E3S Web of Conferences **411**, 01007 (2023) https://doi.org/10.1051/e3sconf/202341101007
APEC-VI-2023

Pre-training method in the tasks of obtaining surrogate models of gas turbine units for gas turbine electric power stations

Grigory Kilin[1], *Boris* Kavalerov[1], *Artem* Suslov[1*], and *Ilya* Tyatenkov[1]

[1]Perm National Research Polytechnic University, 29, Komsomolsky prospect, Perm, 614990, Russian Federation

Abstract. This article focuses on the application of pre-training methods in the task of synthesizing surrogate models. The article emphasizes that pre-training significantly improves the accuracy of surrogate models and speeds up their creation process. The authors examine pre-training's impact on various aspects of surrogate modeling of a gas turbine unit that is part of a gas turbine electric power station, such as reducing computational costs, improving the approximation of complex processes, and optimizing the model synthesis procedure. The work demonstrates specific examples that clearly show how the use of pre-training can significantly improve the performance of surrogate models and optimize the development process. Thus, the authors convincingly argue that pre-training is a key tool for increasing the efficiency of surrogate modeling, capable of significantly reducing the time, costs, and efforts required for the development and use of surrogate models in the energy sector.

1 Introduction

Surrogate models play an important role in modern scientific research as they allow for the approximation of complex and computationally expensive processes using simpler and faster methods [1-2]. One of the definitions of a surrogate model is an emulation of the output (response) of some black-box [1]. In other words, building a model based on available experimental data or data obtained from a complex adequate model.

In this article, authors review the synthesis of surrogate models by comparing methods of learning an artificial neural network. Real world application for surrogate models in the energy sector is a gas turbine unit (GTU) that is part of a gas turbine electric power station (GTEPS). The data used to train neural networks collected by computer simulation of a GTEPS operating at a dedicated load. The authors review a GTU, since:

- It in itself is already a rather difficult element to study due to non-linear dependencies.
- Such units are one of the main components of GTEPS, because they drive synchronous generators (SG).

Synthesizing efficient surrogate models is a challenging task that requires a large amount of data and time. In this context, the use of artificial neural networks [3-4],

* Corresponding author: AISuslov@pstu.ru

especially the possibility of using pre-training [5], is an innovative approach that improves the accuracy and reduces the synthesis time of such models.

The article demonstrate research results that confirm the advantages of pre-trained neuromodel over baseline neuromodel in synthesizing surrogate models. Advantages include increased neuromodel accuracy and reduced synthesis time of a surrogate model. The baseline neuromodel refers to a conventional feedforward neuromodel.

In conclusion, authors discuss possible directions for further development in the energy sector.

2 Materials and methods

The term pre-training refers to the process of training an artificial neural network on a large dataset before using it for a specific task. This allows the model to extract generalized knowledge and improve its performance when fine-tuning on specific data.

In previous studies, pre-training was performed on a small but sufficient dataset [6], consisting of a few thousand points that represented GTEPS operating at a dedicated load in one and several modes of operation. Dataset included data for elements of a GTEPS such as a GTU and a SG. However, in order to synthesize a pre-trained neuromodel of an energy system (or at least a single GTU or a single GTEPS) the volume of a dataset should be several orders of magnitude larger.

Thus, the study began with the merging of experimental data from all previous GTEPS experiments into one dataset. Moreover, the amount of data increased by conducting additional experiments. The resulting dataset includes 971532 points for each variable, which is two orders of magnitude larger than the previously used experimental data.

In order to train hidden layers the implementation algorithm for the pre-training method involve the use of an autoencoder [7-8] with a single hidden layer (Figure 1). The formation of a pre-trained neural network is based on these hidden layers, but with a pre-defined number of neurons in the input and output layers. The number of hidden layers and neurons in them can be arbitrary.

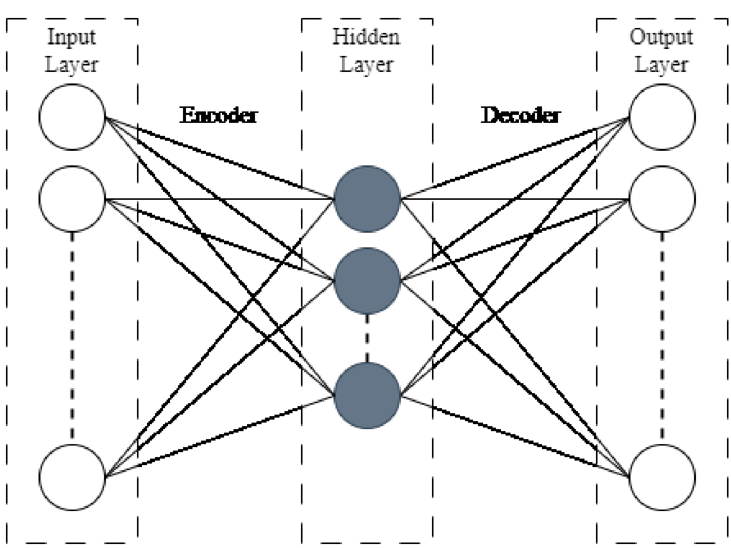

Fig. 1. Autoencoder structure.

