An Energy-Based One Step Ahead of State Prediction with LSTM Model

Milovan Medojevic[*, **]
[*] The Institute for Artificial Intelligence Research and Development of Serbia, Novi Sad, Serbia
[**] EnergyPulse d.o.o, Novi Sad, Serbia
milovan.medojevic@ivi.ac.rs

***Abstract*— In this paper, several LSTM model structures were developed, analyzed, and evaluated for being deployed to predict the future state of the CNC machine tools, one step ahead in terms of energy consumption. For the modeling purpose, self-generated datasets regarding the energy consumption of the CNC milling machine tool were applied for model training and validation in a ratio of 80/20 respectively, while the model evaluation was performed upon the CNC lathing machine tool energy consumption dataset. Furthermore, model rankings were performed based on PROMETHEE (Preference Ranking Organization METHod of Enrichment Evaluation) method which enabled the selection of the most acceptable one in the finite set of alternatives (model structures), linked to multiple criteria considered. All of the generated models generalize the observed problem quite well and can be applied according to the task that has been set.**

I. Introduction

Forecasting energy use is considered a vital and challenging task, in which accurate forecasting can provide valid guidelines for both resource allocation [1] and formulating measures to increase energy efficiency [2], as well as improving the functionality of energy and manufacturing systems through the implementation of effective management decisions [3]. Consequently, the developed forecasting models can be applied in various areas wherever time series appear, such as climate change [4], temperature [5] and solar radiation [6] forecasting, traffic flows [7], inventory changes [8], and so forth. Moreover, fast and accurate forecasting stands as a primary tool in the field of energy-based predictive maintenance. Having this in mind, a variety of research emerges worldwide in the quest for intelligent model development aimed to improve the accuracy of prediction where recently, deep learning became the most used field in prediction [9]. There are studies, in which the long short-term memory (LSTM) neural network was proposed to predict the acceleration of the preceding vehicle by using data from onboard radar sensors [10]. Furthermore, towards capturing the short-term and long-term temporal modes and achieving the day-ahead hourly photovoltaic power forecasting an attention-based long-term and short-term temporal neural network prediction model (ALSM) assembled using the convolutional neural network (CNN), (LSTM), and attention mechanism under the multiple relevant and target variables prediction pattern (MRTPP) was developed [11]. Also, due to the increasing use of solar-assisted water heating systems, to predict energy use, the LSTM neural network, enhanced with the attention mechanism inspired by human vision to pay selective attention to the input data was proposed [12]. Similarly, to cope with the changes and complexity that renewable energy sources bring to the electricity markets, the LSTM-based model was suggested for power grid loss prediction to overcome the limited capability of currently applied methods in the field [13]. LSTM network concepts were also applied for the thermal error prediction which has been considered a dominant factor that seriously hinders the high-accuracy machining of complex parts to overcome the main barriers of data-based models such as weak robustness and low predictive accuracy.

A. Motivation and problem formulation

In this paper, the approach based on the identification of energy flows provides a fundamental basis for the determination of the states and behavior of CNC machine tools within manufacturing systems, while variable determined in this way combines the effects of both, forces and velocities, and their product, power or intensity of energy change, characterized by the dynamic property, includes and reflects complete information about its balance and movement, and therefore exceeds the study of changes in force and motion separately [14].

Having this in mind, several LSTM-based model structures were developed, analyzed, and evaluated to predict the future state of the CNC machine tools, one step ahead in terms of their power draw.

The preferred model could be used as a data-driven prognostics approach [15] that enables the identification of trends/patterns of a developing fault, which, in a defined timeframe, predicts when a predetermined threshold would be reached by using information from historically treated data (trained data).

II. Methodology and obtained Results

A. Model development

In this case, the observed variable is the CNC machine power draw, where the sampling frequency corresponds to one second. A subset (a series of 31 consecutive values) is taken, with the first 30 being used as input to the neural network, while the last 1 is the value the network needs to learn during training. In the training phase, the network learns to predict the next 1 based on 30 input sequences, while in the validation phase, the network should be able to predict the next 1 value based on 30 input sequences of the same dataset. In the evaluation phase, the network should be able to predict the next 1 value based on 30 input sequences of the dataset that represents different machine behavior.

