DOI: https://doi.org/10.46354/i3m.2019.imaaca.005

TIME SERIES MODEL IMPROVING WITH AUTOMATIC SAVITZKY-GOLAY FILTER FOR REMAINING USEFUL LIFE ESTIMATION.

Youssouf Diaf [a], Samir Benmoussa [b], Mohand Djeziri [c]

[a] Agence Spatiale Algérienne, Bouzareah Alger, 16000 Algeria.
[b] Laboratoire d'Automatique et de Signaux de Annaba (LASA), University Badji Mokhtar Annaba, 23000 Algeria.
[c] Laboratoire des Sciences de l'Information et des Systemes (LSIS), UMR CNRS 7296 France.

[a]ydiaf@asal.dz, [b] benmous2a.samir@gmail.com, [c]mohand.djeziri@lis-lab.fr

ABSTRACT
The performance of fault diagnosis and failure prognosis methods is directly related to the quality of sensor measurements. When the systems evolve under extreme conditions of use as in the aeronautical field, the sensors measure are often impacted by strong noise and disturbances. This paper deals with the using of an automatic Savitzky-Golay filter to smooth the sensors data for time series model training and remaining useful life estimation. A proposed window value calculation for the Savitzky-Golay filter is used to improve signal quality. The NARX neuronal network is used as a data-driven approach to model the trend of the smoothed time series data. The proposed approach is successfully applied for degradation-trend modeling and remaining useful life prediction in turbo engines. These systems are considered as a high noise device which mean more challenge in data processing. The results show an improvement in the prediction of the remaining useful life compared to previous works.

Keywords: Remaining Useful Life, Savitzky-Golay filter, NARX neuronal network, smoothing, Gas Turbine Engines

1. INTRODUCTION
Nowadays, competitiveness and the race for performance are increasingly encouraging scientists and industrialists to replace preventive and corrective maintenance strategies with predictive and conditional ones. The deployment of these new maintenance strategies requires an understanding of the current health state of the systems and the future evolution of their degradation process. This knowledge is provided by fault diagnosis and failure prognostic methods (Vichare and Pecht. 2006).

Recently, the common for all the prognostic and diagnostic methods is prognostic and health management. The prognostic and health management (PHM) predict the state of health of equipment or system which improve our knowledge on it and provide a strategy to manage the best CBM maintenance actions. The PHM is based on either physical model, data-driven model or both. Remaining Useful Life (RUL) of the system is one of the key outputs of the methods of failure prognosis

Physical models are not available in most practical cases of industrial systems, because it is very difficult to make modeling assumptions, and formalize the cause-and-effect relationships when the interaction of components and physical phenomena are complex, strongly non-linear and sometimes unknown.

Data-driven failure prognostic methods are the most used in practical cases, and are becoming more and more efficient thanks to developments in data storage and analysis tools and artificial intelligence. The data-driven methods are based on data collected manually or remotely from the system. The data-driven methods create some new challenging problems for scientists, which is sensors.

Actually, in PHM the sensors technology offers more and more accurate measurements and less space to be implemented. These advantages in PHM applications are faced by the problems of how much we need indicators to accurately estimate the state of health and if we have the right indicators, how to prevent the noise effect.

In the most of application data is certainly contaminated with noise, particularly in a harsh environment (high pressure, temperature, vibration ….), where noise becomes very intense and random. In PHM the most comment method used to avoid the noise is smoothing. The smooth methods are multiple and every method has its advantages and disadvantages.

2. SMOOTHING IN DATA-DRIVEN APPROACHES
Recent works on data-driven prognostic algorithms are gating more and more accuracy in damage prognostic. The data-driven methods such neuronal network (Heimes 2008, javed and al. 2011, Bektas and Jones

2016), support vector machine (Bluvband and Porotsky 2016, Schölkopf and Smola 2001, Michael 2001), similarity-based prognostic (Tianyi and al. 2008) and others (Djeziri and Benmoussa and Sanchez 2018, Benmoussa and Djeziri 2017) are used to predict the RUL.

