# An online hybrid prognostics ANFIS-PF method with an application to gearbox for RUL prediction

Atefeh Govahianjahromi[1], Jaehoon Kim[2], Moussa Hamadache[3] and Dongik Lee[4]

[1,2,4]*Department of Electronics Engineering, Kyungpook National University, Daegu, Korea*

a.gavahian.j@gmail.com
destiny-be@hanmail.net
dilee@ee.knu.ac.kr

[3]*Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul, Korea*

*assuom25@yahoo.fr*

## ABSTRACT

Rotary components are dealing with performance degradation phenomenon, which contains the message of unexpected damages. Therefore, prognostics and health management (PHM), has been introduced to calculate and predict the remaining useful life (RUL) in order to prevent costly damages or repairs.

Data-driven, model-based and hybrid-based techniques are three main categories of PHM techniques. From the health monitoring view, the main idea is to use the experimental run-to-failure data as an intelligence-based model for our gearbox and predict the RUL with the probability (model) based method (Particle Filtering). Firstly, to perform our prognostics technique, we require the degradation information from gearbox. Therefore in duration of 10 days, we conduct a run-to-failure experiment for a test bench with initiative fault injection in day 7$^{th}$. Period of last-three-days is considered as run-to-failure signal for proposed algorithm.

After preprocessing the data, we apply a combined prediction method ANFIS-PF, using Adaptive Neuro Fuzzy Inference System (ANFIS) and Particle Filtering (PF). ANFIS used as a prediction model tool, while the particle filter method was used to find a step-ahead behavior of the gear. ANFIS as a powerful data-driven method will model the prediction of degradation data and finally this model is applied to particle filtering to predict a-step-ahead of the gear behavior until failure will happen.

Meanwhile, some important signal characteristics known as condition indicators (CIs) have been extracted from the residual, energy, frequency based data processing. Then, the energy-based health index (HI) is calculated using threshold and sum of distributions, to show the degradation trend of tested gearbox. The online prediction results properly demonstrate the performances of the proposed ANFIS-PF algorithm, to predict the RUL of gearbox system with a 95% confidence boundary distribution.

## 1. INTRODUCTION

Systems with complex mechanical components (ex: gearboxes) have their own useful operating cycles. Therefore it is inevitable to monitor their life cycles while their performances are decreasing until unexpected damage will happen (Cuc, 2002). Recently, condition-based maintenance (CBM) is a preferred solution for monitoring system's life time, maintenance, reliability, safety and many more engineering problems. Generally, CBM is the process consisting of condition monitoring (CM), diagnostics and prognostics (PHM), and predictive maintenance (PM). Among them, prognostics and health management called PHM, as a key process, predicts the future conditions of failing subsystems (Al-Arbi & S. K. 2012). PHM presents the major challenges to CBM, because it makes use of previous observations, current and future information of a subsystem in order to assess the propagation (degradation), diagnose incipient fault and prediction life cycles before final damage (Vachtsevanos et al, 2006, Zio & Enrico, 2012)

Figure 1 enables the significance of applying prognostics to avoid machines meeting an unexpected damage.

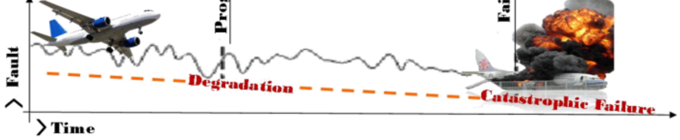



Figure 1. Unexpected failures

The principal goal of prognostics approaches is to calculate the remaining useful life (RUL) of machinery by predicting future propagation at an early level of degradation. Prognostics approaches are classified into three categories: 1) physic-based prognostics, 2) data-driven prognostics, 3) hybrid prognostics (Vachtsevanos et al. 2006). Physic-based or model-based prognostics methods, use explicit mathematical (statistical) formulations known as white box, to formalize physical behavior of degrading components. The data-driven (DD) approaches such as machine learning method or black-box prognostics applies statistical relationship, train and test to the CM data and intelligently provide decision-making information. DD methods rely on the assumption that the statistical characteristics of the system remain steady until an incipient fault occurs in the machine. Finally the hybrid-methods inherit the benefits of both model-based and DD prognostics approaches to find the future behavior of the system (kamran Javed, 2014).