The model development was performed within the Google Colaboratory environment. The initial model structure considers 2 LSTM (tanh activation) and 5 Dense (linear activation) layers, where ADAM is defined as an optimizer, losses are monitored through MSE, MAE and RMSE, while training was scheduled to save the best only epoch in terms of MSE on validation data (Vmse) with patience = 3 and minimum delta = 0.001. Training and validation were performed on a self-generated CNC milling machine tool energy consumption dataset [16] in a ratio of 80/20 respectively, while the generated models were additionally evaluated on a CNC lathing machine tool energy consumption dataset. To perform a controlled experiment in the first iteration, variations were performed in the number of units of the first LSTM layer as given in Table 1.

TABLE I.
VARIATIONS IN THE 1ST LSTM LAYER OF INITIAL EXPERIMENTS

	Units No.			
Layer	EXP1	EXP2	EXP3	EXP4
Input	1			
LSTM	150	250	50	100
LSTM	250			
Dense	256			
Dense	128			
Dense	64			
Dense	32			
Dense	1			

B. *Model rankings and obtained results*

To determine the most favorable model structure, model rankings were performed based on PROMETHEE method which enabled the selection of the most acceptable model architecture in the given finite set of architectures. The criteria considered for rankings are numerical values of MSE, MAE and RMSE generated during model training, validation, and evaluation (Tmse, Tmae, Trmse, Vmse, Vmae, Vrmse, Emse, Emae, Ermse) as well as the amount of time necessary to perform prediction (PredictTime). The second iteration of the controlled experiment considered variations in applied activation functions to the each of existing layers of the previously selected model architecture (Table 2). Similar to the previous iteration, model rankings were performed based on PROMETHEE method.

TABLE II.
VARIATIONS IN APPLIED ACTIVATION FUNCTIONS TO EACH OF THE EXISTING LAYERS

	Model architecture (EXP1)					
	original	a	b	c	d	e**
LSTM	tanh	tanh	relu	relu	tanh	tanh
LSTM	tanh	tanh	relu	relu	Sigmoid*	tanh
DENSE	linear	relu	relu	linear	relu	linear
DENSE	linear	relu	relu	linear	relu	linear
DENSE	linear	relu	relu	linear	relu	linear
DENSE	linear	relu	relu	linear	relu	linear
DENSE	linear	relu	relu	linear	relu	linear
*recurrent activation, **dropout = 0.1						

Table 3 summarizes key criteria for overall model architecture ranking, while the preference results are given in figure 1. Moreover, all of the given criteria are treated equally in terms of preference, except for Pred.Time that could vary due to the changes in hardware infrastructure utilized to train, validate and evaluate the models.

TABLE III.
KEY CRITERIA AND PREFERENCE PARAMETERS FOR MODEL STRUCTURE EVALUATION

	Key criteria									
Model	Tmse	Tmae	Trmse	Vmse	Vmae	Vrmse	Emse	Emae	Ermse	Pred.Time
EXP1	0.4916	0.2508	0.7012	0.4929	0.2442	0.7021	0.6131	0.2389	0.7830	2.235533
EXP2	0.5504	0.2771	0.7419	0.5977	0.2708	0.7731	0.8324	0.2681	0.9124	2.555646
EXP3	0.5999	0.3124	0.7745	0.6791	0.3198	0.8241	0.6470	0.2754	0.8044	2.126676
EXP4	0.6603	0.2995	0.8126	0.6792	0.2907	0.8241	0.5835	0.2670	0.7639	1.207803
EXP1a	0.5853	0.2521	0.7650	0.5679	0.2674	0.7536	0.8071	0.2844	0.8984	1.399784
EXP1b	0.5029	0.2224	0.7092	0.5624	0.2525	0.7499	0.8347	0.2595	0.9136	2.137376
EXP1c	0.5732	0.2707	0.7571	0.6402	0.2982	0.8001	0.6063	0.2417	0.7787	2.727566
EXP1d	0.4931	0.2260	0.7022	0.5315	0.2320	0.7290	0.7158	0.2398	0.8460	2.180069
EXP1e	0.6465	0.2859	0.8040	0.6641	0.3221	0.8149	1.1415	0.5049	1.0684	2.111090
Preference parameters										
Min/Max	min	min	min	min	min	min	min	min	min	min
Weight	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,50
Pref. Fn.	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	U-shape	Usual
Thresholds	%	%	%	%	%	%	%	%	%	%
Q: Indiff.	2,5%	2,5%	2,5%	2,5%	2,5%	2,5%	2,5%	2,5%	2,5%	n/a
P: Pref.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
S: Gaussian	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