Smoothing take an essential part before any training process in prognostic. In (Bluvband and Porotsky 2016), the authors used a fitting technic to remove the noise. They use a monotonic fitting by non-linear regression methods, which used two types of smoothed functions the polynomial functions and the exponential functions. Because the smoothing function didn't guarantee permanent stability, they propose the selection of the right function for each sensor. In (Bektas and Jones 2016), a 9th and a 4th degree polynomial regression was used to filter the data and each polynomial degree is used either for training or testing. (Coble and Hines 2012) propose for reducing the noise a quadratic fitting technic.

In data analysis aspects, using these methods can sometimes lead to an undershoots and a loss of critical information (peaks). Additionally, the performance depends on the appropriate selection of the polynomial order or function which is difficult.

In this work we tend to improve RUL estimation by an automatic smoothing process. An automatic Savitzky-Golay filter (Abraham and Golay 1964, Dombi and Dineva 2018, Zhu and al. 2017) is proposed.

3. METHOD

3.1. Savitzky-Golay Filter

The core of Savitzky-Golay (SG) algorithm is fitting a low degree polynomial in least squares sense on the samples within a sliding window (Abraham and Golay 1964; Steinier, Termonia and Deltour 1972; Dombi and Dineva 2018; Zhu and al. 2017). The main advantage of the SG-filter in contrast to the classical filters which require the characterization and model of the noise process, is that both the smoothed signal and the derivatives can be obtained by a simple and fast calculation.
According to (Abraham and Golay 1964; Steinier, Termonia and Deltour 1972; Dombi and Dineva 2018), The SG filter formalism is given by (1). The input data x_j, y_j is assumed to be equally spaced were $j = 1, \dots., n$ with n the amount of data.

$$g_i = \sum_{i=-m}^{m} c_i \, y_k + i \tag{1}$$

where the window length $M = 2m + 1$, $i = -m, \dots, \lambda, \dots, m$ and λ denotes the index of the center point. A table of SG coefficients is obtained using M = $2m$ + 1 as a window length and k denotes the polynomial degree (Abraham and Golay 1964; Steinier, Termonia and Deltour 1972; Dombi and Dineva 2018).
Usually the Savitzky-Golay filters perform well by using a low order polynomial with long window length or a low degree with a short window. Nonetheless, it is possible to further improve the efficiency with an adaptive smoothing approach based on the classic SG filtering technique that ensures acceptable performance independent of the type of noise process (Dombi and Dineva 2018).
The contribution made in this work tend to automatically calculate the sliding window M value which conserve only the appearance of the signal by annulling noise and outlier's components. The polynomial order is fixed as 2 to avoid any intense wiggling in the smoothed signal. The proposed calculation of the M value is based on the similarity between Kernel estimation and SG filter.
As mentioned in (Bowman and Azzalini 1997), for kernel smoothing regression in case of a normal data vector y the kernel estimator is written as in (2).

$$\tilde{f}(y) = \frac{1}{n} \sum_{i}^{n} w\,(y - y_i; h) \tag{2}$$

Where $\tilde{f}$ the kernel estimator, n the amount of data y and w the kernel function whose variance is controlled by the parameter h. h is called the smoothing function or the bandwidth.
In order to smooth $\tilde{f}$, the optimum formula to calculate h is given by (3).

$$h = \left(\frac{4}{3n}\right)^{1/5} \sigma \tag{3}$$

where σ denotes the standard deviation of the data. The assumption of normality can cause problems of oversmoothing when dealing with non-normal data. To reduce the oversmoothing σ must be adjusted as in equation (4).

$$\tilde{\sigma} = \frac{median\{|y_i - \tilde{\mu}|\}}{0.6745} \tag{4}$$

Where $\tilde{\mu}$ denotes the median of the sample.
In the kernel smoothing regression, the kernel function and h are chosen which is similar to the SG filter where the degree polynomial and the window M must be chosen. This broad analogy between the kernel smoothing regression and the SG filter in the sense that the value of h and M in both methods is reacting the same. For the kernel smoothing regression or the SG filter decreasing M or h makes the data smoother, while decreasing them makes the estimation wiggly. The choose of M is driven by this close similarities between the two methods. The value of M is calculated as in equations (3) and (4).