This paper thus, presents a hybrid PHM method applies the Adaptive Neuro Fuzzy Inference System (ANFIS) method and the Particle Filtering (PF) algorithm. ANFIS as a data driven method will model the prediction of degradation data and finally this model is applied to PF to find a step ahead of the gear behavior until failure happens. The remainder of this work is constituted as follows: in the next section, a brief description of ANFIS and particle filtering methods will be given. In Section.3 the proposed hybrid prognostics methodology (integrated ANFIS-PF) will be presented, Section.4 provides the experimental results and finally the conclusion will be given in section 5.

## 2. BRIEF DESCRIPTION OF ANFIS AND PARTICLE FILTER

The model based prognostics process is illustrated in Figure 2, which requires a model of degradation data. Particle filtering as a model based follow the same procedure hence it needs a model, and what usually we get is only the data from sensors. It is necessary to find a model, based -on the data from CM part (Yu, 2012, An et al, 2011).

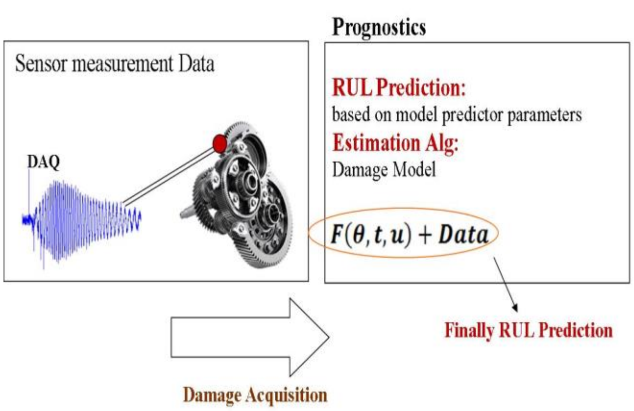
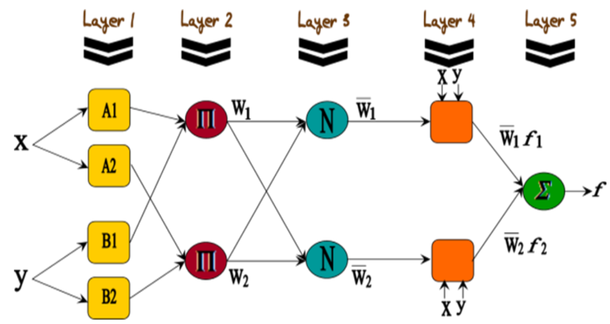



Figure 2. Illustration of the model-based prognostics process

### 2.1. Adaptive Neuro Fuzzy Inference System (ANFIS)

A Neuro-fuzzy system is inspired from the metrics of both neural system and takagi-sugeno (TS). The TS model conveniently provides less number of fuzzy IF-THEN rules. fuzzy inference system which first introduced by sugeno in 1985. The ANFIS structure presented in Figure 3, has been applied to variety engineering systems. It consists of five layers: the fuzzy formation layer, obtaining and-based fuzzy rules layer, normalization of membership functions (MFs) layer, concluding layer and finally the network output layer, respectively (Chen & Matthews, 2013). Table 1 briefly introduces the layers. The following feedforward equations and rules, Eq.(1) and (2) present the ANFIS structure.




$$f = (\bar{w}_1 x)c_{11} + (\bar{w}_1 y)c_{12} + \bar{w}_1 c_{10} + (\bar{w}_2 x)c_{21} + (\bar{w}_2 y)c_{22} + \bar{w}_2 c_{20}$$

Figure 3. Illustration of the model-based prognostics process

where $\bar{w}_i,\ i = 1{,}2$ displays the weight.

