Journal of Mechanical, Civil and Industrial Engineering
ISSN: 2710-1436
DOI: 10.32996/jmcie
Journal Homepage: www.al-kindipublisher.com/index.php/jmcie

JMCIE
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Determining RUL Predictive Maintenance on Aircraft Engines Using GRU

Adryan Fitra Azyus[1] ✉ Sastra Kusuma Wijaya[2] and Mohd Naved[3]

[1]*Faculty of Mathematics and Sciences, Nature/Physics Instrumentation, University of Indonesia, Indonesia*
[2]*Fakultas Matematika dan Ilmu Pengetahuan, Alam/Fisika Instrumentasi, Universitas Indonesia, Indonesia*
[3]*Department of Business Analytics, Jagannath University, D elhi-NCR, India.*

Corresponding Author: Adryan Fitra Azyus, **E-mail**: adryan.fitra01@ui.ac.id

| ABSTRACT

Prognostic and health management (PHM) in the aviation industry is expanding because of its effect on economic and human safety. Advanced maintenance shall be applied to this industry to inform aircraft engine conditions. PdM (Predictive Maintenance) is an advanced maintenance technique that can be applied to the aviation industry because of its high-precision prediction. Combining PdM as a technique to calculate the RUL (Remaining Useful Lifetime) and ML (Machine Learning) as a tool to make high-accuracy predictions is mixed together that accurately forecasts the state of aircraft machine condition and on the best time to get the maintenance or service. In this work, we use the NASA Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) data set. This work proposes GRU to determine RUL on aircraft engines to implement a Predictive maintenance strategy. For the training parameters tested in this study, we used a batch size of 512, a learning rate with Adam optimizer of 0.001, then epochs of 200. The essence of the results of this experiment is to obtain a new method with a simpler calculation process and the epoch value and a faster prediction process compared to other methods used, and the results obtained can approach the original value from an economic point of view and the RUL prediction process using the GRU.

| KEYWORDS

RUL Predictive Maintenance; Aircraft Engines; Prognostic and health management

| ARTICLE INFORMATION

ACCEPTED: 09 December 2022 **PUBLISHED:** 11 December 2022 **DOI:** 10.32996/jmcie.2022.3.3.10

1. Introduction

Machine learning (ML) is a subsection of Artificial Intelligence. This method or algorithm can learn based on training, which is the given information about something, and it can be used in the future when the algorithm is applied. Deep learning (DL) is the development algorithm from ML to fix its limitation. DL implemented an artificial neural network (ANN) for its algorithm and greatly impacted the supervised learning method. DL algorithms, such as image processing and face and speech recognition, are widely employed. The other function of ML and DL is to predict with high accuracy because both algorithms or models can calculate based on training from a dataset or additional information [Wuest, 2016].

The prediction process that needs mathematical methods and techniques is suitable for maintenance strategies such as Prognostic and health management (PHM). Today's popular PHM is on the aviation industry to predict the state of the aircraft engine condition because the capacity of machinery working cannot last forever. Sometimes, it will be broken-down because of out-date operation. Machinery systems that include sensors are just monitoring the state of the machine but cannot make a report of the machine's condition. A maintenance strategy must apply to the scheduled machinery system to avoid the worst event (failure) and get information about a machine's status. Predictive maintenance (PdM) is strategy maintenance, but continuously monitoring the state of the machine, the maintenance performed optimally. In other words, PdM indicates the state of the machine to perform a maintenance schedule based on historical data, integrity factors, statistical inference methods, and engineering approaches [Susto, 2012].

Prognostic and health management (PHM) in the aviation industry is important because of its effect on economic and human safety. Advanced maintenance shall be applied to this industry to inform aircraft engine conditions. PdM is an advanced maintenance technique that can be applied to the aviation industry because of its high-precision prediction. It can reduce the cost of operation and increase safety by calculating the remaining helpful time (RUL) of aircraft engines [Si, 2011].

