Remaining Useful Life estimation methods for predictive maintenance models: defining intervals and strategies for incomplete data

A. Delmas*, M. Sallak*, W. Schön* and L. Zhao**

** Laboratoire Heudiasyc, Sorbonne Universités, Université de Technologie de Compiègne, CNRS, Heudiasyc UMR 7253, CS 60319, 60203 Compiègne Cedex, France {adrien.delmas, mohamed.sallak, walter.schon}@utc.fr*
*** SECTOR group, 12 avenue du Québec BP 636 Villebon sur Yvette, 91965 Courtaboeuf cedex, France linda.zhao@sector-group.net*

Abstract. Predictive maintenance allows the drawbacks of corrective and preventive maintenance to be overcome by estimating the Remaining Useful Life (RUL) of components. This value helps undertake required tasks at the right time: when the component is deteriorated but right before a failure occurs.
We propose an imprecise RUL estimation method with the use of Hidden Markov Models (HMM). HMMs have been used for prognosis in several studies. Here, we fit the training data by a polynomial and then we frame it with 2 other polynomials of the same degree. This results in 3 distinct trained HMMs: the first is used for the RUL estimation and the other 2 provide the lower and higher bounds. Finally, we propose a strategy that can be adopted when dealing with incomplete or imprecise data. This work has been applied to the IEEE 2008 PHM challenge which was a competition aiming at evaluating prognosis methods.

1. Introduction

Maintenance is an essential step in the life cycle of every complex system. The European standard BS EN 13306 (BSI, 2010) defines maintenance as the *combination of all technical, administrative and managerial actions during the life cycle of an item intended to retain it in, or restore it to, a state in which it can perform the required function.* In other words, maintenance allow ensuring the nominal operating state of the system in conditions that guarantee the safety of the user and the environment, along with its availability. However, it is usually pretty expensive and the system is not available during maintenance tasks. Moreover, current habits are hardly ever optimal:

- Corrective maintenance involves waiting for a failure before repairing or replacing the affected components
- Preventive maintenance involves replacing the components at regular intervals, according to a fixed planning

The former strategy usually induces additional costs (Lee, Ni, Djurdjanovic, Qiu, & Liao, 2006) and an important unavailability of the system (Palem, 2013). Besides, a failure can raise safety issues as the systems can no longer perform all or part of the functions it was designed for. Preventive maintenance reduces significantly the risk of unexpected failures as well as the downtimes but can cause high maintenance costs because each operation is performed without knowing the actual state of the system. Therefore, some components are replaced whereas they are still in a nominal operating state or slightly degraded (Lee, Ni, Djurdjanovic, Qiu, & Liao, 2006) (Le, 2016).
In order to overcome these drawbacks, a new form of maintenance has emerged: predictive maintenance (Palem, 2013). It consists of a regular surveillance of the organs of the system (Bartelds, et al., 2004), which allow evaluating their health state and their proper functioning. It is then possible to forecast the appearance of a system failure before it occurs (Lee, Ni, Djurdjanovic, Qiu, & Liao, 2006) (Le, 2016) (Jardine, Lin, & Banjevic, 2006). Consequently, the adequate maintenance tasks can be planned at just the right time (Mercier & Pham, 2012).
The interest in prognosis has been growing for a few decades and many solutions have been developed, tested and improved over the years. These methods can be classified into 3 main categories (Le, 2016) (Vachtsevanos, Lewis, Roemer, Hess, & Wu, 2006) (Soualhi, Razik, Clerc, & Doan, 2014) (Roemer, Byington, Kacprzynski, & Vachtsevanos, 2006):

- *Physical model-based prognosis* that rely on mathematical or physical models of the degradation phenomenon.

- *Evolutionary or trending models* that are based on temporal data sets and indicators that can describe the evolution of the health state.
- *Experience-based prognosis* which is based on the experience and the expert knowledge learned, for example, on the history of failures or degradation of the system

Data driven models allow performing prognosis from data sets produced by means of surveillance. In particular, stochastic models are of interest for RUL estimation because their implementation is quite simple and understandable. Among them, Markov process are dynamic and multi-state models that can model a degradation pretty faithfully. Moreover, hidden Markov models (HMM) take into account the fact that it is not possible to measure directly the degradation of a component. Besides, these models have proven to be effective in various domains and have been used for predictive maintenance purposes in several studies (Sikorska, Hodkiewicz, & Ma, 2011) (Tobon-Mejia, Medjaher, Zerhouni, & Tripot, 2011) (Geramifard, Xu, Zhou, & Li, 2012) (Tang, Makis, Jafari, & Yu, 2015). They are based on Markov chains that represent the health state of the studied component.