Table 1. Description of ANFIS layers

| ANFIS Layers | Description |
|---|---|
| Layer1 | Performing fuzzy formation |
| Layer2 | Performing fuzzy rules with "AND" operator |
| Layer3 | Normalizing the MFs |
| Layer4 | Concluding layer |
| Layer5 | Network output layer |

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y),\ \ i = 1{,}2 \tag{1}$$

$$\bar{w}_i = \frac{w_i}{w_1+w_2}, \begin{cases} f_1 = a_1 x + b_1 y + c_1 \\ f_2 = a_2 x + b_2 y + c_2 \end{cases} \rightarrow f = \frac{w_1 f_1 + w_2 f_2}{w_1+w_2} \tag{2}$$

where $a_i, b_i$ and $c_i$ , $i = 1{,}2$ are the antecedent and consequence parameters of ANFIS structure.

Normalization of membership functions (MFs) in layer 3 plays an important role in clustering the inputs of ANFIS algorithm, thus, in this work, the MFs are Gaussian type and computed as follows:

$$\mu_{A_i}(x) = e^{-\frac{1}{2}(\frac{x-m_i}{\sigma_i})} \quad (3)$$

where, $\sigma_i$ is the width of MFs and $m_i$ is the center of them. ANFIS is well-train when using the gradient decent form and it has expert knowledge. (Chen & Matthews, 2013). Various prognostics for machinery, use four input variables (showing the previous operating conditions) (Chen & Vachtsevanos et al 2011). In presented work, three inputs assigned with 8 IF-Then rules for prediction is designed.

### 2.2. Particle filtering (PF) for prediction

The particle filter is one of the main concerns to solve nonlinear prediction problems. The PF takes the Bayesian technique to perform the dynamic state estimation, in which it attempts to accurately present the probability distribution function (PDF) of given data. Essentially, observed states displayed with PFD values (Yu, J 2012). The stronger (in weights) the prediction PDF is, the more accurate the state estimation will be. An important step to get the robust PF is that of resampling stage, which deletes particles with low weights and multiplies those with high weights. This helps to prevent the degeneracy phenomenon (Orchard 2007, Orguner 2007). The steps of PF algorithm has been explained shortly (Hao et al 2016):

#### 2.2.1. Initialization:

$N$ number of particles are created based on the initial state of the system. They initialized with a Gaussian distribution.

#### 2.2.2. Prediction:

The prior probability distribution $p(x_t|x_{t-1})$ at time $t$ is evaluated thanks to the state model, the evaluated state $x_{t-1}$ at time $t-1$ and the inputs of the system $u_t$, which is the vibration signal.

#### 2.2.3. Update:

In case of new available measurement, the weights are calculated as follows:

$$\omega_k^i \alpha \omega_{k-1}^i . \frac{p(y_k|x_k^i)p(x_k^i|x_{k-1}^i)}{p(x_k^i|x_{k-1}^i, y_k)} \; i = 1,2, ... \quad (4)$$

Weight computation requires the knowledge of the likelihood distribution that measure the matching between each particle and the latest measurement which expression comes from the measurement model;

#### 2.2.4. Resampling:

The particles with low weights are eliminated and the other ones duplicated. Resampling is equivalent to modifying the random measure by improving the exploration of the state space at $t+1$.

Update and resampling are the most delicate tasks of the PF procedure (Orchard 2007, Orguner 2007).

- The likelihood-distribution $p(y_k|x_k^i)$ that measure the matching between each particle ($N = 100$) and the latest measurement which expression comes from the measurement model;
- The prior-distribution $p(x_k|x_{k-1})$
- The importance-density $p(x_k|x_{k-1}, y_k)$

### 3. PROGNOSTICS USING ANFIS-PF (ONLINE/OFFLINE)

Model-based methods rely on the availability of a good dynamic model of the system. Such models can be derived using either physical modeling, or online parameter estimation, system identification or by using artificial intelligence methods. In our work, the degradation model created by ANFIS, as was shown in Figure 2, which it is expressed as a function of usage conditions $U$, elapsed time $t$, and model parameters $h$.