1.1 Machine Learning

To validate the result of using ML and DL methods, evaluation is important to indicate the performance of the model that has been trained. *Root Mean Squared Error* (RMSE) is a frequent evaluation method for regression, indicating how concentrated the output is around the line of best fit on prediction data. The evaluation method for the classifying technique is *Accuracy* defined parameter, which indicates output values of prediction compared with the actual value; p*recision* describes a level of accuracy between the output of prediction compared with all data. *A recall* is defined as the level of correct result compared to the number of results that should have been returned, which means the higher the recall value, the higher data collected. No data will be missed from accuracy and precision [Wuest, 2016]. Furthermore, the combination of PdM as a technique to calculate the RUL and ML as a tool to predict high accuracy is mixed together to accurately forecast the state of the aircraft machine condition and the best time to get the maintenance or service [Theissler, 2021].

1.2 GRU (Gated Recurrent Unit) Model

One of the best and most popular DL models to make a prediction, especially to determine PdM on aircraft engines, is long short-term memory (LSTM) neural networks that are applied to be specialized in extracting sensor temporal information [Kong, 2019]. Although the accuracy of LSTM is higher than other algorithms, the training time of LSTM is much longer than that of other algorithms. How to reduce the training time under the premise of ensuring high accuracy is still challenging. The gated recurrent unit (GRU) is a special case of LSTM. It has a shorter training time than LSTM. GRU networks are mainly used in classification and are seldom applied in regression problems [Wang, 2018]. A Gated recurrent unit (GRU) is a variant of LSTM that contains an update gate, reset gate, activation, and candidate activation, respectively. The update gate controls how much historical and new information needs to be forgotten in the current state. The reset gate controls how much information is available from the candidate state. The candidate activation can be regarded as new information at present. If much further information is retained, the old data will be less, and vice versa. GRU simplifies the input gate and forgets the gate of LSTM into an update gate, combining the cell and hidden states. Therefore, the GRU unit retains the advantages of LSTM and further reduces the model training time by reducing the parameters in the model. The GRU has fewer parameters than LSTM, so the trend in deep learning models applied to time series analysis is toward GRU.

There is various implementation of GRU models, such as forecasting photovoltaic [Wang, 2018], speech recognition [Ravanelli, 2018], and predicting the remaining useful lifetime (RUL) on battery Lithium and other machinery [Song, 2018]. So, in this paper, we work on an experiment to determine RUL on an aircraft engine for predictive maintenance using the gated recurrent unit (GRU) model. Therefore, some research uses the same method and object or dataset. Still, in this work, we modify the parameters and use a server to calculate the process for flexibility and economic purposes.

2. Method

In this study, we tried to experiment with using the GRU (Gate Recurrent Unit) architecture by building several GRU units. For benchmarking, we used the CMAPSS Dataset. This dataset can be obtained from the NASA website for turbofan engines.

Therefore, various datasets are available for turbofan engines or aircraft. The CMAPSS data is the most popular one. Much research is conducted using this dataset, so it is easy to compare and develop more detail on the model architecture [Vollert, 2021]. CMAPSS Dataset has four pieces of data, namely FD001, FD002, FD003, and FD004, each divided into Train, Test, and RUL data. The size of each dataset is different, where FD002 and FD004 are two times larger than FD001 and FD003 because the four data have different machine ids. So, to carry out training and testing, we each run it separately depending on the data code [Ramasso, 2014].

In processing data to become GRU input, of course, we change the size of the data dimensions because we know that in Deep Learning, the input data must have high dimensions, especially LSTM and GRU. For this reason, in each dataset, 26 columns indicate the characteristics of this RUL, with the number of rows being the total cycle of each machine ID, so we need to perform a dimensional transformation to a higher level so that more feature levels are absorbed by the artificial neural network, namely the GRU that was built. The process of transforming this dimension from 2 dimensions to 3 dimensions is carried out by determining the size of the data sequence to be 50, with the number of features being the total column other than the machine ID column, which amounts to 26. So, a two-dimensional data size (50, 26) will be formed, which can then be determined as batch size for the training process. The label of the data in the form of the RUL value is the value of the highest cycle of each machine id, which number depends on the number of machine IDs.