Pre-trained neural network compared with the baseline neural network. The baseline neural network formed with the same architecture and hyperparameters as a pre-trained neural network, but it uses feedforward method to learn.

3 Results and Discussion

The synthesis of a pre-trained neural network reviewed in this article concluded with these steps (Figure 2):

- Performance of three training procedures based on the collected dataset. Each training procedure included neural network based on an autoencoder consisting of 5 input neurons, a single hidden layer with 20 neurons, and 5 neurons in the output layer.
- Selection of three neuromodels with the smallest training error from each training.
- Afterwards, the hidden layers were extracted from these autoencoder networks and used to create a pre-trained neural network consisting 3 input neurons, three hidden layers with 20 neurons in each hidden layer, and 5 neurons in the output layer.

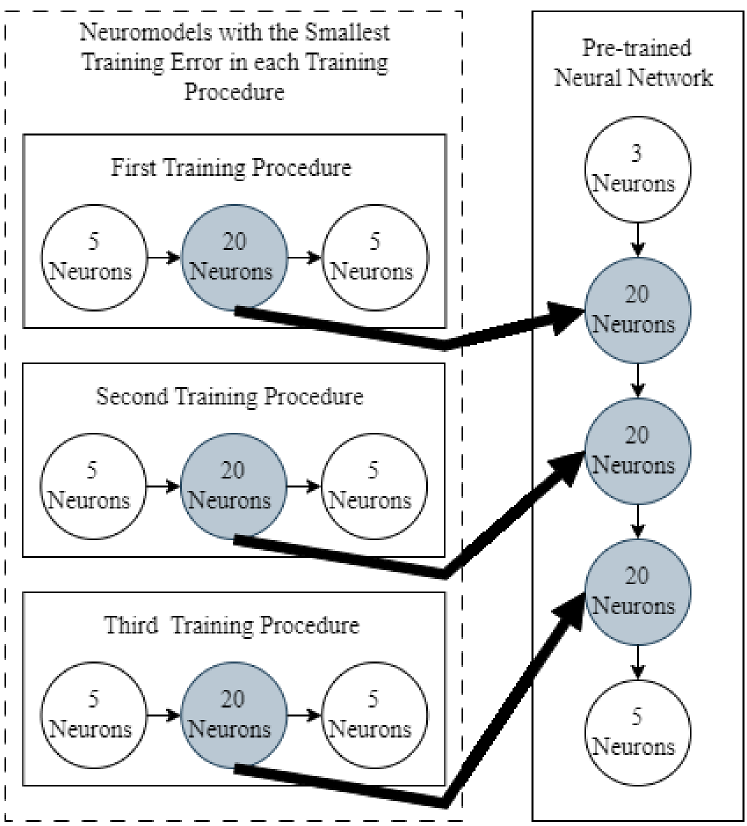

Fig. 2. The synthesis of a pre-trained neural network.

Lastly to confirm the advantages of pre-trained neural network over baseline neural network in synthesizing surrogate models, these neural networks were trained on a data from computer modeling of the one GTEPS operation at a dedicated load of a 1000 kW (Figure 3) and subsequent load increase to 2000 kW.

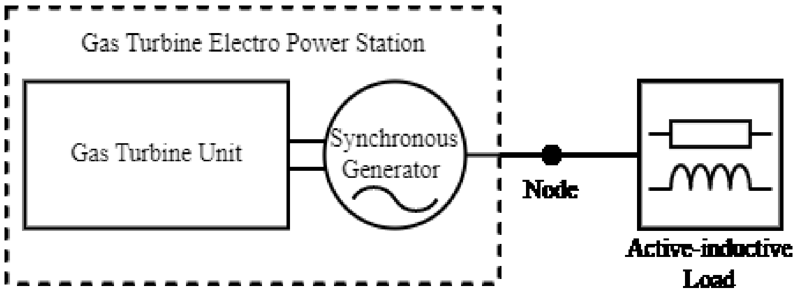

Fig. 3. Diagram of a gas turbine power station operating at a dedicated load.

The baseline neural network had the same architecture and hyperparameters as the pre-trained neural network, except that its three hidden layers were not pre-trained, but the weights of their neurons were randomly initialized at the beginning of training.