However, these studies consider simplified cases, in particular, perfect data sets, whereas in real applications the data can be altered. For example, a sensor failure might result in some imprecise or even missing measurements, or several sensors can give conflicting information (Krüger, Measures of Conflicting Evidence in Bayesian Networks for Classification, 2013) (Krüger, Detection of Failing Sensors by Conflicting Evidence in Bayesian Classification, 2014).

This paper proposes a general RUL prediction method with the use of HMM. This RUL is provided with an interval thanks to a new methodology. Finally, we propose a strategy that can be adopted when dealing with incomplete or imprecise data.

The paper is organized as follow: a second section presents HMM and their use for RUL estimation. In the third section, the method is applied to the 2008 IEEE PHM challenge which was a competition aiming at evaluating the performance of prognosis models. The fourth section presents our method for determining the interval. Finally, the last section proposes our strategy in case of uncertain data.

2. HMM for RUL estimation

An HMM can model the degradation of a component, each state of the model representing a health state of the component. Such a model is composed of a Markov process with states that are not directly observable but can be revealed thanks to collected data called observations.

Knowing the model and the probabilities of transitions, it is possible to calculate the mean time of sojourn in each state of the model. Therefore, once the health state of the component has been estimated, we can deduce its RUL (Sikorska, Hodkiewicz, & Ma, 2011).

An HMM is described by the following parameters (Rabiner, 1989) (Degirmenci, 2014):

- The state transition probability distribution $A = \{a_{ij}\}$, where a_{ij} is the transition probability from state number i to state number j. In other words, it is the probability that the system be in the state S_j at the time $t+1$ knowing that it is in the state S_i at the time t.
- The observation probability distribution of each state $B = \{b_j(k)\}$, where $b_j(k)$ is the probability of producing the observation v_k knowing that the system is in the state S_j.
- The initial distribution $\pi = \{\pi_i\}$, where π_i is the probability that the system is in the state S_i at the initial time.

A HMM is usually defined by the 3 matrices and is written $\lambda = (A, B, \pi)$.

The estimation of a RUL requires the estimation of the current health state of the component: we have an observation sequence and a trained HMM that represents the component and we would like to know the current state the model is in. Among the existing methods, the Viterbi algorithm gives the most optimal sequence of states according to the maximum likelihood criterion. The last state of this sequence corresponds to the current health state.

Once the current health state known, we can estimate the RUL of the component. In most cases, the final state of the HMM is terminal, meaning that once in this state, the model as a 100% chance of staying in this state.

This means that the transition probability matrix has the following form: $A = \begin{pmatrix} Q & R \\ 0 & 1 \end{pmatrix}$, where Q is a square matrix of size $N-1$, R is a matrix of size $(N-1) * 1$ and 0 is a zero matrix of size $1 * (N-1)$.
By recurrence, $A^n = \begin{pmatrix} Q^n & [I + Q + \cdots + Q^n]R \\ 0 & 1 \end{pmatrix}$
We have $\lim_{n\to\infty} Q^n = 0$, so $\lim_{n\to\infty} A^n = \begin{pmatrix} 0 & FR \\ 0 & I \end{pmatrix}$, where $F = (I - Q)^{-1}$ is called the fundamental matrix of the model. Each element f_{ij} of F is the mean time of the transition from state number i to state number j.
Then, from any state i, the mean time before entering the failed state (state N) is given by: $E_i = \sum_{j=1}^{N} f_{ij}$
Therefore, the RUL can be obtained by subtracting the time spent in the current state from the mean time before entering the failed state.

3. Application to the IEEE 2008 PHM challenge

This method has been applied to the IEEE 2008 PHM challenge which is a competition sponsored by IEEE in order to evaluate prognostic models (Le, 2016). It consists of 3 data sets: one for model training, one for assessing and improving the model, and the last one for evaluating and comparing the models. The data was built by simulation with NASA's C-MAPSS model (Commercial Modular Aero-Propulsion System Simulation). This model simulates an airplane engine. It takes 3 parameters as inputs and gives 21 sensor measurements as outputs (Saxena, Goebel, Simon, & Eklund, 2008). Figure 1 presents the engine as simulated by the model.