3.2. Data Sets

The data used for this work is from the Prognostics Data Repository, which is a collection of data sets that have been donated by various universities, agencies, or companies. Mostly these are time series data from some nominal state to a failed state (NASA s.d.).

The data is an engine degradation simulation that was carried out using C-MAPSS (Commercial Modular Aero-Propulsion System Simulation). Four different sets are simulated under different combinations of operational conditions and fault modes. The data set was provided by the Prognostics CoE at NASA Ames (Saxena and Goebel 2008).

The C-MAPSS data set consists of multiple multivariate time series which is divided into training and test subsets. Each subset contains a time series data of 100 engines (a fleet of engines of the same type). In each cycle the data record: the engine number, the cycle number, three operational setting and 21 sensors measurement. The measured data is contaminated with high sensor noise (Saxena, Goebel, Simon and al. 2008).

Our objective is to enhance the prediction of the remaining operational cycles and to provide the vector of true Remaining Useful Life (RUL) values for the test data using only the sensor data. Only the first set of data will be used to test the proposed approach.

3.3. Data-Driven Method

Regarding it successful use in many works, the nonlinear autoregressive network with exogenous inputs (NARX) is used in this work to estimate directly the RUL value from data. The NARX network was successfully applied to estimate the RUL value from operation data in various applications (Bektas and Jones 2016; Rai and Upadhyay 2017; Santoso, Prahasto and Widodo 2013).

NARX is a recurrent dynamic network which has exogenous inputs with feedback connections enclosing several layers of the network (Billings 2013). As shown in figure 1, NARX model is based on the linear autoregressive network model, which is commonly used in time-series modeling. The NARX model is defined by equation (5).

$$y(t) = f(y(t-1), y(t-2), \ldots, y(t-n\_y), \quad u(t-1), \ldots, u(t-n\_u)) \qquad (\mathbf{5})$$

where the next value of the dependent output signal $y(t)$ is regressed on previous values of the output signal and previous values of an independent (exogenous) input signal. The NARX network can predict next value for an input signal, filtering nonlinear noise or model a nonlinear dynamic system.

As mentioned in (Bektas and Jones 2016; Narendra and Parthasarathy 1991), the best training processes use the true output instead of feeding back the estimated output, figure 2. This training process enhance the output accuracy of the network and the training is simply performed by static backpropagation. More details on the NARX model used in this work is available in (Bektas and Jones 2016; Billings 2013).

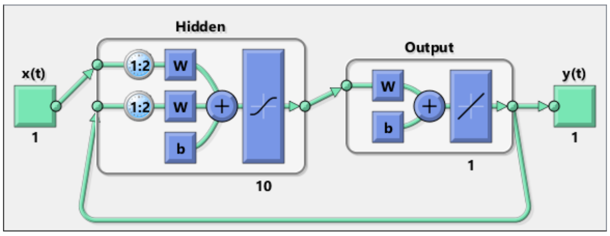

Figure 1: A MATLAB representation of NARX network, where X(t) is the network inputs (measured data) and Y(t) is the network output.

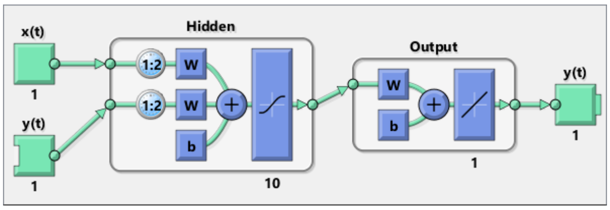

Figure 2: training the NARX network in MATLAB where x(t) is the training data inputs y(t) is the training targets.

4. RESULTS AND DISCUSSION

All the simulations are done using MATLAB R2016a. The first simulation part is testing the proposed SG filter on some signals and compare the smoothed signal with the kernel smoothing regression. Figure 3 shows the application of the proposed SG filter Vs a kernel smoothing regression on benchmark signal (Cichocki and Amari 2003). the filtered signal with SG filter offers more smoothness for the original signal. In figure 4 the signal used is from C-MAPSS data set (Saxena, Goebel, Simon and al. 2008). The resultants signal is very smooth compared to the kernel smoothing regression.