Figure 4 shows that the errors decreased during all training process. Additionally, it is noticeable that by continuing training procedures the training errors can decrease. The training process on the large dataset took a considerable amount of time. Therefore, the necessary and sufficient number of training iterations were selected to demonstrate the validity of the assumption that pre-training increases neuromodel accuracy and reduced synthesis time of a surrogate model.

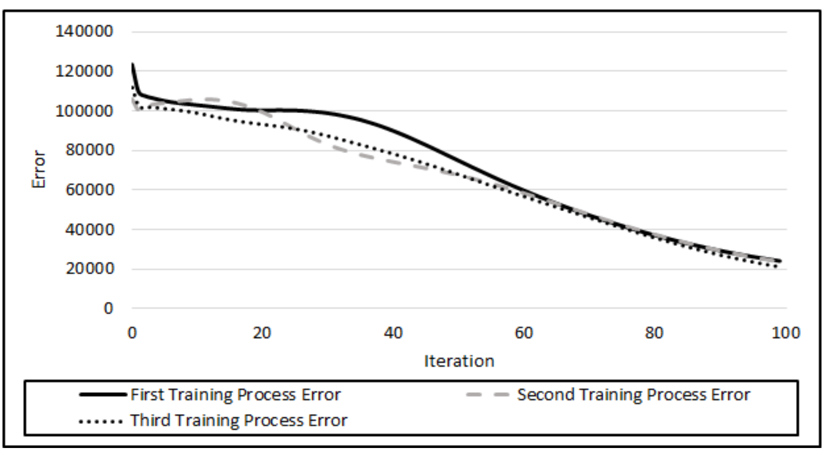

Fig. 4. Comparison of training error changes for autoencoders on the collected dataset (solid black line - change in error for the first training; grey dashed line - change in error for the second training; black dotted line - change in error for the third training).

Table 1 shows the best autoencoder neuromodels with the smallest errors during all training process.

Table 1. Autoencoder neuromodels with smallest training error.

Number of a training process	Smallest error, units	Iteration with the smallest error
First	24070	100
Second	23321	100
Third	21288	100

Figure 5 shows the graphs comparing the changes in training errors between the pre-trained neural network and the baseline neural network.

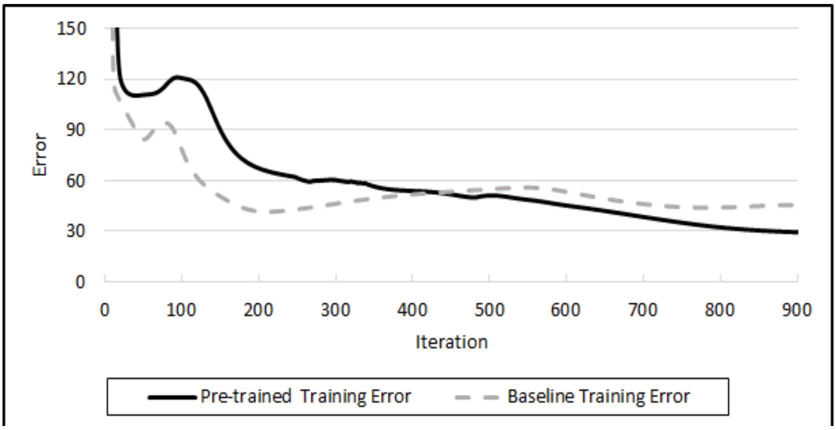

Fig. 5. Comparison of training error changes for the pre-trained neural network and the baseline neural network (black solid line – pre-trained neural network training error, grey dashed line – baseline neural network training error).

In Figure 5, it is clearly visible that:

- By the 900th iteration of training, the error of the pre-trained neural network becomes 35.4% smaller than that of the baseline neural network.
- It is also noticeable that after 200 iterations of training, the baseline neural network stopped decreasing the training error (and at a certain point even started to increase), while the training error of the pre-trained neural network continued to decrease.

Table 2 shows baseline and pre-trained neuromodel with smallest training error. The best training error of a pre-trained neuromodel is 29% smaller than the best training error of a baseline pre-trained neuromodel

Table 2. Smallest training error of a baseline and pre-trained neural models with three hidden layers.

Training process name	Smallest error, units	Iteration with the smallest error
Pre-trained neural network	29.35	900
Baseline neural network	41.53	210

Thus, the experimental results shown in Figure 5 confirm the advantages of pre-trained neural network over baseline neural network in synthesizing surrogate models.