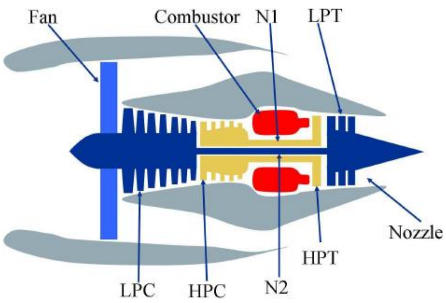

Figure 1. Airplane engine as simulated for the 2008 PHM challenge (Saxena, Goebel, Simon, & Eklund, 2008)

In order to create a HMM describing the degradation of the engine, we need an indicator of this degradation. In this study, we use the same indicator as (Le, 2016) which was originally proposed by (Le Son, Fouladirad, Barros, Levrat, & Iung, 2013). It is constructed from a selection of 7 of the 21 sensors. Figure 2 show this indicator for one of the 218 units that make the training data set.

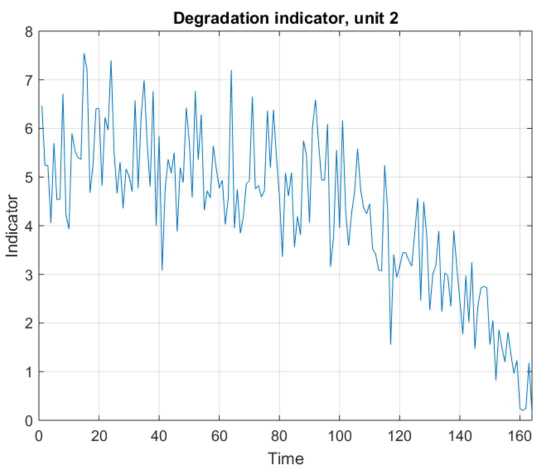

Figure 2. Degradation indicator for the unit number 2

Once the indicator defined, the next step is to build the HMM that will model the engines. To do so, we need to determine the health state and the observation emitted at each time step. Let $S_i(t)$ and $V_i(t)$ be respectively the state of unit number i and the observation emitted by unit i at time t. To determine $S_i(t)$ we proceed as follow:

- We fit the curve of the health indicator by a 3rd degree polynomial P_i
- We define $P_{max}(i) = \max_{1 \le t \le T^i_{failure}} (P_i(t))$, where $T^i_{failure}$ is the failure time of unit i. Let P_{MAX} be the 95th percentile of P_{max}
- The value given to $S_i(t)$ depends on the value of $P_i(t)$ compared to some thresholds:
  - If $P_i(t) \ge \frac{N-2}{N-1} P_{MAX}$ then $S_i(t) = 1$ (nominal state)
  - If $\frac{N-k}{N-1} P_{MAX} > P_i(t) \ge \frac{N-k-1}{N-1} P_{MAX}, 2 \le k \le N-2$ then $S_i(t) = k$
  - If $P_i(t) \ge \frac{1}{N-1} P_{MAX}$ then $S_i(t) = N-1$
  - $S_i(T^i_{failure}) = N$

Figure 3 (left) shows how the states are defined for a 6 states HMM

For the observation values, the process is quite similar:

- Let $D_i(t)$ be the indicator made by unit number i at time t, and $D_{max}(i) = \max_{1 \le t \le T^i_{failure}} D_i(t)$
- Let D_{MAX} be the 95th percentile of D_{max}
  - If $D_i(t) \ge \frac{N-1}{N} D_{MAX}$ then $V_i(t) = 1$
  - If $\frac{N-k}{N} P_{MAX} > D_i(t) \ge \frac{N-k-1}{N} P_{MAX}, 1 \le k \le N-1$ then $V_i(t) = k+1$

Figure 3 (right) shows how the observation symbols are defined for a 6 states HMM

