# *A New Health Assessment Prediction Approach: Multi-Scale Ensemble Extreme Learning Machine*

BERGHOUT Tarek

*Laboratory of Automation and Manufacturing Engineering, University of Batna 2, Batna 05000, Algeria;*

t.berghout@univ-batna2.dz

**Keywords:**

**remaining useful life;**
**c-mapss;**
**extreme learning**
**machine;**
 **prognostic and**
**health management;**
 **neural networks.**

**Abstract-**This work can be considered as a first step of designing a future competitive data-driven approach for remaining useful life prediction of aircraft engines. The proposed approach is an ensemble of serially connected extreme learning machines. The results of prediction of the first networks are scaled and fed to the next networks as an additive features to the original inputs. This feature mapping allows increasing the correlation of training inputs with their targets by holding new prior knowledge about the probable behavior of the target function. The proposed approach is evaluated under remaining useful estimation using a set of "time-varying" data retrieved from the public dataset C-MAPSS (Commercial Modular Aero Propulsion System Simulation) provided by NASA. The prediction performances are compared to basic extreme learning machine and proved the effectiveness of the proposed methodology.

## 1.   Introduction

Remaining Useful Life (RUL) of complex systems based on conventional prediction paradigms such as physical modeling can be considered as very difficult task due to the need of the deep knowledge of systems components and its interior interactions. Besides, most physical models are constructed under limited conditions which is the appropriate explanation of the  poor generalization[1].

Nowadays, the availability of historical data of operating systems due the continuous growth in sensors technologies makes Machine Learning (ML) tools receiving a growing attention. Replacing classical predictions paradigms with training tools is the appropriate solution for complexity and human intervention reduction.

Ensembles learning[2]–[5], Hybrid algorithms [6]–[9] and deep learning families[10]–[13] are the most investigated ML approaches for RUL estimation. However prediction algorithms based on old training algorithms such as backpropagation for neural networks and Lagrange minimization for support vector machines are very expensive and time consumers.

In recent decades, a new fast and accurate training approach named Extreme Learning Machine (ELM) has been spread wide to fit many prediction applications under different architectures[14]. ELM was firstly created to remove the barriers between biological learning and Artificial Neural Networks (ANN) by introducing a new tuning theory. In ELM, the hidden nodes parameters need not to be tuned. In fact, the output weights are the only responsible for universal approximation[15]. Therefore, in this context a new data-driven approach is reconstructed based on ELM in the aim of producing a very fast and

accurate learner for "time-varying" data. The developed Multi-scale Ensemble Extreme Learning Machine (MSE-ELM) is evaluated and compared to basic ELM and proved the efficacy of the new feature mapping.

The remainder of this work can be addressed as follow: in section 2, the description of the used materials in this study is briefly introduced. Section 3 elaborates the proposed approach adopted in this work. Experiments, results and discussions are showcased in Section 4, Finally section 5 is the conclusion of this work.

## 2. Materials

C-MAPSS dataset is a publically provided dataset as a benchmark for constructing and evaluating health state estimators of aircraft engines under operating conditions[16]. Retrieved data is divided into four datasets according to several failure modes and operating conditions. Each subset carries a set of 100 different degradation profiles of the engine defined by time series measurements of 24 different sensors. According to prognostic health management challenge in 2008 [17] the target function for each life cycle, describes a non cumulative degradation as a piecewise stochastic linear function. **Figure 1** gives an insight about the behavior of sensors measurements and the RUL target function in one life cycle of the engine.
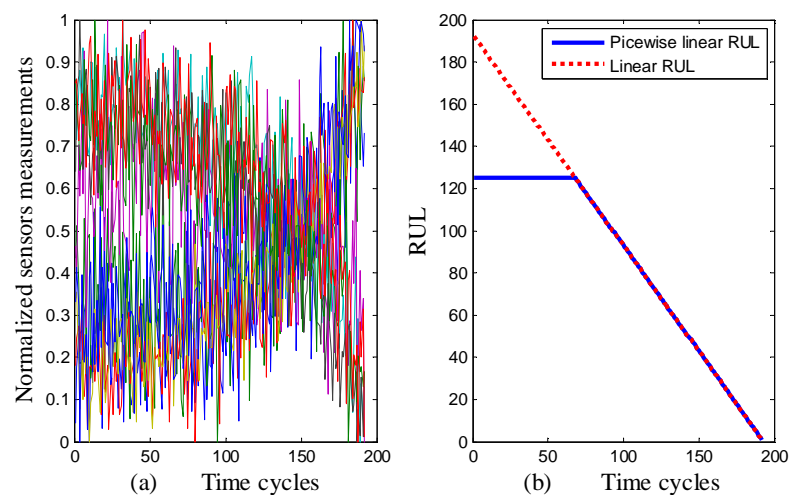


Figure: 1 . a) sensors measurments behavior in one lif cycle. b) RUL target function.

