. The mathematical model of the collector is defined as a function of: $T_{col} = f(T_{inc}, v_p, I, T_{out}, U)$, where T_{col} is the temperature of the liquid at the outlet of the collector, T_{inc} is the temperature of the liquid at the inlet to the collector, v_{col} is the volumetric flow rate in the collector cycle, I is the solar radiation, T_{out} is the temperature of the collector's environment, U is the heat loss coefficient of the collector.

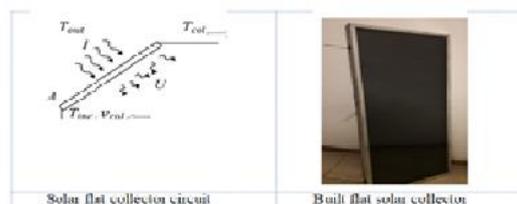


Figure 2. Solar flat collector

The authors propose a new architecture for neural networks based on ordinary differential equations. The idea is to apply the new architecture for practical problems of accident prediction (the problem of extrapolation of time series) and classification (classification of accidents based on historical data). The new architecture has several advantages over existing RNN architectures, such as:

- Efficiency. In training neural differential equations, it is not necessary to compute the gradient through all the operations of the numerical method, and there is also no need to store intermediate results.
- Flexibility of working hours. Compared to solutions of differential equations, where the grid spacing determines the accuracy of the model and affects the computation time, in neural differential equations we can explicitly control the balance between numerical accuracy and computational costs.
- Number of parameters. Comparing ResNet, the advantage of the proposed architecture is the smaller number of parameters, since no new parameters are required for each layer.

The methodological basis of the research will be: Machine learning algorithms, artificial intelligence techniques, the theory of differential equations - these directions allow you to build a model for predicting system accidents.

The theory of database management (non-relational databases) - these systems allow you to establish the optimal storage of large time series.

ANN modeling is done in two stages; the first step is to train the network and the second is to test the network with data that is not used for training. Although the training is time consuming, they are very quick to resolve any operation. ANN is trained with a suitable training method to perform a specific function by adjusting the weight values. The process continues until the network output matches the output. Changing weights and biases reduces the error

between the mains output and the desired output. The training process ends when the error falls below a certain value. The network uses a learning mode in which inputs are presented to the network along with the desired result and the weights are adjusted so that the network tries to produce the desired output.

The network architecture selected using input and output parameters is shown in Figure 3. The input parameters for this present work are solar radiation received per unit area, heat pump blower and ambient temperature. Solar radiation is the energy of solar irradiation that falls on a flat solar collector to heat water appropriately over time in a specific cycle of operation. The second input is the heat pump blower and the real work is done for a delivery depth of 0.5m. The third input parameter, the ambient temperature is also recorded for each heating time value. The output parameters in this work are the heating time for each cycle of operation, the efficiency of the heat pump and the flow from the system.

Mathematically, the efficiency of the system can be written in the following form,

$$\eta = \frac{\rho_w * V_w * g * H}{I_\beta} \quad (1)$$

where H the head in m, V_w is the volume of raised water and I_β is the intensity of solar radiation in W / m².

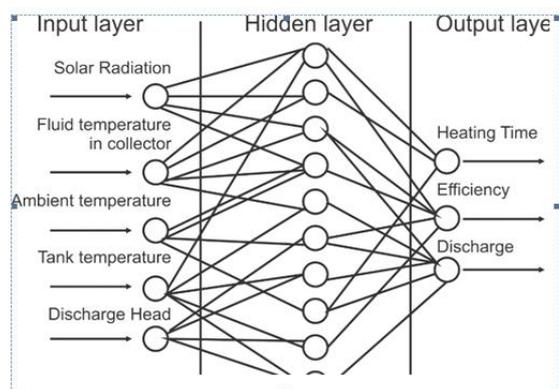


Figure 3. Network architecture for a flat panel solar collector with a heat pump

The obtained experimental data is used to train the network.