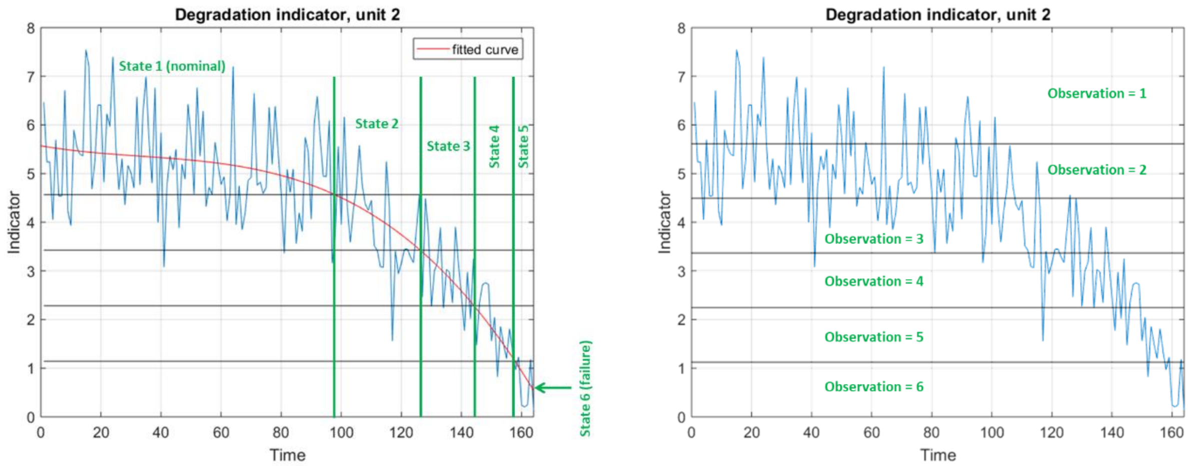

Figure 3. Definition of the states (left) and the observation symbols (right) for each time step for unit number 2

These steps result in the creation of a data set we can actually use for model training. Therefore, the next step is to train our HMM with this data thanks to the Baum-Welch algorithm.

We now have a trained model of the engines and can begin estimating the RULs with the 2nd data set. As before, we construct the health indicator from the 7 chosen sensors and we follow the same procedure to define the observation symbols. Then, we use the Viterbi algorithm to find the most likely state sequence and we take the latest sate of this sequence: it is the current health state of the studied engine. Finally, we calculate the RUL thanks to the method described earlier.

4. Defining an interval for the RUL estimation

Like every estimation, the RUL obtained as described in the previous section is a value subject to uncertainties. Therefore, it is important to evaluate the scale of this uncertainty. Consequently, in this section we propose a method to determine an interval for the RUL values. The idea is to create 2 other HMMs that will give an upper and a lower bound for the RUL. The procedure is as follow:

- We calculate the error between the indicator D_i and the fitted polynomial P_i at each time step: $dist_i(t) = P_i(t) - D_i(t), 1 \le i \le N_{units}, 1 \le t \le T^i_{failure}$

- This error is split in 2: $dist_{sup}^{i}$ and $dist_{inf}^{i}$ that store respectively the positive and negative values of $dist_i$
- We calculate the mean values of $dist_{sup}^{i}$ and $dist_{inf}^{i}$ over a sliding window. Let's call them $dist_{sup}^{i}$ and $dist_{inf}^{i}$
- We create 2 new polynomials: $\begin{cases} P_{sup}^{i}(t) = P_i(t) + \overline{dist_{sup}^{i}} \\ P_{inf}^{i}(t) = P_i(t) + \overline{dist_{inf}^{i}} \end{cases}$
- These new polynomials result in 2 new sets of state sequence for each unit in the training data set. We then train the model as earlier and obtain 2 new HMMs. One of them will give an upper bound and the other will give a lower bound for our estimated RUL.

Figure 4 (left) shows the fitted curve P_i along with the upper and lower bounds P_{sup} and P_{inf} for one particular units in the training data set.

Knowing the RUL along with its interval, we can plot the survival function of the studied engine with the upper and lower bounds. Figure 4 (right) shows this survival function and the 2 bounds for one particular unit from the test data set. We chose to present the results as survival functions because it is easier to interpret and it provides enough information for maintenance policy implementation.