## 3. Methods

In 2004[15], ELM were proposed as the new training scheme for Single hidden Layer Feedforward Neural networks (SLFN), than it was extended to fit a verity of architectures for different applications[14]. Unlike backpropagation algorithm, the basic training rules of ELM can be simply addressed in three steps subject to minimize the objective function in equation (1). $\beta$ is the output weights matrix, $H$ is feature mapping of the entire training set and $T$ are the desired targets.

$$E = \left\| T - H\beta \right\|_2 \tag{1}$$

1. The one needs to generate the hidden nodes parameters: weights *A* and biases *B*, randomly from any probability distribution as full rank and linearly dependent matrices respectively.
2. Calculate and activate the hidden layer for a "single-batch" of data according to any chosen activation function *G* as shown in equation (2). *X* refers to the entire training inputs.

$$H = G(AX + B) \qquad (2)$$

3. Determine analytically the output weights $\beta$ depending on the Pseudo-inverse of the *H* matrix using equation (3).

$$\beta = H^{-1}T \qquad (3)$$

Theoretically and without considering over-fitting and other ill posed problems, it is mentioned in ELM theories that the more the number of hidden nodes, the more the accuracy the network may achieved [15]. However, experimentally in higher dimensional problems such as the feature space of the current studied dataset, more hidden nodes can delay the training process or obscure the computer memory.

In the proposed approach and based on [18], we constructed a new multilayer neural network architecture based on ELM learning rules. The new neural network that is showcased in **Figure 2,** allows the use of small ELM learners to map the training inputs accurately to achieve more satisfaction and fast generalized approximation.

Our contribution attempts to prove that: "By scaling the predicted target of one network the same as inputs and considered it as an additive feature to the original inputs, the reconstructed data at the finale learner will carries a prior knowledge about the probable targets.". The new feature mapping in this case does not consider the hidden layers before the final feature mapping. In fact, the results of last learner are the final decision of the training model.
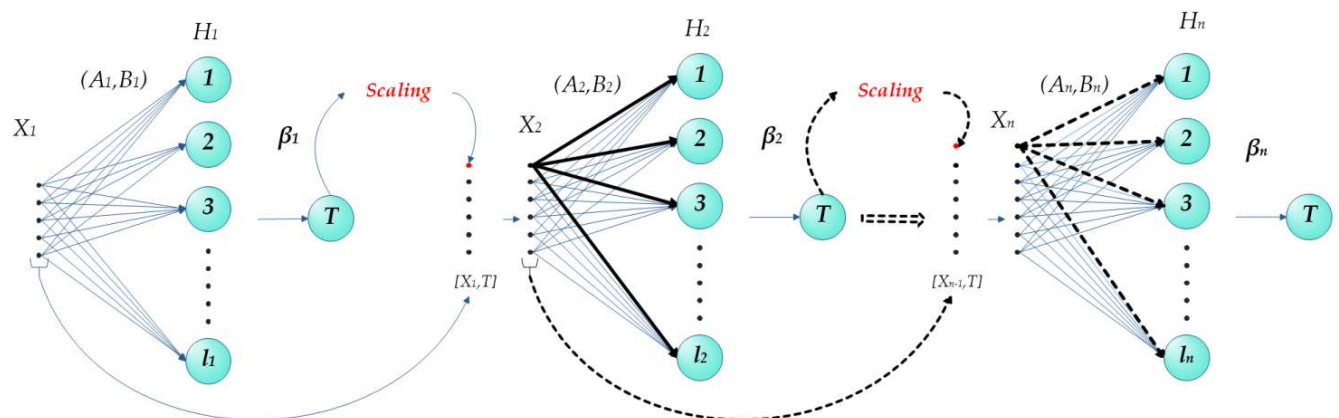


Figure: 2.Architecture of the proposed Network.