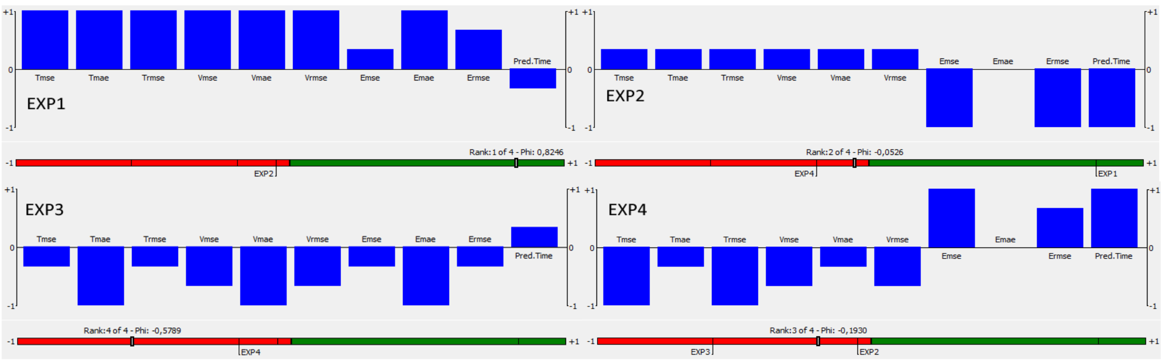

Figure 1. Preference results given in normalized scale

As could be seen from figure 1, the model generated within EXP1 is characterized by the highest preference among all key criteria, and therefore this model was selected for further optimization in the second iteration. Also, it is important to mention that model generated within the EXP4 provided results of high preference in terms of evaluation (Emse and Emae) although its training and validation did not result in the desired performance. To illustrate the overall preference indicator (PI) considering the key criteria given, GAIA webs were deployed (Figure 2).