Then, the states of model parameters were predicted by the PF algorithm (Orchard & M. E. 2007). But in case of data-driven techniques, some models have to be created based on a given data. In this work the model is the fuzzy-inference-system. Integration of destructed model, with degrading data displays system's health status. The cycles to failure is known as remaining useful life (RUL) and should be predicted to help operators not to meet the unwanted damages. In this paper we used ANFIS as a damage-model-predictor, which goes to PF to estimate the health state. Following previous works (Chen & Vachtsevanos et al 2011), in this work we also used three-input ANFIS estimator to find the current state and applying it to PF that will give us a step-a-head prediction as follows.

$$\{x_{t-3}, x_{t-2}, x_{t-}, x_t\} \quad (5)$$

The fault growth formulation is presented as follows:

$$\hat{x}_k + w_{k-1} \quad (6)$$

where, $w_{k-1}$ is the process noise and $\hat{x}_k$ is the current state and can be calculated as:

$$\hat{x}_k = y_k(x_{k-1}, x_{k-2}, x_{k-3}, x_k) \quad (7)$$

ANFIS nonlinearly can find the $\hat{x}_k$ within the process noise which follows the Gaussian distribution behavior. Where, $y_k$ is a non-linear mapping function which maps the nonlinear relationship between input and output.

The detailed steps for prognostics are given bellow:

***Step1:*** The health index from extracted features called condition indicators (CIs) of raw data, is created (Qu & Bechhoefer 2012). The analysis algorithms applying to vibration/acoustic signal are, time-based analysis, frequency and amplitude modulation analysis, energy-based analysis and residual analysis and some of important extracted CIs from analysis algorithms are: Root-mean-square, peak-to-peak, crest-factor, Kurtosis.

***Step2:*** Using the combination of all CIs energy, to find a single signal to show the degradation trend and threshold. This single signal is known as health index (HI), (He & Bechhoefer 2012). HI is constructed using two levels of alerts, called Alarm and warning. We consider that the HI is normalized between zero and one and is defined as:

$$HI = \frac{0.5}{THR}\sqrt{\sum_{i=1}^{n} Y_i^2} \tag{8}$$

where sum of Raleigh distribution of CIs is donated by $\sum_{i=1}^{n} Y_i^2$ and THR is the threshold, and $n$ is number of CIs (Bechhoefer & He 2011, September).

***Step3:*** Off-line training for ANFIS, 80% of available data used for training and the rest for the test. The training phenomenon is continuously done until the best and optimal values of $a_i, b_i$ *and* $c_i$ , $i = 1,2$ parameters defined in Eq.(2) are achieved with lower RMSE.

Then it will be used as the model-predictor for Particle filtering. The ANFIS a three-order trained one (which has to be updated for training part).

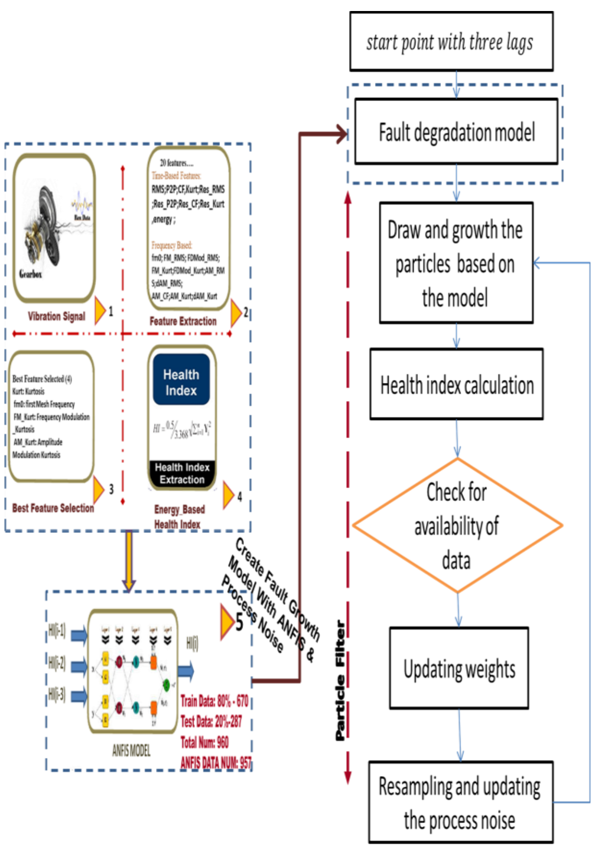



Figure 4. Flowchart of proposed ANFIS-PF algorithm.