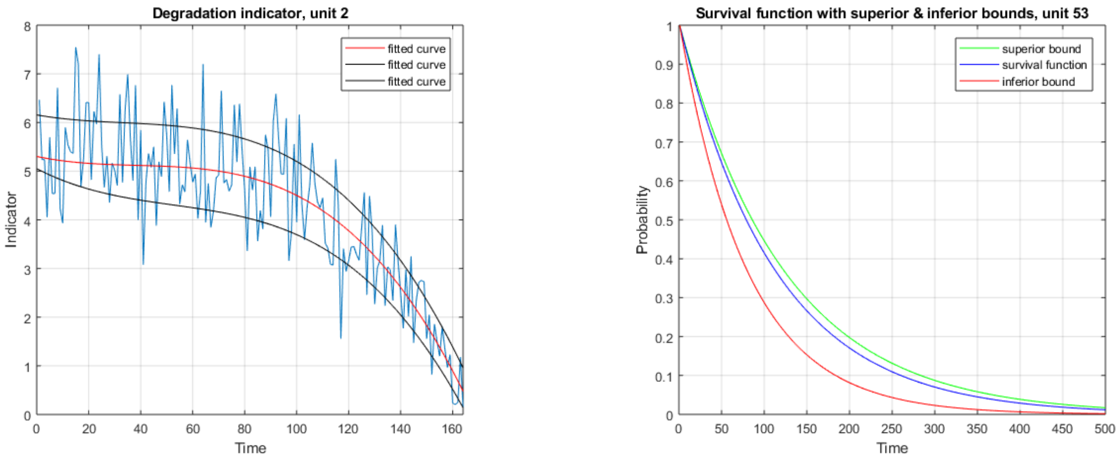

Figure 4. Upper and lower bounds of the polynomial P_i (left). Survival function obtained for unit number 53 with the upper and lower bounds (right)

5. Dealing with uncertain data

Our RUL estimation method is based on a data set that can describe the health state of the component. Therefore, it relies highly on the quality of the collected data and on the confidence we can grant it. However, part of this data can be altered for various reasons and this may have an impact on our results. A HMM being memoryless by definition, if the uncertain data is in the beginning of the observation sequence it will not influence the RUL estimation. But if this data is quite recent, it can impact the result. Consequently, we propose several strategies to overcome this issue:

- Safe attitude: we consider that every piece of uncertain data is an observation of the maximum value (6 in the example presented above). In other words, we consider the worst-case scenario and maximize the safety
- Unsafe attitude: we consider that every piece of uncertain data is an observation of value 1. In other words, we consider the most optimistic scenario
- Medium attitude: we consider that every piece of uncertain data is an observation of mean value (3 or 4 in our example).
- Blind attitude: we make no assumption about the uncertain data. We know each piece of data belongs to the set $\{1,2,\ldots,M\}$ (in our example $M = 6$) so we consider every case and search the maximal and minimal values of the RUL of the component.

These strategies were applied to the 2008 PHM challenge. We artificially altered the data of some units from the test data set. Let T_f^i be the time of the latest piece of data available for unit i. We consider that all the data collected between $T_f^i - 10$ and $T_f^i - 5$ included is damaged and we apply our strategies. Figure 5 shows the survival function obtained by the different attitudes described earlier. The blind attitude considers every possible case, in particular the worst and the most optimistic ones. Therefore, the maximal and minimal values given by this attitude are the same as the unsafe and safe attitudes respectively.

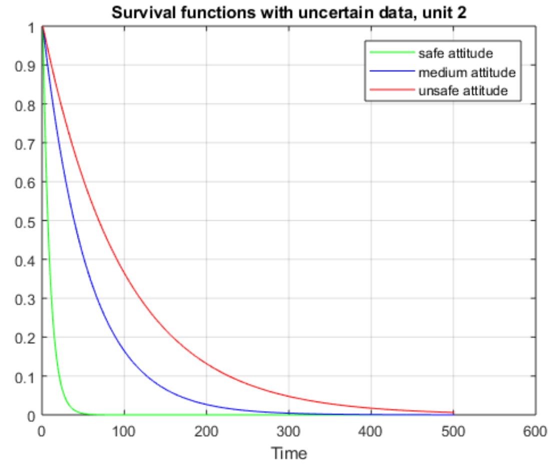

Figure 5. Survival functions obtained with our strategies when dealing with uncertain data

This figure highlights the benefits of such strategies for decision making. Indeed, it helps chose the right time for the maintenance tasks. As an example, we consider that the maintenance should be done when the component has a 0.5 chance to fail. Then, the safe attitude will result in the planification of the maintenance in a very close future while the unsafe attitude will recommend to do the maintenance within the next 70 time steps. In practice, whether one strategy or the other is adopted will depend on several parameters such as the price of the component or whether it is safety related or not.

On the other hand, this figure also shows the limits of the method: as the number of uncertain pieces of date increases, the result of the safe strategy tends toward 0. It can be seen in Figure 5 that the survival function for the safe attitude decreases very quickly towards 0 It shows that the result of the RUL estimation is nearly useless because it gives very little time to plan the required maintenance tasks.

