

A Switching Ensemble Approach for Remaining Useful Life Estimation of Electrolytic Capacitors

S. Al-Dahidi, F. Di Maio & P. Baraldi

Energy Department, Politecnico di Milano, Milan, Italy

E. Zio

Chair System Science and the Energy Challenge, Fondation Electricité de France (EDF), CentraleSupélec, Université Paris Saclay, Grande Voie des Vignes, 92290 Chatenay-Malabry, France

Energy Department, Politecnico di Milano, Milan, Italy

ABSTRACT: We consider the problem of predicting an equipment Remaining Useful Life (*RUL*), based on data from a fleet of similar components working under different operating conditions. Two data-driven prognostics models are considered, an Homogeneous Discrete-Time Finite-State Semi-Markov Model (*HDTFSSMM*) and a Fuzzy Similarity-Based (*FSB*) model. The predictions from the two models are ensemble with a switching approach that selects the most accurate model at each stage of the equipment life to guarantee the best performance throughout the degradation progression. A real case study is worked out, regarding a fleet of aluminum electrolytic capacitors used in electric vehicles powertrains. In this specific case, the *HDTFSSMM* is selected at the early stages of the equipment degradation and the *FSB* model at the last stages of life. Results are compared with those achieved by each individual model, showing that the proposed ensemble approach is able to provide more accurate *RUL* predictions throughout the entire equipment life.

1 INTRODUCTION

In industries such as nuclear, oil and gas, automotive and chemical, unforeseen equipment failures are extremely costly in terms of repair costs and lost revenues (Baraldi et al. 2015). To anticipate failures, predictive maintenance approaches are being developed, based on the assessment of the actual equipment degradation condition and on the prediction of its evolution for setting the optimal time for maintenance (McCall 1965; Wei et al. 2013; Zhao et al. 2010). The underlying concept is that of failure prognostics, that consists in predicting the Remaining Useful Life (*RUL*) of the equipment undergoing degradation, i.e., the amount of time the equipment can continue performing its functions under its design specifications (Zio 2012; Tobon-Mejia et al. 2011; Jardine et al. 2006).

In practice, efficient failure prognostics avoids the system unscheduled shutdowns, defines efficient maintenance strategies, exploits the full *RUL* for operation and, hence, increases the system availability and safety while reducing maintenance costs (Tobon-Mejia et al. 2011; Al-Dahidi et al. 2015).

Approaches for *RUL* estimation can be generally categorized into model-based and data-driven (Tobon-Mejia et al. 2011; Sankavaram et al. 2009; Chiang et al. 2001). Model-based approaches use physics models to describe the degradation behaviour of the equipment (Chiang et al. 2001; Heng et al. 2009). Despite the fact that these approaches lead

to accurate prognostics results, uncertainty arising due to the assumptions and simplifications of the adopted models may pose limitations on their practical deployment (Chiang et al. 2001; Di Maio et al. 2013). On the other side, data-driven prognostics approaches do not use any explicit physical model, but rely exclusively on the availability of process data related to equipment health to build (black-box) models that capture the degradation and failure modes of the equipment (Heng et al. 2009; Van Tung & Yang 2009).

In fact, the availability of condition monitoring data from similar equipment, forming what in the industrial context is called a fleet (Medina-Oliva et al. 2014), motivates the development of data-driven prognostics approaches that capitalize on the information contained in such data to estimate the equipment *RUL*. In practice, a fleet of P pieces of equipment, which have different and/or similar technical features, can undergo different usage with different operating conditions, e.g., a fleet of highly standardized steam turbines of pressurized water reactors nuclear power plants (Al-Dahidi et al. 2016a), and can provide a wide knowledge concerning the equipment behaviour and, thus, improve the efficiency of the fault prognostics task (Medina-Oliva et al. 2014). However, a main difficulty is that the equipment typically experiences different operating conditions, which influence both the condition monitoring data and the degradation processes that physically determine the *RUL* (Al-Dahidi et al. 2016b).

Therefore, individual data-driven prognostics models might not provide satisfactory predictions of the *RUL* in terms of accuracy: each model can provide accurate *RUL* predictions under some operating conditions but less accurate in others (Hu et al. 2010).