For the calculation of the RUL value for each cycle, it can be done by subtracting the maximum value of the cycle with each cycle and then adding it up with the RUL value for each machine id; the formula is as follows:

$$RUL = RUL_{True\ from\ dataset} + RUL_{cycle\_max-cycle} \tag{1}$$

Then each machine id has a different number of cycles so that we can manually determine the maximum life of each machine to find the point of degradation (decrease in machine performance) while the algorithm is as follows:

$$if\ n\ in\ RUL >= MAXLIFE : newRUL = MAXLIFE\ else\ newRUL = n \tag{2}$$

So, the RUL graph form with cycle will be like in the following image:

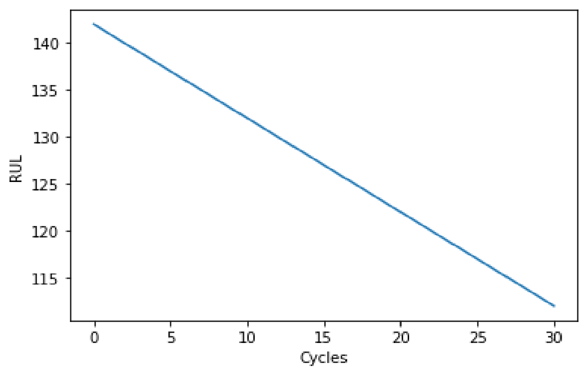
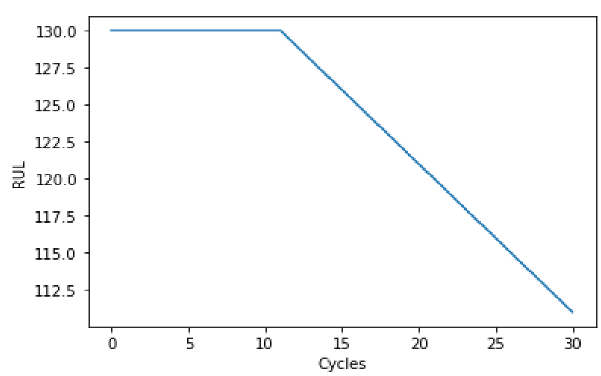


FIGURE 1. RUL graph

Figure 1 shows the relationship between RUL and engine cycles, where the highest RUL is 140. Then, for example, if the max life of the engine is set above 130, the RUL degradation point occurs on the 12th cycle. The graphic form will be like Figure 2.



The architecture of the GRU can be seen in Figure 3, where the GRU architecture is divided into an input layer, a GRU layer, a Fully Connected layer, and an output layer. For the 3-dimensional input layer (none, 50, 26), each time it is processed through the GRU layer and the Fully Connected layer, it will be activated with a sigmoid function to produce output in the form of a single number, namely the RUL value. As for the experiment, the number of GRU architectures that were tried were three architectures by combining the number of GRUs, namely 2 GRU to 4 GRU. In theory, the deeper the network is created, the deeper the learning process will be and the greater the reach of feature levels. For the training parameters tested in this study, we used a batch size of 512, a learning rate with Adam optimizer 0.001, and then epochs of 200.

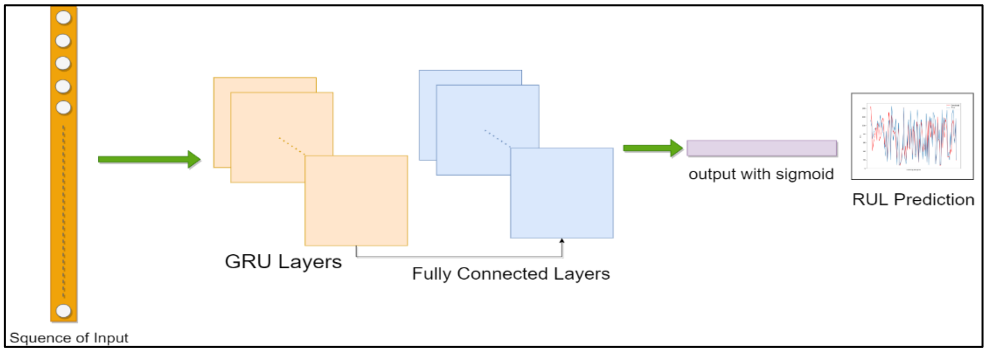


FIGURE 2. GRU model architecture

On the main configuration for the experiment, the number of GRU architectures tried was three, as shown in Table 1 below. In theory, the deeper the network is created, the deeper the learning process will be and the greater the reach of feature levels. For the training parameters tested in this study, we used a batch size of 512, a learning rate with Adam optimizer of 0.001, and then epochs of 200. This experiment also has been done with a server for the calculation because it is flexible and can be done faster than a personal computer with high specifications.