### 4.   Experiments, results and discussion

In the present comparative study, the subset that is named FD001 from C-MAPSS dataset is considered. In this case, only one failure mode and one operating condition are investigated.  As a preprocessing step of our data, sensors whose indices are {7, 8, 9, 12, 14, 16, 17, 20, 25, 26} are chosen for the application and their measurements are scaled using min-max normalization according to the literature [1].

The results of this comparative study between ELM and proposed MSE-ELM are carried out using the architecture addressed in **Table 1**. To study the algorithms performances under the same conditions we make the hidden nodes in SLFN architecture equal to the sum of hidden neurons in the entire MSE-ELM network.

Table: 1. Table of hyper parameters.

| Methods | Hidden layers | Activation function |
|---------|---------------|---------------------|
| **ELM** | **1** | **sigmoid** |
| **MSE-ELM** | **4** | **sigmoid** |

As a first advantage of the new architecture, the training based basic ELM makes the SLFN much sensitive towards the risk of over-fitting before the expected accuracy is achieved. Weakness of SLFN towards empirical risk is already proven by **Figure 3** where the networks is well trained based on ELM and over-fits earlier in testing phase. Unlike basic ELM, the proposed architecture allows the networks to extend positively to achieve more accuracy.
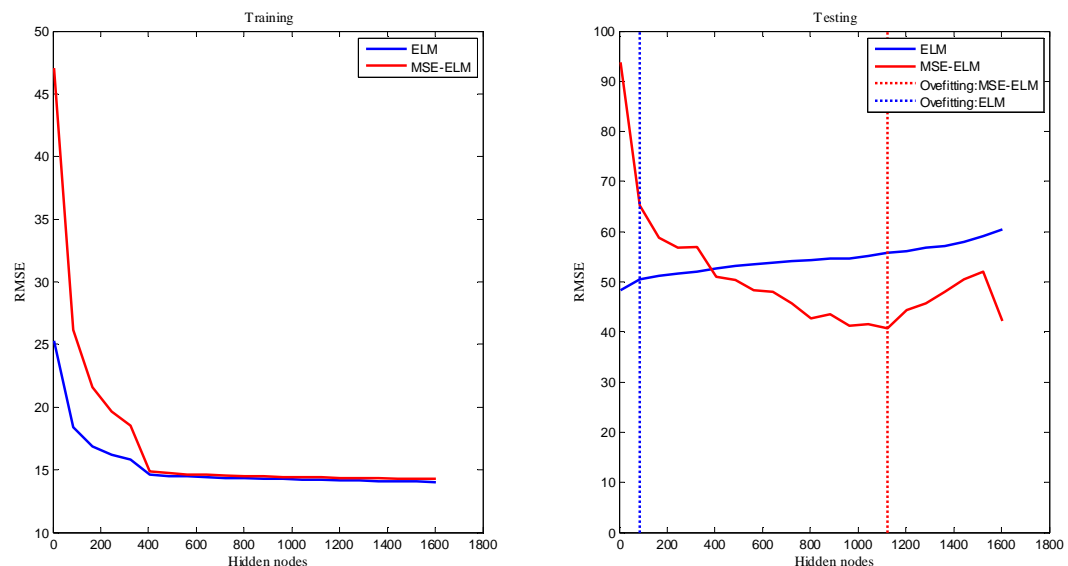


Figure: 3. Accuracy results during growth in hidden nodes

The ensemble of small size of ELM learners can contribute in time consumption reduction during training as showcased in **Figure 4** where clearly MSE-ELM outperform ELM in training phase. In the testing phase the MSE-ELM takes more time