To overcome this, ensemble approaches, based on the aggregation of multiple model outcomes, have been introduced due to the superior robustness and accuracy with respect to single models (Polikar 2006) and the possibility of estimating the uncertainty of the predictions (Baraldi et al. 2013). In this context, the objective of the present work is to develop an ensemble approach composed by different data-driven prognostics models for aggregating their *RUL* predictions, to guarantee the best performance throughout the equipment degradation progression.

Two data-driven prognostics approaches are considered: 1) an Homogeneous Discrete-Time Finite-State Semi-Markov Model (*HDTFSSMM*) (Al-Dahidi et al. 2015; Al-Dahidi et al. 2016b) and 2) a Fuzzy Similarity-Based (*FSB*) model (Di Maio et al. 2013). The former approach entails building a degradation model, estimating its parameters and using the model within a direct Monte Carlo (*MC*) simulation scheme to estimate the *RUL* of an equipment (Cadini et al. 2009), whereas the latter model evaluates the similarity between the test degradation trajectory and the available fleet run-to-failure training trajectories, and to use the *RULs* of these latter to estimate the *RUL* of the former, considering how similar they are (Di Maio et al. 2013).

The ensemble procedure is based on a switching strategy that selects the most accurate model at each stage of the equipment life. More specifically, the approach entails: an offline selection of the optimal switching time t_{opt} that corresponds to the time of which the average value the Accuracy Index (*AI*) on a validation set is minimized (i.e., small *AI* values indicate more accurate predictions) (Saxena et al. 2010); then switching time is used for predicting online the *RUL* of a new equipment of the fleet.

The proposed approach is applied to a real case study regarding a fleet of aluminum electrolytic capacitors used in electric vehicles powertrains. In this specific case, the *HDTFSSMM* is selected to be the most suitable model for *RUL* prediction at the early stage of the equipment degradation, whereas the *FSB* model when the equipment approached the end-of-life. The performance of the proposed approach is compared with the performance of each individual *HDTFSSMM* and *FSB* models showing that the proposed ensemble approach is able to provide more accurate *RUL* predictions throughout the entire life of the equipment.

The remaining of this paper is organized as follows. In Section 2, the two prognostics models are briefly recalled. In Section 3, the proposed ensemble approach is illustrated. In Section 4, the real case

study is described together with the results obtained. Finally, some conclusions are drawn in Section 5.

2 DATA-DRIVEN PROGNOSTICS MODELS

This Section illustrates the basics of the two prognostics models used in this work for the prediction of the *RUL*. In particular, we consider the Homogeneous Discrete-Time Finite-State Semi-Markov Model (*HDTFSSMM*) proposed by some of the authors in (Al-Dahidi et al. 2015; Al-Dahidi et al. 2016b) (Subsection 2.1) and the Fuzzy Similarity-Based (*FSB*) model of literature proposed in (Di Maio et al. 2013) (Subsection 2.2).

Let us assume the availability of multi-dimensional P degradation trajectories regarding a fleet of P pieces of equipment working under different operating conditions. Each p -th trajectory is a z -dimensional trajectory of Z signals representative of the equipment behaviour and of the operating conditions the equipment is subjected to. The available data are divided into P_{train} training, P_{valid} validation and P_{test} testing sets for the purpose of building the individual models, developing the proposed ensemble approach and verifying its performance with respect to each individual model.

Among the training trajectories, P_{train}^c are assumed to be complete run-to-failure trajectories (i.e., trajectories that last all the way to the instance when the degradation state reaches the threshold value beyond which the equipment loses its functionality) and $P_{train}^{lc} = P_{train} - P_{train}^c$ are incomplete run-to-failure trajectories (i.e., trajectories that do not reach the failure threshold).