Here, 100 datasets are used to train the network. Input data is divided into training sets (70%), test kits (15%) and testing sets (15%). This simulation used a single hidden layer learning algorithm. By lowering the propagation of error from the output layers to the bottom layer, the algorithm optimizes the synaptic weight (hidden layer and input layer). The ANN model approximates the desired result from each iteration so that errors are reduced and the errors are then passed to the ANN to adjust the weights. The network is trained to achieve the selected error target. In this work, the sigmoid transfer function Tan is used as the activation function for the buried layer and the output layer.

$$RMSE = \sqrt{\frac{(X_i - Y_i)^2}{N}} \quad (2)$$

$$CC = \frac{N \sum (X_i * Y_i) - \sum X_i \sum Y_i}{\sqrt{(N \sum X_i^2 - (\sum X_i)^2)(N \sum Y_i^2 - (\sum Y_i)^2)}} \quad (3)$$

$$APE = \frac{1}{N} \sum \left(\frac{abc(X_i - Y_i)}{X_i} \right) * 100 \quad (4)$$

where are X_i – experimental values and Y_i – ANN predicted output values.

III. RESULTS

In this study, experiments with a heat pump were carried out for a depth of 0.5 m. Experimental work was carried out continuously for one week. The flat plate solar collector was exposed to radiation every day from 7:00 am and the heat pump was ready to perform its first cycle by 11:00 am.

Figure 8 shows the time variation of solar radiation on a sunny day. The maximum solubility of solar radiation per day with experimental data was between 11:00 and 17:00.

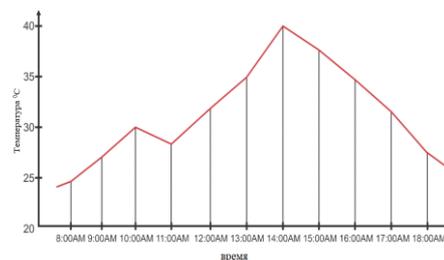


Figure 4. Change in solar radiation over time on a sunny day

The heat pump gives the best performance during this time interval and the experimental results are exactly the same. The heating time at noon is much shorter compared to other times.

Figure 5 shows the change in ambient temperature over time. The graph shows the data of the solar radiation curve, which is the solar radiation, which directly affects the ambient temperature.

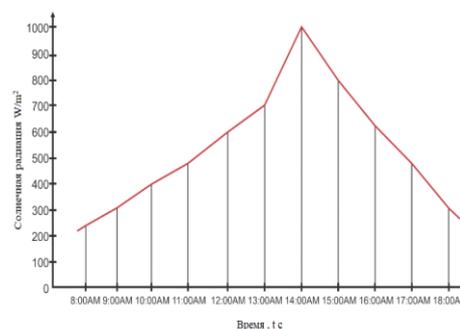


Figure 5. Change in ambient temperature over time

Figure 6 shows the exact change in pressure over time when the heat pump is at a depth of 0.5 m and fluctuates between the two limits.

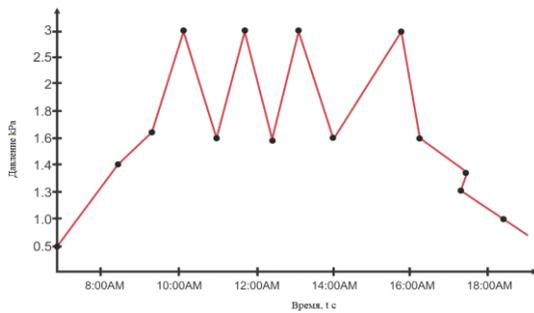


Figure 6. The exact change in pressure over time

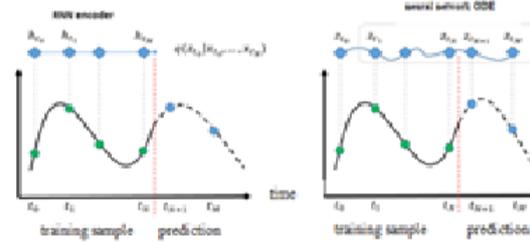
Table 2. Number of cycles per day, consumption and efficiency