TABLE 1. Three combinations of GRU architectures for an experiment.

Model	GRU	Total Parameter Training
gru1	1	375,937
gru2	2	409,089
gru3	3	421,289

3. Results

In this work, the experiments for predicting RUL estimates on aircraft engines using the C-MAPSS dataset with the regression method and the GRU deep learning algorithm were completed using 200 epochs with a 12 GB RAM server and 9 AMD epyc cores as well as OpenVZ virtualization. The result from the experiment compares with the original RUL on the C-MAPSS dataset.

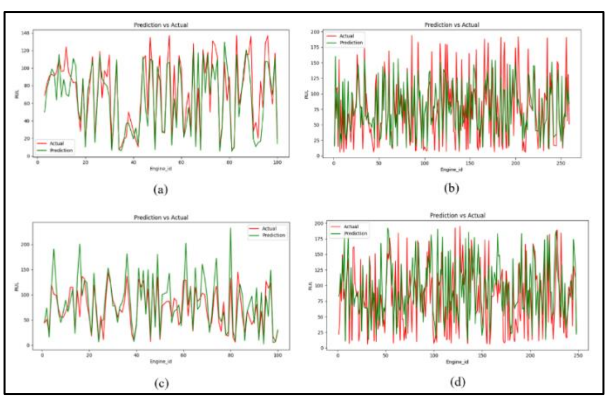


FIGURE 3. General predictable between RUL data set and experiment (a)FD001 (b)FD002 (c)FD003 (d)FD004 (Source: Author)

The C-MAPSS dataset is preprocessed first. Not a cleaning process is carried out but a windowing process, extracting the data more deeply to produce a dataset ready to be processed using the GRU algorithm. The results of the prediction process with the regression method can be seen in Figure 4.

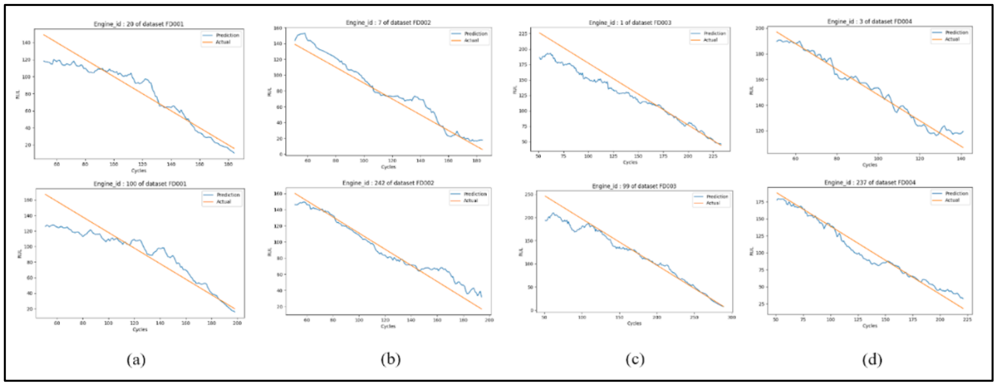


FIGURE 4. Degradation pattern between RUL data set and experiment (a) result from dataset FD001 using model_gru1(b) result from dataset FD002 using model_gru3 (c) result from dataset FD003 using model_gru3 (d) result from dataset FD004 using model_gru3 (Source: Author)

Figure 4 is the result of the estimated RUL for a machine, in this figure only shows the initial and final images for each dataset. This figure also shows that the estimated RUL for predictions is close to the actual value of the RUL for all machines. The average results of all machines can be seen in figure 3, which shows the results of comparing RUL predictions for each machine in the dataset. Figure 4 is to validate all evaluation values for each machine cycle on all datasets so that in table 1, we get the average value for each machine, the model used, and the value evaluation for the validation process. From this table, it can also be seen that the models used are different for each dataset. This is because the fittest model for each dataset is different. Still, the gru3 model is the most dominant because this model has the lowest RMSE value for the datasets FD002, FD003, and FD004, while the gru1 model is the fittest for the FD001 dataset. It can be interpreted that the characteristics of each dataset are different so that it will produce different values for each model that is applied to each dataset.

TABLE 2. Experiment results for each machine using the GRU model.