compared to SLFN based ELM due to the multilayer architecture, but this difference can be neglected compared to training time consumption ratio.
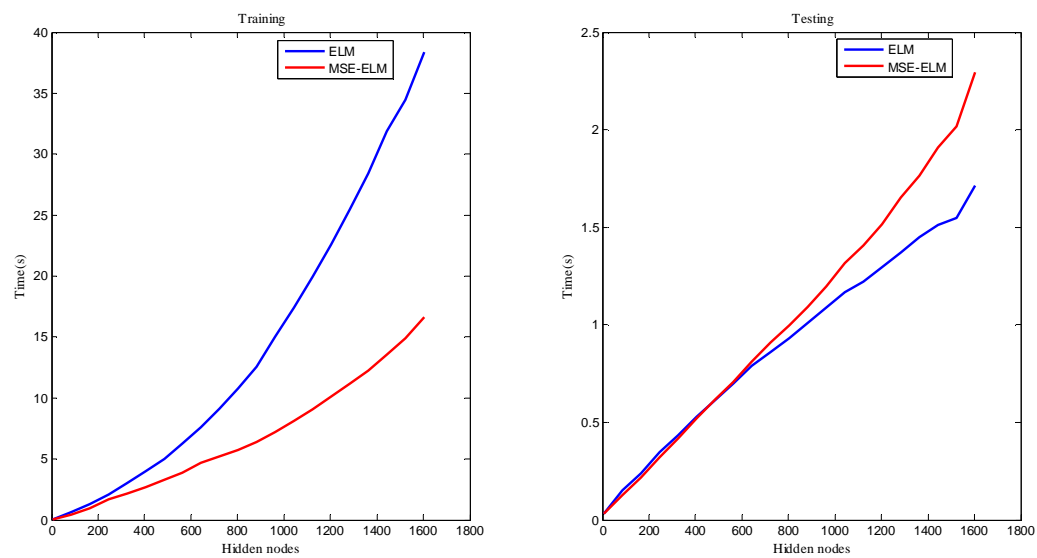


Figure: 4. Time comsuption atitude during growth of hidden nodes.

The RUL prediction results of FD001 test set of both algorithms are shown in **Figure 5.** The new architecture clearly proves the accuracy of proposed new feature mapping methodology by enhancing prediction accuracy.
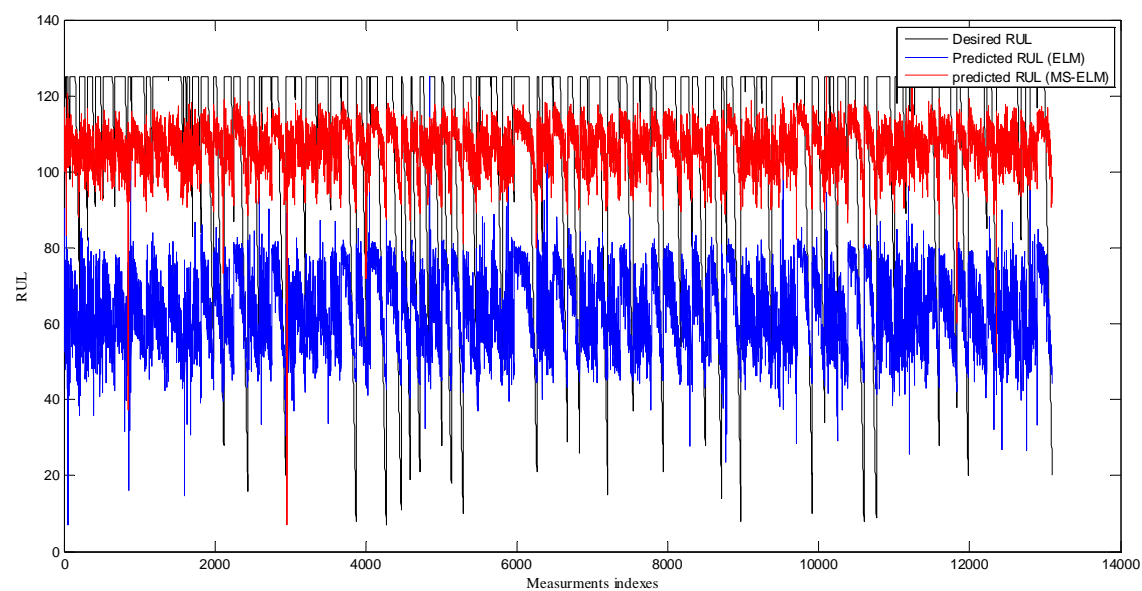


Figure: 5. RUL prediction results.

## 5.   Conclusion

The present works attempts to prove that the theory of the new feature mapping based ensemble of small neural networks and feature scaling can enhance the prediction model in term of accuracy and time consumption. After evaluating the proposed approach under higher dimensional "time-varying" data obtained from C-MAPSS simulation software, and the results strongly supports the credibility of the new training scheme.