No	Pressure (m)	Number of cycles per day	Unloading (liter/day)	Efficiency
1	0,5	6	30	0.16

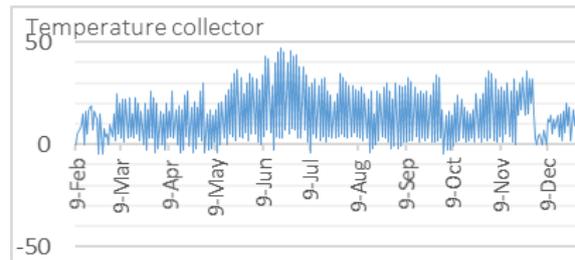
Table 2 shows the number of cycles per day, flow rate and operating efficiency performed for the heat pump. The efficiency depends on the number of cycles per day and the performance of the heat pump.

The efficiency depends on the number of cycles per day and the performance of the heat pump.

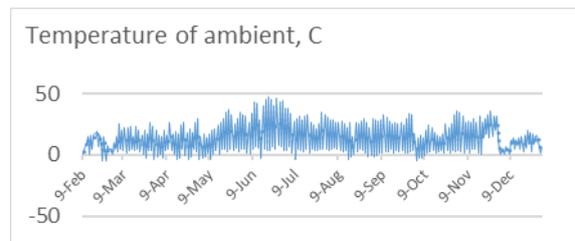
Considering the problem of time series prediction, they mainly use a recurrent neural network (RNN). The principle of operation of RNN is as follows: the training sample is fed to the input of the RNN (encoder), the result will usually be a normal distribution, on the basis of which a random variable is generated $z(t_0) : q(z(t_0) | \mu, \sigma)$, where the μ, σ mathematical expectation and standard deviation of the normal distribution. Then $z(t_0)$ another RNN (decoder) is fed to the input $z(t+1) = z(t) + f(z(t), t)$, which $f(z(t), t)$ restores the values from the test data. This neural network uses discrete transformations, where is some non-linear transformation. This transformation coincides with the ODE scheme of Euler's methods, in fact, you can change the direction from a discrete transformation to a continuous trajectory $z(t)$. This replacement allows you to create a new architecture for neural network ODEs. Figure 7 shows a diagram of the operation of neural network ODEs and its difference from RNN.



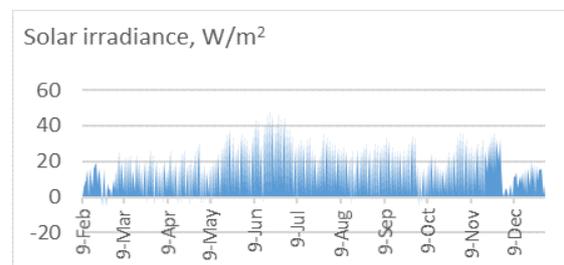
a)



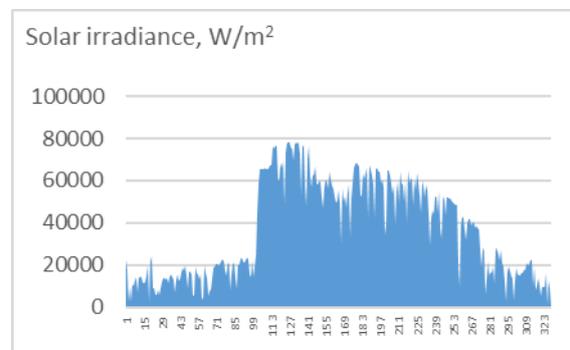
b)



c)



d)



d)

Figure 8. (a,b,c,d) simulation result comparing ambient temperature and solar irradiance

A flat solar collector was chosen for the verification; data were collected for 3 months of observation from technical devices. Temperature sensors were installed in the

accumulator tank, at the outlet from the flat solar collector, voltage sensors on the pump and vibration.