***Step4***: On-line, prediction of RUL, with the help of the fault growth model for any 4th order available data ANFIS-PF will predict the a-step-ahead.

Numbers of Multi-step-ahead prediction which is out of this paper scope prediction also can be obtained by taking the expectation of the update model (Chen & Vachtsevanos et al 2011). The Gaussian distribution $p \sim N(\mu_p, \sigma_p^2)$ as a process noise also is updating.

The proposed ANFIS-PF machine-condition-prognostics algorithm flowchart is shown in Figure 4.

## 4. EXPERIMENTAL RESULTS

The proposed ANFIS-PF for prognostics is employed to a drive-train including transmission gearbox (Figure 5). The test rig consists of a gearbox with 4 gears (2-pinions and 2-spurs). As an initiative fault applying to drive-train, we cut (10%) a high-speed-shaft side pinion gear tooth, and we run it continuously almost to failure for 9 days (Figure 6). We only used the last three days of the whole experiment which we faced starting-of-degradation. The goal was applying the day-seven-data to the proposed method to predict the day-eight and day-nine status.

Four acceleration sensors and an acoustic emission sensor were mounted vertically and horizontally on the gearbox case high-speed-side and low-speed-side, to have the metrics of both vibration and acoustic information.

The accelerometers were the ICP industrial accelerometer model No. IMI 608-A11, with a sampling device of (NI-DAQ 9234), demodulation board of AD8339 and acoustic emission sensor of R15a 150KHZ W/SMA CONN – ALPHA.


Figure 5. The drive-train includes gearbox and two motors.

After preprocessing, the vibration and acoustic emission raw signals, which a sample of them is shown in Figure 7, some important characteristics known as condition indicators (CIs) has been extracted from the data, which are, root mean square (RMS), crest factors and kurtosis from, the residual, energy, and frequency based data processing.

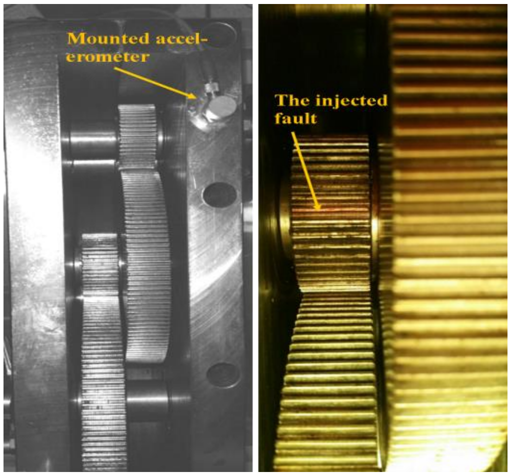



Figure 6. Gearbox of the test bench and the injected initiative injected fault on pinion gear.