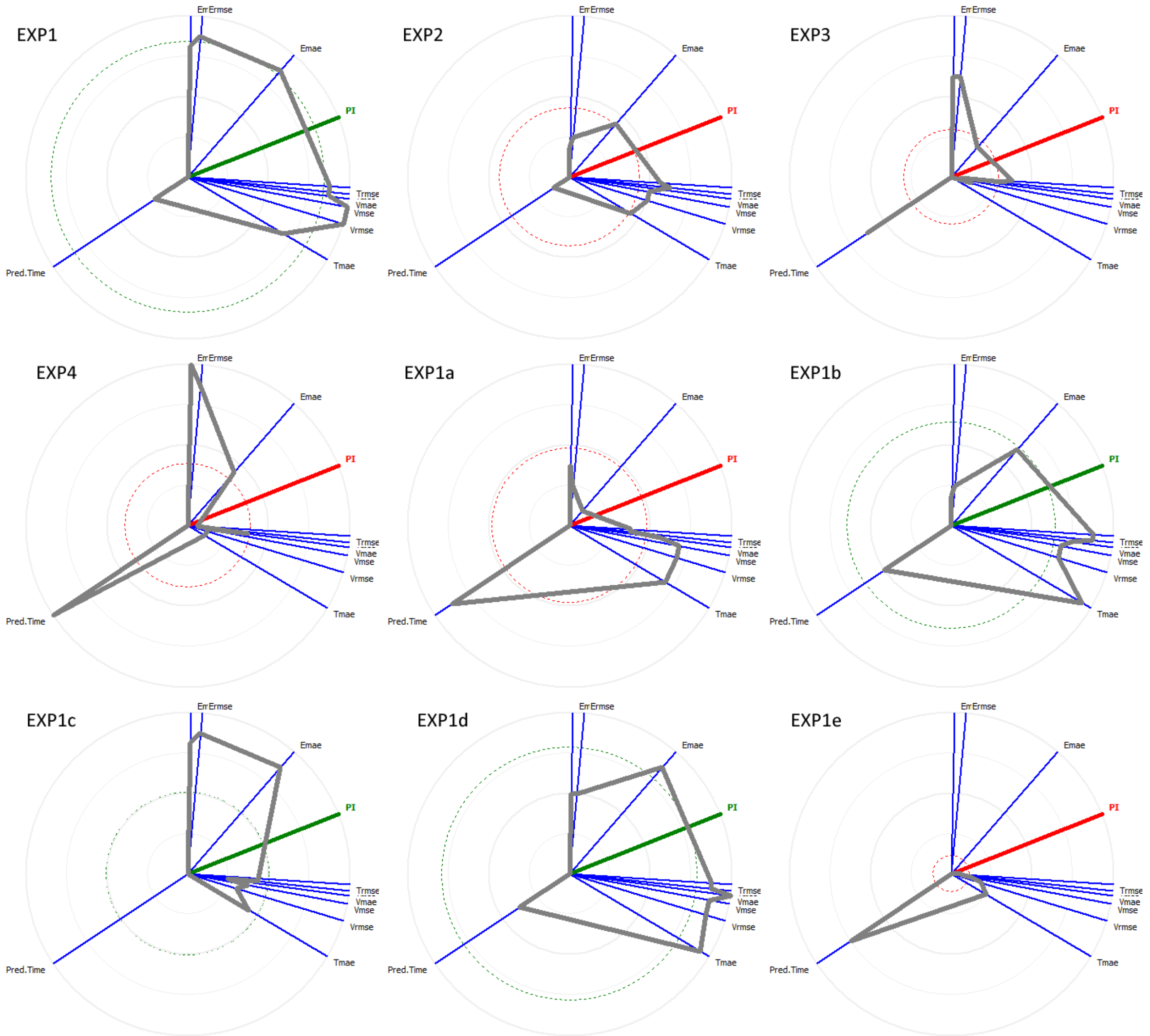

Figure 2. Overall model evaluation in terms of key criteria via GAIA webs

Here, the GAIA web tool is used to provide an enhanced spiderweb display for one action (model) enabling a graphical representation of the unicriterion net flow scores for the selected action. The variables (criteria) are equally spaced around the center of the display indicating that the shape of the spiderweb is highly dependent on the arbitrary order of the criteria. Also, the criteria axes are oriented as in the plane (the GAIA plane) where the criteria expressing similar preferences are placed close to each other enabling more meaningful visuals through generated spiderweb shape. Finally, the decision axis as well as a dotted circle corresponding to the multicriteria net flow score of the actions is visible in green if PI score of an analyzed action is positive or in red if it is negative. This enables a general overview of observed models in terms of normalized preference among the set criteria. The final rankings are shown in the form of a normalized scale given in figure 3.

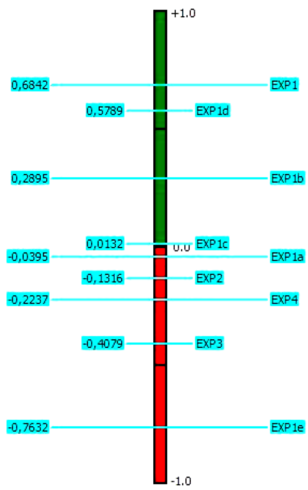

Figure 3. Overall model rankings in terms of net flow scores given in the normalized scale

From figure 3 it could be concluded that models generated within experiments EXP1, EXP1d, EXP1b, and EXP1c have positive PI scores while the model within EXP1 remained characterized with the highest preference after several optimization procedures during the second iteration of controlled experimentation. Therefore, this model architecture is given in figure 4 due to its highest preference results and favoritism for deployment into the production process.