Based on data from sensors on temperature and pressure, we can predict future problems in the installation, such as pump breakdown, depressurization, and more.

Additionally, environmental data were collected for one year. Figure 9 shows the high correlation between solar radiation and temperature in the collector and reservoir. If the density of solar radiation is higher, then the temperature is more monotonic, and the standard deviation of the temperature in the storage is close to the average value, for example, from July to August.

Figure 9 shows the temperatures in the manifold, reservoir and pump operating mode. We see the correlation between the ambient temperature and the temperature of the heated water. When the temperature in the manifold reaches 40 ° C, the pump switches over and starts transferring the heated liquid to the heat exchanger inside the storage tank and receives the chilled liquid in the manifold cycle for heating. The second reason for the cooling of the fluid in the cycle is associated with weather conditions (decrease in ambient temperature, sunset).

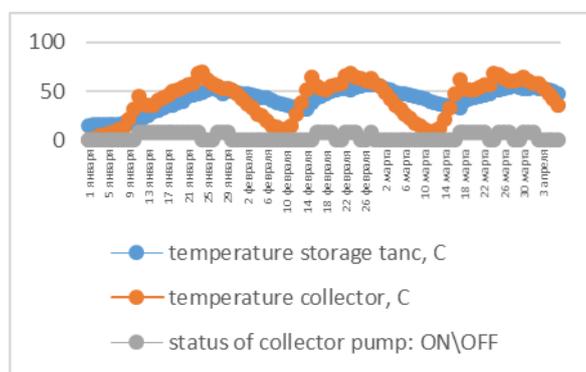


Figure 9. Ambient and manifold and reservoir temperatures with pump operating mode

To assess the accuracy, the model was trained on historical data to capture the seasonality and trend of the time series. The model prediction has been made over the last month. The RSME metric was used to assess the accuracy of the model.

Verification of the considered model has shown good results in predicting the future behavior of the pump load. In Figure 10, you can see that the model predicted the future temperature of the water from the loop and there are no large deviations.

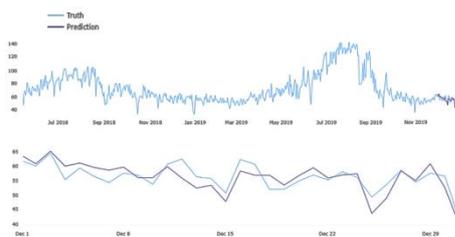


Figure 10. Predicting the temperature in the solar collector circuit

Figure 11 shows the prediction of the temperature in the solar collector circuit. We can observe the difference

between the predicted temperature in the model and the actual one. This may be a signal that there are deviations from the standard temperature indicators in the system. The model prompts you to inspect the installation.

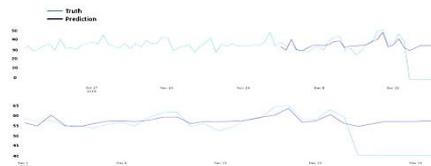


Figure 11. Abnormal deviation of temperature indicators in the battery tank

The model also showed abnormal deviations in the pressure (bar) readings in the unit loop. In Figure 12, you can see large deviations between the predicted values and the original readings from the sensors.

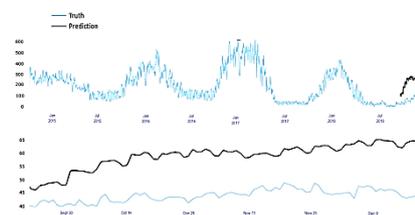


Figure 12. Abnormal deviation in atmospheric pressure reading in the installation

IV. CONCLUSION

The created model was modeled entirely for a solar heating system. On the basis of a mathematical model of a solar flat collector, a block diagram was developed, a block diagram of the collector model. The result of the simulation is the temperature of the fluid leaving the reservoir. A flat solar collector was chosen for the verification; data were collected for 3 months of observation from technical devices. Temperature sensors were installed in the accumulator tank, at the outlet from the flat solar collector, voltage sensors on the pump and vibration. Based on data from sensors on temperature and pressure, we can predict future problems in the installation, such as pump breakdown, depressurization, and more. The difference between the predicted temperature in the model and the actual temperature was observed. This may be a signal that there are deviations from the standard temperature indicators in the system. The model prompts you to inspect the installation. The model also showed abnormal deviations in the pressure (bar) readings in the unit loop. You can also notice large deviations between the predicted values and the original readings from the sensors.