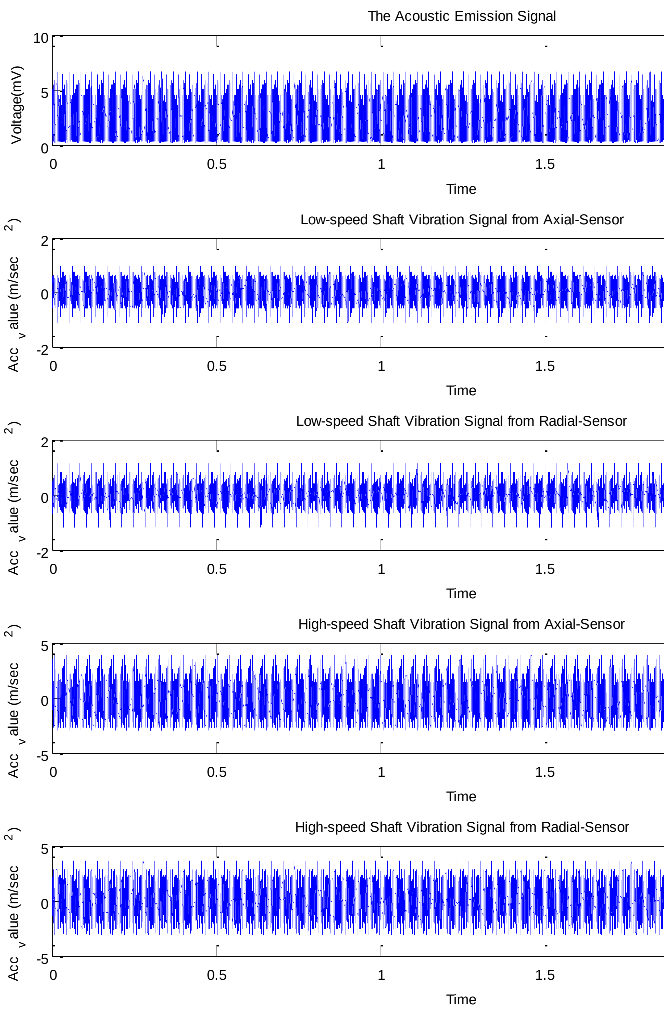



Figure 7. Sample of captured vibration and AE raw data.

Among 45 CIs, best four features were selected named, (a) Residual RMS from vibration signal (CI1), (b) Residual RMS from AE signal (CI2), (c) Amplitude Modulation Kurtosis and (d) Energy based Kurtosis from vibration data (Figure 8). As we need one signal containing useful information of all those CIs, an energy-based HI has been calculated using the threshold and the sum of distributions. The 95% confidence boundary is also determined.

The health index (HI), shows the condition of the gear based on the important CIs. From the whole run-to failure data, we consider the HI (Figure 9) for the first 7 days of experiment for training and testing the FIS-model, then we entered it to ANFIS-PF for 2 days. It should be noted that the next step was defined by next minute.

We consider normalized threshold, therefore we meet two alert levels: warnings and alarms. The probability of false alarm (PFA) is generated by exceeded HI from 0.5. the warnings and alarms are deals with probability of exceeded HI from (0.75) and (1.0), respectively (Bechhoefer & He, 2011).

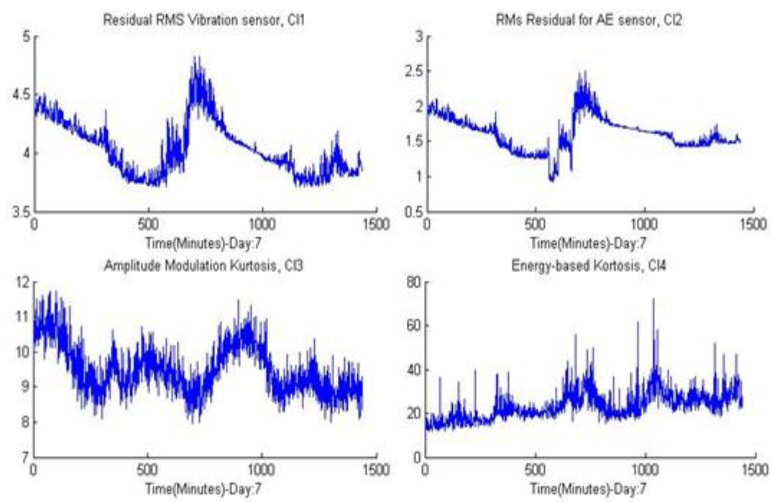



Figure 8. Four best condition indicators.

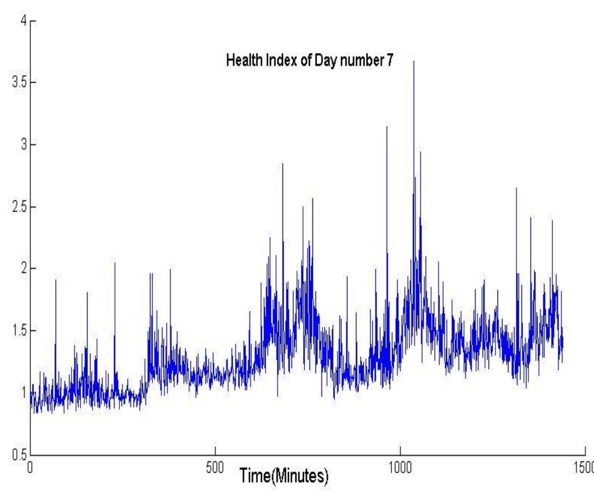



Figure 9. The HI for 7 days

In this work, the HI of day-seven is given to train and test (Figure 10) our system model (ANFIS) and the particle filter uses the best model of our work. This model is saved and updated every 2 hours.