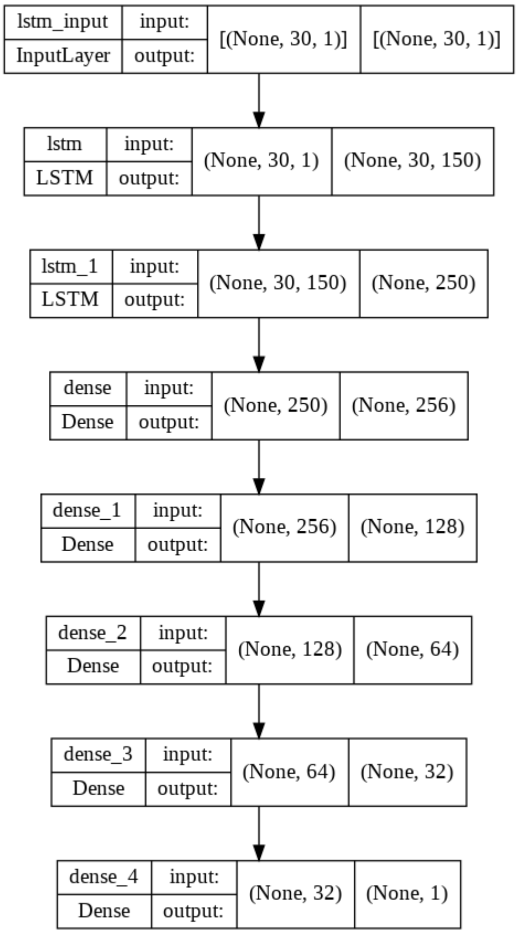

Figure 4. EXP1 model architecture with layer characteristics

C. *Model deployment*

In the subsequent phase model with the highest preference (EXP1) was selected for deployment. The deployment results are given in the figure 5. Figure 5 illustrates the model evaluation phase only where it turned out that it generalizes the observed problem quite well and can be applied according to the task that has been set.

The results (Table 3) showed that the prediction effect of the chosen model during evaluation on the dataset for CNC lathe was better than the one generated during validation phase for CNC mill. Also the prediction effect for both, the high and low power draw periods turned out to be better in evaluation phase.

However, it is necessary to stress out that in practical application, the prediction effect of this model for different operating modes could diverse, which in depth indicates that mechanism and solutions need to be further analyzed.

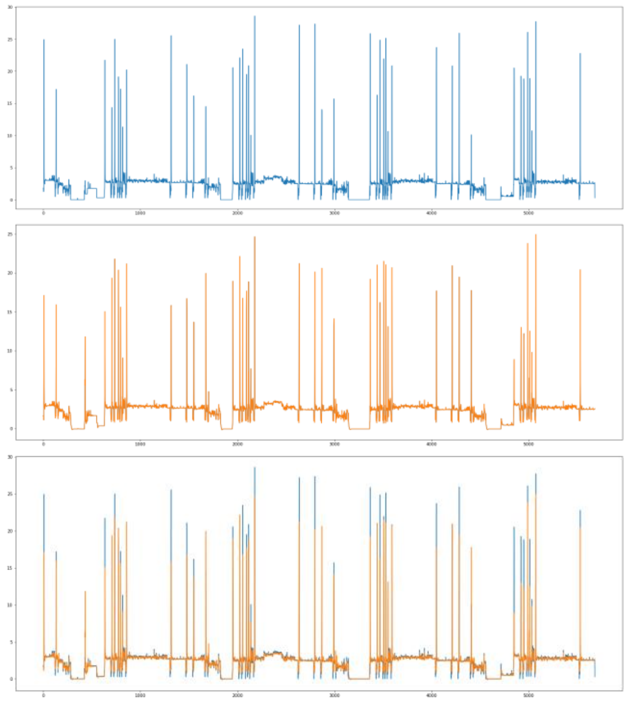

Figure 5. (1) Actual behavior profile from the validation set of data for HAAS SL 20 HE, (2) Predicted behavior profile of the validation data set for HAAS SL 20 HE, and (3) Comparative profiles of actual and predicted behavior of the validation data set for HAAS SL 20 HE

III. CONCLUSION

In this paper, a deep neural network has been designed specifically to predict power draw time series. The LSTM network architecture has been proposed due to its ability to deal with sequential data as its main characteristic is memory for retaining temporal relationships in the long term. A quest to find the best values of the hyper parameters such as number of layers, number of LSTM cells for layer, learning and so on, have been carried out in two iterations, both within the controlled environment. Finally, several LSTM model structures were developed, analyzed, and evaluated for being deployed to predict the future state of the CNC machine tools, one step ahead in terms of energy consumption. Self-generated dataset of the CNC milling machine energy consumption was applied for model training and validation, while the model evaluation was performed on the energy consumption dataset for CNC lathing machine.