References

[1] RENA, 2018. Opportunities to accelerate national energy transitions through advanced deployment of renewables

[2] Niu, G., 2017. Data-Driven Technology for Engineering Systems Health Management. Springer Singapore, Singapore.

[3] Lui, Y., Zhao, G., Peng, X., Hu, C., 2017. Lithium-ion battery remaining useful life prediction with long short-term memory recurrent neural network. Annu. Conf. Progn. Heal. Manag. Soc. 1–7.

- [4] de Keizer, A.C., Vajen, K., Jordan, U., 2011. Review of long-term fault detection approaches in solar thermal systems. *Sol. Energy* 85, 1430–1439.
- [5] Ghritlahre, H.K., Prasad, R.K., 2018. Application of ANN technique to predict the performance of solar collector systems - A review. *Renew. Sustain. Energy Rev.* 84, 75–88.
- [6] Kalogirou, S., Lalot, S., Florides, G., Desmet, B., 2008. Development of a neural network-based fault diagnostic system for solar thermal applications. *Sol. Energy* 82, 164–172.
- [7] Lipton, Z.C., Berkowitz, J., Elkan, C., 2015. A Critical Review of Recurrent Neural Networks for Sequence Learning. *Proc. ACM Int. Conf. Multimed. - MM '14* 675–678.
- [8] Srivastava, S., Lessmann, S., 2018. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. *Sol. Energy* 162, 232–247.
- [9] Leva, S., Dolara, A., Grimaccia, F., Mussetta, M., Ogliari, E., 2017. Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power. *Math. Comput. Simul.* 131, 88–100.
- [10] Correa-Jullian, C., Cardemil, J.M., Droguett, E.L., Behzad, M., 2019. Assessment of Deep Learning Techniques for Prognosis of Solar Thermal Systems. *Renew. Energy* 145, 2178–2191.
- [11] Kalogirou, S.A., Agathokleous, R., Barone, G., Buonomano, A., Forzano, C., Palombo, A., 2019. Development and validation of a new TRNSYS Type for thermosiphon flat-plate solar thermal collectors: energy and economic optimization for hot water production in different climates. *Renew. Energy* 136, 632–644.
- [12] Souliotis, M., Kalogirou, S., Tripanagnostopoulos, Y., 2009. Modelling of an ICS solar water heater using artificial neural networks and TRNSYS. *Renew. Energy* 34, 1333–1339.
- [13] Keshuov, S., Omarov, R., Tokmoldayev, A., Omar, D., Kunelbayev, M., Amirseit, S. Hybrid system for using renewable sources of energy for local consumers in agriculture 2017, *Journal of Engineering and Applied Sciences*, (5), 1296-1306
- [14] Djamila Bouhalouan, Bakhta Nachet, Abdelkader Adla, "A Knowledge-Intensive Decision Support System for Industrial Machines Maintenance", *WSEAS Transactions on Information Science and Applications*, pp. 41-52, Volume 17, 2020
- [15] Mahmoud Mohamed El-Ghobashy El-Hagar, "Simple and Efficient Methods to Perform the Optimum Matching between a Reciprocating Piston Pump and a Wind Machine", *WSEAS Transactions on Applied and Theoretical Mechanics*, pp. 216-221 Volume 15, 2020.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Murat Kunelbayev, Abdildayeva Assel has organized and executed the experiments. Taganova Guldana carried out the simulation results, interpretation (discussion) and verification of results.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US