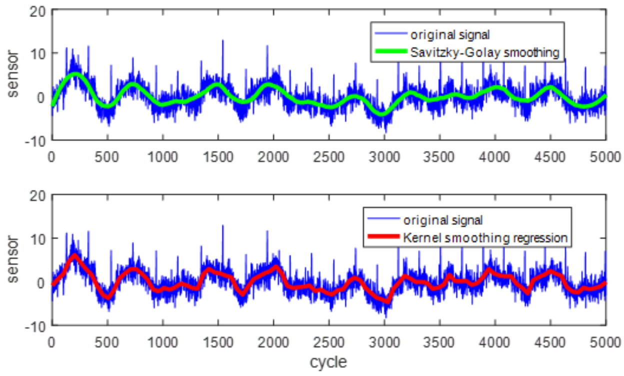

Figure 3:shows the application of the proposed SG filter Vs a kernel smoothing regression on benchmark signal (Cichocki and Amari 2003).

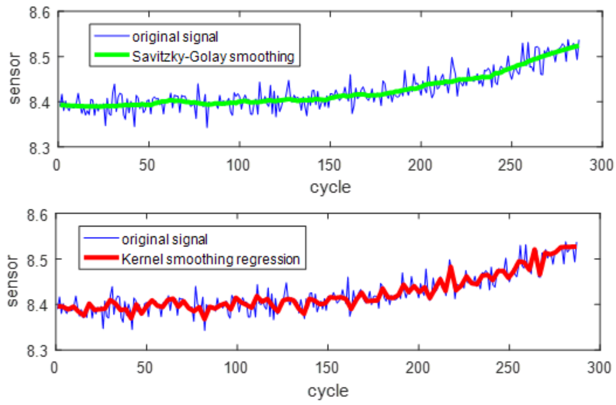

Figure 4:shows the application of the proposed SG filter Vs a kernel smoothing regression on C-MAPSS data set (Saxena, Goebel, Simon and al. 2008).

The simulation steps are:

1. Loading the training data file
2. Choose the best indicators for RUL prediction based on data trendability (Al-Dahidi and al. 2016).
3. Smooth the sensor data with Savitzky-Golay filter.
4. Train multiple NARX networks architecture on open loop (Bektas and Jones 2016).
5. Chose the best network based on the means of training prognostic metrics (Saxena, Celaya and al. 2008).
6. Load the test data.
7. Smooth the test data.
8. Calculate the RUL vector of the test data.

Figures 5 show clearly the performance of the SG-filter on a monotonic indicator of the first engine in the first sets of training data. It can be observed that the smoothing is very good.
The automatic window calculation gives a smoothing which conserve only the appearance of the signal and shows clearly the monocity of data series.

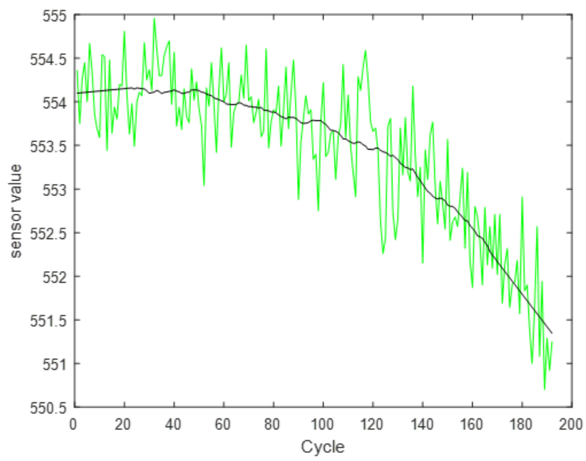

Figure 5: Performance of the SG filter on sensor signal C-MAPSS data (Saxena, Goebel, Simon and al. 2008). clear chart: original sensor signal - black line: smoothed signal.