In order to select the most preferred model architecture PROMETHEE method was applied. The method revealed that models generated within experiments EXP1, EXP1d, EXP1b, and EXP1c have positive PI scores while the model within EXP1 remained characterized with the highest preference after several optimization procedures during the second iteration of controlled experimentation. The generated model generalize the observed problem quite well (Table 3) and can be applied according to the task that has been set.

Future work will be directed towards the fusion of different deep learning models and novel techniques in order to exploit the different advantages of each of them with aim to obtain enhanced predictions for different this and other real-world problems.

REFERENCES

[1] J. Liu, Y. Chen, J. Zhan, and F. Shang, "An On-Line Energy Management Strategy Based on Trip Condition Prediction for Commuter Plug-In Hybrid Electric Vehicles," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 3767–3781, May 2018.

[2] B. Han, D. Zhang, and T. Yang, "Energy consumption analysis and energy management strategy for sensor node," in *Proceedings of the 2008 IEEE International Conference on Information and Automation, ICIA 2008*, 2008, pp. 211–214.

[3] H.-T. Pao, "Forecast of electricity consumption and economic growth in Taiwan by state space modeling," *Energy*, vol. 34, no. 11, pp. 1779–1791, 2009.

[4] S. Rasp and S. Lerch, "Neural Networks for Postprocessing Ensemble Weather Forecasts," *Mon. Weather Rev.*, vol. 146, no. 11, pp. 3885–3900, Nov. 2018.

[5] Q. Zhang, H. Wang, J. Dong, G. Zhong, and X. Sun, "Prediction of Sea Surface Temperature Using Long Short-Term Memory," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1745–1749, Oct. 2017.

[6] F. Nomiyama, J. Asai, T. Murakami, and J. Murata, "A study on global solar radiation forecasting using weather forecast data," in *2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2011.

[7] H. Yin, S. C. Wong, J. Xu, and C. K. Wong, "Urban traffic flow prediction using a fuzzy-neural approach," *Transp. Res. Part C Emerg. Technol.*, vol. 10, no. 2, pp. 85–98, Apr. 2002.

[8] Y. Zuo and E. Kita, "Stock price forecast using Bayesian network," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 6729–6737, Jun. 2012.

[9] A. Salam and A. El Hibaoui, "Energy consumption prediction model with deep inception residual network inspiration and LSTM," *Math. Comput. Simul.*, vol. 190, pp. 97–109, Dec. 2021.

[10] B. Tian, G. Wang, Z. Xu, Y. Zhang, and X. Zhao, "Communication delay compensation for string stability of CACC system using LSTM prediction," *Veh. Commun.*, vol. 29, p. 100333, Jun. 2021.

[11] J. Qu, Z. Qian, and Y. Pei, "Day-ahead hourly photovoltaic power forecasting using attention-based CNN-LSTM neural network embedded with multiple relevant and target variables prediction pattern," *Energy*, vol. 232, p. 120996, Oct. 2021.

[12] A. Heidari and D. Khovalyg, "Short-term energy use prediction of solar-assisted water heating system: Application case of combined attention-based LSTM and time-series decomposition," *Sol. Energy*, vol. 207, pp. 626–639, Sep. 2020.

[13] J. Tulensalo, J. Seppänen, and A. Ilin, "An LSTM model for power grid loss prediction," *Electr. Power Syst. Res.*, vol. 189, p. 106823, Dec. 2020.

[14] M. S. Medojević, M. M. Medojević, and P. D. Villar, "Simulation-based design of solar photovoltaic energy generation system for manufacturing support," *Therm. Sci.*, vol. 25, no. 4 Part A, pp. 2517–2538, 2021.

[15] M. Medojevic, P. Díaz Villar, I. Cosic, A. Rikalovic, N. Sremcev, and M. Lazarevic, Energy management in industry 4.0 ecosystem: A review on possibilities and concerns," in *Proceedings of the 29th DAAAM International Symposium*, 2018, pp. 674–0680.

[16] M. M. Medojević, B. B. Tejić, M. S. Medojević, and M. V. Kljajić, "Design and development of IIoT-based system for behavior profiling of nonlinear dynamic production systems based on energy flow theory," *Therm. Sci.*, vol. 26, no. 3 Part A pp. 2147–2161, 2022.