In the next step the particle filter will estimate the next minute of HI using the current data (day-eight and day-nine are) till the failure stage in day-nine as can be seen in Figure 11.

Hence the confidence boundary has been applied to the online prediction, within a step-ahead of the failure behavior.

In particle filtering stage for every minutes by using the particles (Num-of Particles=100) exceeding the mentioned threshold is checked to calculate PFA.

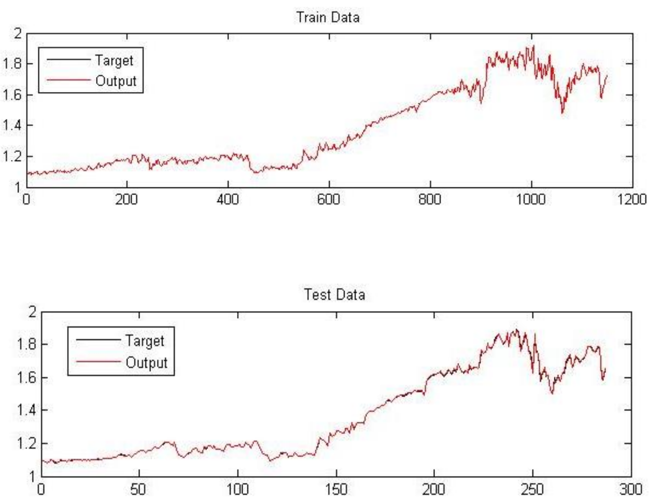



Figure 10. ANFIS Train & Test stages.

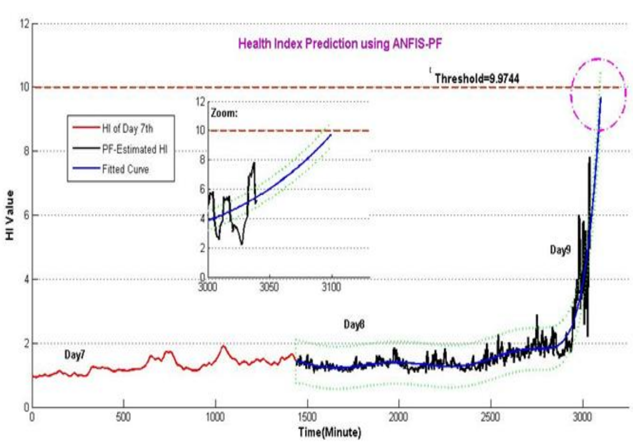



Figure 11. Prediction of Health index using ANFIS-PF and Confidence bounds.

## 5. CONCLUSION

This paper addressed an online prognostics methodology, the ANFIS-PF algorithm, which uses an artificial intelligence method and statistical algorithm, ANFIS and particle filtering, respectively. The best trained ANFIS applied as a predictor model with three lags, which toke the historical data for train and test procedures. Meanwhile the FIS model was exploited to particle filtering for mapping states and to find a step ahead of the gear status. To validate the performance of proposed ANFIS-PF algorithm, a drive-train with fault injection on the gearbox, was run for 9 days –to failure built. Best selection of features (four) created the single-signal (HI), which represents the health condition of the gearbox life cycles. The system of HI was modeled using ANFIS, and a step a-head prediction was achieved using the particle filtering algorithm. The designed ANFIS-PF algorithm is a hybrid-prognostics tool, consists of data driven and model based methods. Thus, it inherited both their metrics. The accuracy of the proposed method is predicting the gearbox RUL.