Prognostic metrics were calculated for evaluation, table 1 contains metric results for test data. The metrics are summarized in detail in (Saxena, Celaya and al. 2008). The results show an excellent prediction for the NARX model when trained with filtered data from SG method. Table 2 shows the twenty-first RUL predicted compared to (Bektas and Jones 2016). It's noticed the superiority of the proposed approach compared to the previous work.

Table 1: Summary of result on test data using (Saxena, Celaya and al. 2008) metrics

Metrics	Results
Score	205
FPR (%)	48
FNR (%)	52
MSE	37
MAE	4.4
MAPE (%)	12.6
Accuracy (%)	79

Table 2: Results of estimated RUL compared to (Bektas and Jones 2016) for the 20 first test engines

Test engine N°	Our work	[4] work	True RUL
1	110	116	112
2	94	112	98
3	64	43	69
4	80	79	82
5	92	88	91
6	87	111	93
7	95	93	91
8	93	107	95
9	107	118	111
10	104	93	96
11	91	85	97
12	131	78	124
13	102	84	95
14	106	98	107
15	80	97	83
16	86	100	84
17	53	52	50
18	23	39	28
19	84	113	87
20	15	26	16

5. CONCLUSION

An approach based on an automatic SG filter and NARX model is introduced in this work. The combination of a powerful data-drive model such the NARX model, applied successfully in many applications, with a high-performance filter with automatic parameters lead to excellent results in modeling health state of a system in PHM application.

The proposed approach is validated using the first data set of C-MAPSS data for gas turbo engine and compared to previous works using the same model. The result shows a considerable improvement on RUL prediction accuracy that indicates the importance of selecting the proper filtering process before training.

The challenge is to adopt this approach with the other data sets of C-MAPSS.

REFERENCES

Abraham, Savitzky, and Marcel J. E. Golay. 1964. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." Analytical Chemistry (American Chemical Society) 36 (8): 1627–1639. doi:10.1021/ac60214a047.

Al-Dahidi, Sameer, Francesco Di Maio, Piero Baraldi, and Enrico Zio. 2016. "Remaining useful life estimation in heterogeneous fleets working under variable operating conditions." Reliability Engineering & System Safety 156: 109–124.

B., Scholkopf, and Smola A. J. 2001. Learning with kernels:support vector machines, regularization, optimization, and beyond. The MIT Press.

Bektas, Oguz, and Jeffrey A. Jones. 2016. "NARX Time Series Model for Remaining Useful Life Estimation of Gas Turbine Engines." EUROPEAN CONFERENCE OF THE PROGNOSTICS AND HEALTH MANAGEMENT SOCIETY 2016.

Benmoussa, Samir, and Mohand Arab Djeziri. 2017. "Remaining useful life estimation without needing for prior knowledge of the degradation features." IET Science, Measurement & Technology 11 (8): 1071 - 1078.

Billings, Stephen A. 2013. Nonlinear System Identification: NARMAX Methods in the Time, Frequency and Spatio-Temporal Domains. Wiley.

Bluvband, Zigmund, and Sergey Porotsky. 2016. Prognostic and Health Management: God Is in Details.

Bowman, Adrian W., and Adelchi Azzalini. 1997. Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations . New York: CLARENDON PRESS OXFORD.

Cichocki, Andrzej, Shun-ichi Amari, Krzysztof Siwek, and Toshihisa Tanaka and al. n.d. http://www.bsp.brain.riken.jp/ICALAB. ICALAB Toolboxes.

Coble, Jamie, and J. Wesley Hines. 2012. "Identifying Suitable Degradation Parameters for Individual-Based Prognostics." In Diagnostics and Prognostics of Engineering Systems_ Methods and Techniques, 145. IGI Global.

Djeziri, Mohand Arab, Samir Benmoussa, and Roberto Sanchez. 2018. "Hybrid method for remaining useful life prediction in wind turbine systems." Renewable Energy 116: 173 - 187.

Dombi, József, and Adrienn Dineva. 2018. "Adaptive Multi-round Smoothing Based on the Savitzky-Golay Filter." International Workshop Soft Computing Applications. 446-454.