Conclusion

In this paper, we have shown how to calculate components RUL from their degradation state thanks to HMMs. Then, we proposed a solution to give an interval for this estimated RUL. Finally, we proposed a strategy that can be adopted when dealing with uncertain data. These methods have been tested on the IEEE 2008 PHM challenge. The results show they can be significant decision-making support tools thanks to the additional knowledge they provide. However, this application has also underlined some limits.

Future works will try to apply these methods to a real-life case. We will also try to improve them and overcome the drawbacks we discussed in this paper.

References

Bartelds, G., Biemans, C., Boller, C., Breidne, M., Claesson, Å., Delebarre, C., . . . Worden, K. (2004). *Health Monitoring of Aerospace Structures – Smart Sensor Technologies and Signal Processing.* Southern Gate, Chichester, West Sussex PO19 8SQ, England: John Wiley & Sons Ltd.

BSI. (2010). Maintenance - Maintenance terminology. Standard, British Standard Institution.

Degirmenci, A. (2014). Introduction to Hidden Markov Models. Harvard University, 1-5.

Geramifard, O., Xu, J., Zhou, J., & Li, X. (2012). A Physically Segmented Hidden Markov Model Approach for Continuous Tool Condition Monitoring: Diagnostics and Prognostics. *IEEE Transactions on Industrial Informatics, 8*(4), 964 - 973.

Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*(20), 1483-1510.

Krüger, M. (2013). Measures of Conflicting Evidence in Bayesian Networks for Classification. *16th International Conference on Information Fusion.* Istanbul, Turkey.

Krüger, M. (2014). Detection of Failing Sensors by Conflicting Evidence in Bayesian Classification. *17th International Conference on Information Fusion (FUSION).* Salamanca, Spain.

Le Son, K., Fouladirad, M., Barros, A., Levrat, E., & Iung, B. (2013). Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliability Engineering & System Safety, 112*, 165 - 175.

Le, T. (2016). *Contribution to deterioration modeling and residual life estimation based on condition monitoring data.* PhD thesis, Université Grenoble Alpes.

Lee, J., Ni, J., Djurdjanovic, D., Qiu, H., & Liao, H. (2006). Intelligent prognostics tools and e-maintenance. *Computers in Industry*(57), 476–489.

Mercier, S., & Pham, H. (2012). A preventive maintenance policy for a continuously monitored system with correlated wear indicators. *European Journal of Operational Research*(222), 263–272.

Palem, G. (2013). Condition-based maintenance using sensor arrays and telematics. *International Journal of Mobile Network Communications & Telematics, 3*(3).

Rabiner, L. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE, 77*(2), 257 - 286.

Roemer, M. J., Byington, C. S., Kacprzynski, G. J., & Vachtsevanos, G. (2006). An overview of selected prognostic technologies with application to engine health management. *ASME Turbo Expo 2006: Power for Land, Sea & Air.* Barcelona, Spain.

Saxena, A., Goebel, K., Simon, D., & Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. *Proceedings of 2008 international conference on prognostics and health management*.

Sikorska, J., Hodkiewicz, M., & Ma, L. (2011). Prognostic modelling options for remaining useful life estimation by industry. *Mechanical Systems and Signal Processing*(25), 1803-1836.

Soualhi, A., Razik, H., Clerc, G., & Doan, D. (2014). Prognosis of Bearing Failures Using Hidden Markov Models and the Adaptive Neuro-Fuzzy Inference System. *IEEE Transactions on industrial electronics, 61*(6), 2864 - 2874.

Tang, D., Makis, V., Jafari, L., & Yu, J. (2015). Optimal maintenance policy and residual life estimation for a slowly degrading system subject to condition monitoring. *Reliability Engineering and System Safety*(134), 198-207.

Tobon-Mejia, D., Medjaher, K., Zerhouni, N., & Tripot, G. (2011). Hidden Markov models for failure diagnostic and prognostic. *2011 Prognostics and System Health Management Conference.* Shenzhen, China.

Vachtsevanos, G., Lewis, F. L., Roemer, M., Hess, A., & Wu, B. (2006). *Intelligent Fault Diagnosis and Prognosis for Engineering Systems.* Hoboken, New Jersey: John Wiley & Sons, Inc.