Further research in this area will focus on conducting extensive investigations to assess the efficacy of pre-training (Figure 6) by leveraging the following:

- Exploration and analysis of alternative operating modes of the identical gas turbine electric power station, including scenarios like load-shedding mode, to evaluate the impact on surrogate model performance.
- Collection and examination of experimental data derived from the operation of a distinct gas turbine electric power station, aiming to assess the generalizability and robustness of the pre-training approach.
- Acquisition and analysis of experimental data pertaining to various components within the power system, such as synchronous generators, asynchronous motors, and other relevant elements, to explore the potential benefits of pre-training across a wider scope.
- Emphasis on utilizing empirical data obtained from full-scale experiments rather than relying solely on computer simulations, with the goal of validating the effectiveness and practical applicability of the pre-training methodology.
- The synthesis of neural network based control systems for power supply system components.

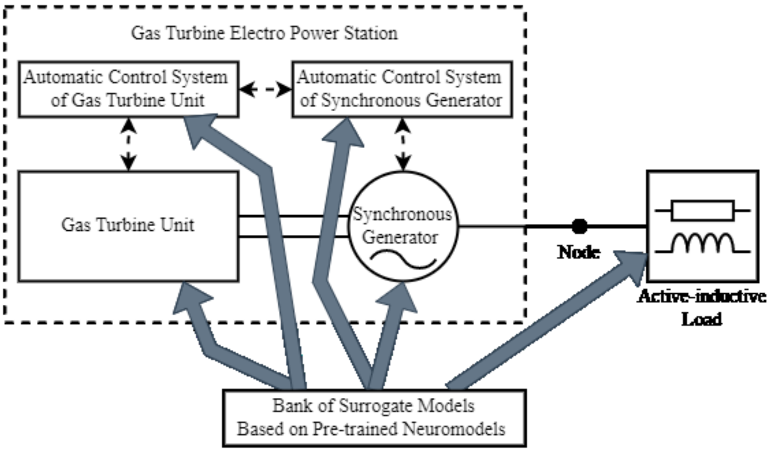

Fig. 6. Possible use of pre-training in surrogate modeling.

4 Conclusions

In this study, authors thoroughly examined the application of pre-training using autoencoders in the synthesis of surrogate models, demonstrated its potential to improve model accuracy and reduce the time required for synthesis and showed that pre-training method helps to reduce computational costs and improve the approximation of complex processes. Results of the paper demonstrated the advantages of pre-training and its practical significance for real world application for a gas turbine unit that is part of a gas turbine electric power station.

In conclusion, pre-training is a promising area of research in the energy sector. It opens up new possibilities for developing and using more effective and accurate models, helping to solve complex scientific and engineering problems with fewer time and resource costs.

Ideally, the procedure for obtaining a surrogate model based on a pre-trained neural network should take only a few training iterations [9-10], which should drastically reduce the time required for synthesizing surrogate models.

Acknowledgements

The research was carried out with the support of the Russian Science Foundation grant No. 23-29-00381, https://rscf.ru/project/23-29-00381/.

References

1. A. Forrester, A. Sobester, A. Keane, Engineering design via surrogate modelling: a practical guide (John Wiley & Sons, New York City, 2008)
2. S. Grihon, S. Alestra, D. Bettebghor, E. Burnaev, P. Prikhodko, *Comparison of different techniques for surrogate modeling of stability constraints for composite structures*, in Proceedings of 1st International Conference on Composite Dynamics, DYNACOMP, 22-24 May 2012, Archanon, France (2012)
3. S. Haykin, Neural networks: a complete course (Williams Publishing House, Moskow, 2006)
4. K. Robert, Basic search neural networks (Williams Publishing House, Moskow, 2001)

5. D. Erhan, Y. Bengio, A. Courville, P. Vincent, P. Manzagol, S. Bengio, J. Machine Learning Research, **11,** 625 (2010)
6. G. Kilin, B. Kavalerov, A. Suslov, I. Tyatenkov, Modeling gas turbine electro power station typical operating modes using pre-trained artificial neural network, Earth and Environmental Science, **990,** 012028 (2022)
7. C.Y. Liou, C.W. Cheng, J.W. Liou, D.R. Liou, Intelligent Science and Intelligent Data Engineering, **7751**, 458 (2013)
8. A. Khurshudov, Learning large-scale images features using multi-layer sparse autoencoder, Herald of Computer and Information Technologies, **2**, 27 (2014)
9. X. Sun, B. Wang, Z. Wang, H. Li, He. Li, K. Fu, IEEE J. of Selected Topics in Applied Earth Observations and Remote Sensing, **14**, 2387 (2021)
10. I. Drori, S. Zhang, R. Shutteleworth, L. Tang, A. Lu, E. Ke, K. Liu, L. Chen, S. Tran, N. Cheng, R. Wang, N. Singh, T. Patti, J. Lynch, A. Shporer, N. Verma, E. Wu, G. Strang, Proc. of the National Academy of Sciences, **119**, 2123433119 (2022)