C-MAPSS	Model	RMSE	MSE
FD001	model_gru1	20.5502224	427.197998
FD002	model_gru3	34.28032684	1200.295044
FD003	model_gru3	31.74353981	1026.914917
FD004	model_gru3	44.75546265	2023.009277

Furthermore, the results of this experiment will be compared with the results of the previous experiment. Some experiments using other methods such as CNN, MLP to SVR with the same dataset will compare the evaluation value, namely RMSE. from the results of the comparison in table 2, it is found that the RMSE value for each database can be different depending on the accuracy of the model used to predict the estimated RUL value for each dataset, for the GRU algorithm the lowest RMSE value is obtained to predict the estimated RUL value on the FD001 set. From the overall comparison data, the GRU still tends to be better than some other algorithms, although not the best. However, the GRU algorithm can be developed again to get a lower evaluation value.

TABLE 2. Comparison **of** the experiment results in this paper with another [11].

Model	C-MAPSS			
	FD001	**FD002**	**FD003**	**FD004**
MLP	37.56	80.03	37.39	77.37
SVR	20.96	42	21.05	45.35
RVR	23.8	31.3	22.37	34.34
GRU	**20.55**	**34.28**	**31.74**	**44.75**

The advantages of using GRU to predict the estimated value of RUL are in terms of processing efficiency and costs used for research because the RUL method can be processed even using a medium specification PC so that in the future, machine learning and deep learning processing can be done more cheaply.

4. Conclusion

In this paper, the GRU algorithm can predict the estimated RUL of the C-MAPSS dataset with 200 epochs and a light computational process. From the 4 C-MAPSS datasets, it turns out that they have different characteristics for each dataset, so not all datasets can be applied to the same GRU model. In this paper, the cnngru1 model with the configuration gets the lowest RMSE value for the FD001 dataset and the cnngru3 model for the other datasets so that the model is appropriate for the use of that dataset. In the future, the prediction process can use a different model configuration for each dataset to produce a model according to the characteristics of the C-MAPSS dataset. Adding epochs of up to 300 or 500 is believed to increase prediction results with a lower RMSE value.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Kong, Z., Cui, Y., Xia, Z., and Lv, H. (2019). Convolution and long short-term memory hybrid deep neural networks for remaining useful life prognostics. Applied Sciences, 9(19), 4156

[2] Ramasso E and Saxena, A. (2014). Review and Analysis of Algorithmic Approaches Developed for Prognostics on CMAPSS Dataset, In *International Conference of the Prognostics and Health Management Society*

[3] Ravanelli, M., Brakel, P., Omologo, M. and Bengio, Y. (2018). Light-gated recurrent units for speech recognition. In: *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(2), 92-102 (IEEE, 2018).

[4] Song, Y., Li, L., Peng, Y. and Liu, D. (2018). Lithium-ion battery remaining useful life prediction based on GRU-RNN, In:*2018 12th international conference on reliability, maintainability, and safety (icrms)*, 317-322

[5] Si, X, Wang, W, Hu, C.H and Zhou, D. (2011). Remaining useful life estimation - a review on the statistical data-driven approaches. *European Journal of Operational Research*. *213*, 1–14

[6] Susto, G. A., Member, S., Beghi, A., & Luca, C. D. (2012). A predictive maintenance system for epitaxy processes based on filtering and prediction techniques. *IEEE Transactions on Semiconductor Manufacturing. 25*, 638–649

[7] Sateesh B, Giduthuri, P Z and Xiao-Li L. (2016). Deep convolutional neural network-based regression approach for estimation of remaining useful life, In *International conference on database systems for advanced applications*, 214-228

[8] Theissler, A, P´erez-Vel´azquez, J, Kettelgerdes, M and Elger, G. (2021). Predictive Maintenance enabled by Machine Learning: Use cases and challenges in the automotive industry, *Reliability Engineering & System Safety*

[9] Vollert, S and Andreas T. (2021). Challenges of machine learning-based RUL prognosis: A review on NASA's C-MAPSS data set, In:*2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 1-8

[10] Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications, *Production & Manufacturing Research. 4*, 23–45

[11] Wang, Y., Liao, W. and Chang, Y. (2018). Gated recurrent unit network-based short-term photovoltaic forecasting. Energies, 11(8), 2163