## ACKNOWLEDGEMENT

This research was supported by the International Research & Development Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning of Korea (Grant number: NRF-2012K1A3A7A03057508)

## REFERENCES

Cuc, A. I. (2002). Vibration-based techniques for damage detection and health monitoring of mechanical systems (Doctoral dissertation, University of South Carolina).'

Vachtsevanos, G., Lewis, F. L., Roemer, M., Hess, A., & Wu, B. (2006). Intelligent fault diagnosis and prognosis for engineering system. Hoboken, NJ: John Wiley & Sons, Inc

Orguner, U., & Demırekler, M. (2007). Analysis of single Gaussian approximation of Gaussian mixtures in Bayesian filtering applied to mixed multiple-model estimation. International Journal of Control, 80(6), 952-967.

Orchard, M. E. (2007). AParticle filtering-based framework for on-line fault diagnosis and failure prognosis (Doctoral dissertation, Georgia Institute of Technology).

Bechhoefer, E., Clark, S., & He, D. (2010). A state space model for vibration based prognostics. In Annual Conference of the Prognostics and Health Management Society (pp. 10-16).

Chen, C., Zhang, B., Vachtsevanos, G., & Orchard, M. (2011). Machine condition prediction based on adaptive neuro–fuzzy and high-order particle filtering. IEEE

Transactions on Industrial Electronics, 58(9), 4353-4364.

Bechhoefer, E., He, D., & Dempsey, P. (2011, September). Gear health threshold setting based on a probability of false alarm. In Annual Conference of the Prognostics and Health Management Society (Vol. 2011).

Yu, J. (2012). Health condition monitoring of machines based on hidden Markov model and contribution analysis. IEEE Transactions on Instrumentation and Measurement, 61(8), 2200-2211.

Zio, Enrico. (2012) "Prognostics and health management of industrial equipment." Diagnostics and prognostics of engineering systems: methods and techniques: 333-356

He, D., Bechhoefer, E., Dempsey, P., & Ma, J. (2012). An integrated approach for gear health prognostics.

Al-Arbi, S. K. (2012). Condition Monitoring of Gear Systems using Vibration Analysis (Doctoral dissertation, University of Huddersfield).

Qu, Y., Bechhoefer, E., He, D., & Zhu, J. (2012). A new acoustic emission sensor based gear fault detection approach. International Journal of Prognostics and Health Management, 4(2), 32-45.

Chen, B., Matthews, P. C., & Tavner, P. J. (2013). Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS. Expert Systems with Applications, 40(17), 6863-6876.

An, Dawn, Joo-Ho Choi, and Nam Ho Kim. (2013) "Prognostics 101: A tutorial for particle filter-based prognostics algorithm using Matlab." Reliability Engineering & System Safety 115: 161-169.

kamran Javed. (2014). A Robust and Reliable Data-driven Prognostics Approach Based on Extreme Learning Machine and Fuzzy Clustering (Doctoral dissertation).

Hao, L., Bian, L., Gebraeel, N., & Shi, J. (2016). Residual Life Prediction of Multistage Manufacturing Processes With Interaction Between Tool Wear and Product Quality Degradation. IEEE Transactions on Automation Science and Engineering.

## BIOGRAPHIES

**A. Govahianjahromi.** Received the M.S. degrees in electronics engineering from DECS Lab, Kyungpook National University, Daegu, Korea, in 2015 and currently, she is a Ph.D candidate at same laboratory, meanwhile she works as a condition monitoring engineer at iba-Korea System Co., Ltd. Her research interest focuses on the prognostics and health management (PHM), condition monitoring (CM) and condition based maintenance (CBM).

**Dongik Lee.** Received the B.S. and M.S. degrees in control engineering from Kyungpook National University, Daegu, Korea, in 1987 and 1990, respectively, and the Ph.D degree in complex systems control engineering from Sheffield University, Sheffield, U.K., in 2002. He is currently a Professor of the School of Electronics Engineering, Kyungpook National University. His research interest focuses on the design of real-time control & monitoring system for various safety-critical applications, including intelligent robots, wind turbines, and automobiles.