Dombi, József, and Adrienn Dineva. 2018. "Adaptive Multi-round Smoothing Based on the Savitzky-Golay Filter." In Advances in Intelligent Systems and Computing, by József Dombi and Adrienn Dineva, 446-454.

Heimes, Felix O. 2008. "Recurrent neural networks for remaining useful life estimation." Proceedings of the Annual Conference on Prognostics and Health Management Society.

javed, kamran, rafael gouriveau, ryad zemouri, and noureddine zerhouni. 2011. "Improving data-driven prognostics by assessing predictability of features." Proceedings of the Annual Conference of the Prognostics and Health Management Society.

K., Javed, Gouriveau R., Zemouri R., and Zerhouni N. 2011. "Improving data-driven prognostics by assessing predictability of features." Proceedings of the Annual Conference of the Prognostics and Health Management Society.

M, Tipping. 2001. "Sparse Bayesian learning and the Relevance Vector Machine." Journal of Machine Learning Research 1 (3): 211–244.

Michael, E. Tipping. 2001. "Sparse Bayesian learning and the Relevance Vector Machine." Journal of Machine Learning Research 1 (3): 211–244.

Narendra, Kumpati S., and Kannan Parthasarathy. 1991. "Learning Automata Approach to Hierarchical Multiobjective Analysis." IEEE Transactions on Systems, Man and Cybernetics 20 (1): 263–272.

NASA. n.d. prognostic-data-repository. NASA. https://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/.

O., Heimes F. 2008. "Recurrent neural networks for remaining useful life estimation." Proceedings of the Annual Conference on Prognostics and Health Management Society.

Rai, Akhand, and S.H. Upadhyay. 2017. "The use of MD-CUMSUM and NARX neural network for anticipating the remaining useful life of bearings." Measurement 111: 397-410.

Santoso, Gunawan Budi, Toni Prahasto, and Achmad Widodo. 2013. "Prognosis of bearing damage performance to industrial system using nonlinear autoregressive with exogenous (NARX)." 2013 International Conference on QiR.

Saxena, Abhinav, and Kai Goebel. 2008. Turbofan Engine Degradation Simulation Data Set. Edited by NASA Ames Research Center. Moffett Field, CA. http://ti.arc.nasa.gov/project/prognostic-data-repository.

Saxena, Abhinav, Jose Celaya, Edward Balaban, Kai Goebel, Bhaskar Saha, Sankalita Saha, and Mark Schwabacher. 2008. "Metrics for evaluating performance of prognostic techniques." 2008 International Conference on Prognostics and Health Management. 1-17.

Saxena, Abhinav, Kai Goebel, Don Simon, and Neil Eklund. 2008. "Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation." Proceedings of the Ist International Conference on Prognostics and Health Management (PHM08). Denver CO.

Schölkopf, Bernhard, and Alexander J. Smola. 2001. Learning with kernels:support vector machines, regularization, optimization, and beyond. The MIT Press.

Steinier, Jean, Yves Termonia, and Jules Deltour. 1972. "Smoothing and differentiation of data by simplified least square procedure." Analytical Chemistry 44 (11): 1906-1909. doi:10.1021/ac60319a045.

T., Wang, Yu J., Siegel D., and Lee J. 2008. "A similarity-based prognostics approach for Remaining Useful Life estimation of engineered systems." Proceedings of the Annual Conference on Prognostics and Health Management Society.

Tianyi, Wang, Jianbo Yu, David Siegel, and Jay Lee. 2008. "A similarity-based prognostics approach for Remaining Useful Life estimation of engineered systems." Proceedings of the Annual Conference on Prognostics and Health Management Society.

Vichare, Nikhil M., and Michael G. Pecht. 2006. "Prognostics and Health Management of Electronics." IEEE Transactions on Components and Packaging Technologies 29 (1): 222-229.

Zhu, Gao Feng, Hong Qiu Zhu, Chun Hua Yang, and Wei Hua Gui. 2017. "Improved Savitzky-Golay filtering algorithm for measuring a pharmaceutical vial's oxygen content based on wavelength modulation spectroscopy." Journal of Optical Technology 84 (5): 355.