MSE-ELM is a not complicated offline training approach for neural networks designed to study the accuracy of the proposed feature mapping. Therefore, the future approaches will focus on enhancing this algorithm by integrating online and adaptive learning attitudes attempting to produce a more competitive 'data-driven' approach to other tools in the literature.

## References

[1]     T. Berghout, L. Mouss, O. Kadri, L. Saïdi, and B. Mohamed, "Aircraft Engines Remaining Useful Life Prediction with an Improved Online Sequential Extreme Learning Machine," *Appl. Sci.*, vol. 10, no. 3, 2020.

[2]     L. Wen, Y. Dong, and L. Gao, "A new ensemble residual convolutional neural network for remaining useful life estimation," *Math. Biosci. Eng.*, vol. 16, no. 2, pp. 862–880, 2019.

[3]     K. Shen, "Data assimilat ion by coupling un certain support vec tor machine wit h ensemble kal man filt er," pp. 15–17, 2012.

[4]     B. Wang, Y. Lei, N. Li, and J. Lin, "An improved fusion prognostics method for remaining useful life prediction of bearings," *2017 IEEE Int. Conf. Progn. Heal. Manag. ICPHM 2017*, pp. 18–24, 2017.

[5]     J. Liu, M. Zhang, H. Zuo, and J. Xie, "Remaining useful life prognostics for aeroengine based on superstatistics and information fusion," *Chinese J. Aeronaut.*, vol. 27, no. 5, pp. 1086–1096, 2014.

[6]     C. Ordóñez, F. Sánchez Lasheras, J. Roca-Pardiñas, and F. J. de C. Juez, "A hybrid ARIMA–SVM model for the study of the remaining useful life of aircraft engines," *J. Comput. Appl. Math.*, vol. 346, pp. 184–191, 2019.

[7]     A. Al-Dulaimi, S. Zabihi, A. Asif, and A. Mohammadi, "A multimodal and hybrid deep neural network model for Remaining Useful Life estimation," *Comput. Ind.*, vol. 108, pp. 186–196, 2019.

[8]     P. J. García Nieto, E. García-Gonzalo, F. Sánchez Lasheras, and F. J. De Cos Juez, "Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability," *Reliab. Eng. Syst. Saf.*, vol. 138, pp. 219–231, 2015.

[9]     S. M. Lee, T. S. Roh, and D. W. Choi, "Defect diagnostics of SUAV gas turbine engine using hybrid SVM-artificial neural network method," *J. Mech. Sci. Technol.*, vol. 23, no. 2, pp. 559–568, 2009.

[10]    L. Ren, Y. Sun, H. Wang, and L. Zhang, "Prediction of bearing remaining useful life with deep convolution neural network," *IEEE Access*, vol. 6, pp. 13041–13049, 2018.

[11]    X. Li, Q. Ding, and J. Q. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Reliab. Eng. Syst. Saf.*, vol. 172, pp. 1–11, 2018.

[12]    H. Yang, F. Zhao, G. Jiang, Z. Sun, and X. Mei, "applied sciences A Novel Deep Learning Approach for Machinery Prognostics Based on Time Windows," 2019.

[13]    J. Deutsch and D. He, "Using deep learning based approaches for bearing remaining useful life prediction," *Proc. Annu. Conf. Progn. Heal. Manag. Soc. PHM*, vol. 2016-Octob, pp. 1–7, 2016.

[14]    G. Huang, "An Insight into Extreme Learning Machines : Random Neurons , Random Features and Kernels," *Springer New York LLC*, vol. 6, no. 3, pp. 376–390, 2014.

[15]    G. Bin Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," *IEEE Int. Conf. Neural Networks - Conf. Proc.*, vol. 2, no. August, pp. 985–990, 2004.

[16]    A. Saxena, M. Ieee, K. Goebel, D. Simon, and N. Eklund, "Damage Propagation Modeling for Aircraft Engine Prognostics," *Response*, 2008.

[17]    F. O. Heimes, "2008 INTERNATIONAL CONFERENCE ON PROGNOSTICS AND HEALTH MANAGEMENT Recurrent Neural Networks for Remaining Useful Life Estimation," 2008.

[18]    J. Zhu, N. Chen, and W. Peng, "Estimation of Bearing Remaining Useful Life Based on Multiscale Convolutional Neural Network," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3208–3216, 2019.