2.1 The Homogeneous Discrete-Time Finite-State Semi-Markov Model (*HDTFSSMM*)

The Homogeneous Discrete-Time Finite-State Semi-Markov Model (*HDTFSSMM*) entails three phases: 1) an offline identification of the degradation levels (health states) of an *HDTFSSMM* that are explained by the different operating conditions experienced by the equipment during its life; 2) the estimation of the model parameters; 3) the use of the inferred degradation model in a direct Monte Carlo (*MC*) simulation to estimate the *RUL* of a new equipment of the fleet (Cadini et al. 2009). For more details on the *HDTFSSMM* model, the interested reader may refer to (Al-Dahidi et al. 2016b).

This model has been selected because it is capable of benefiting from the availability of also incomplete run-to-failure degradation trajectories during the training phase for enhancing the *RUL* estimation, rather than relying solely on the complete run-to-failure trajectories, compared to other data-driven prognostics models, e.g., Fuzzy Similarity-Based (*FSB*) model (Di Maio et al. 2013).

2.2 The Fuzzy Similarity-Based (FSB) model

The idea underpinning this model is to evaluate the similarity between the test trajectory and the P_{train}^C available complete run-to-failure reference trajectories, and to use the *RULs* of these latter to estimate the *RUL* of the former, considering how similar they are (Di Maio et al. 2013). The similarity is quantified by resorting to the definition of an “approximately zero” fuzzy set taken as a bell-shaped function whose parameters can be set by following a trial and error procedure on a validation set of complete run-to-failure trajectories. For more details on the Fuzzy Similarity-Based (FSB) model, the interested reader may refer to (Di Maio et al. 2013).

3 THE SWITCHING ENSEMBLE APPROACH FOR DATA-DRIVEN PROGNOSTICS

The proposed approach is sketched in Figure 1 and entails, first an offline selection of the optimal switching time t_{opt} among all the possible switching times $t_{sw} = [t_{sw}^{min}, t_{sw}^{max}]$, where t_{sw}^{min} is the first available measurement time and t_{sw}^{max} is the longest end-of-life among the fleet, that minimizes the Accuracy Index (*AI*) over the P_{valid} validation trajectories, i.e., the relative error of the *RUL* prediction (Saxena et al. 2010). The *AI* evaluation metric can be expressed as (Saxena et al. 2010):

$$AI = \frac{\sum_{p_{valid}=1}^{P_{valid}} AI_{p_{valid}}}{P_{valid}} \quad (1)$$

where

$$AI_{p_{valid}} = \sum_{t=1}^{T_{p_{valid}}} \frac{|\widehat{RUL}_{p_{valid}}(t) - RUL_{p_{valid}}(t)|}{RUL_{p_{valid}}(t)} \quad (2)$$

and $T_{p_{valid}}$ is the failure time of each p_{valid} -th equipment.

Then, an online usage for predicting the *RUL* of P_{test} pieces of equipment. In other words, the optimal switching time t_{opt} represents the time up to which Model 1 is used for providing the *RUL* estimates at the early stage of the equipment life and beyond which Model 2 is used when the equipment approaches the end-of-life.

4 CASE STUDY

In this Section, the potentiality of the proposed ensemble approach is demonstrated on a real case study regarding a fleet of $P = 150$ aluminum electrolytic capacitors used in electric vehicles powertrain, adapted from (Al-Dahidi et al. 2016b; Rigamonti et al. 2016).

4.1 Description of dataset

The main degradation mechanism of electrolytic capacitors is the vaporization of the electrolyte, whose degradation speed is largely influenced by the component working temperature (Kulkarni et al. 2010).

According to (Rigamonti et al. 2016), the Normalized Equivalent Series Resistance (ESR^{norm}) is considered as a degradation indicator and its evolution has been simulated for a fleet of capacitors using a physics-based model (Rigamonti et al. 2016). The simulated degradation trajectories begin with an initial ESR^{norm} value equal to 100% and continue until when the ESR^{norm} exceeds the failure threshold of 200% (Venet et al. 1993). During the capacitor life, the following $Z=2$ signals are measured:

1. $ESR^{measured}$, the direct measurements of the degradation indicator, i.e., the measurements of the Normalized Equivalent Series Resistance (ESR^{norm});
2. the temperature experienced by the capacitor (T), which represents the operating condition most influencing the degradation process of the capacitor.

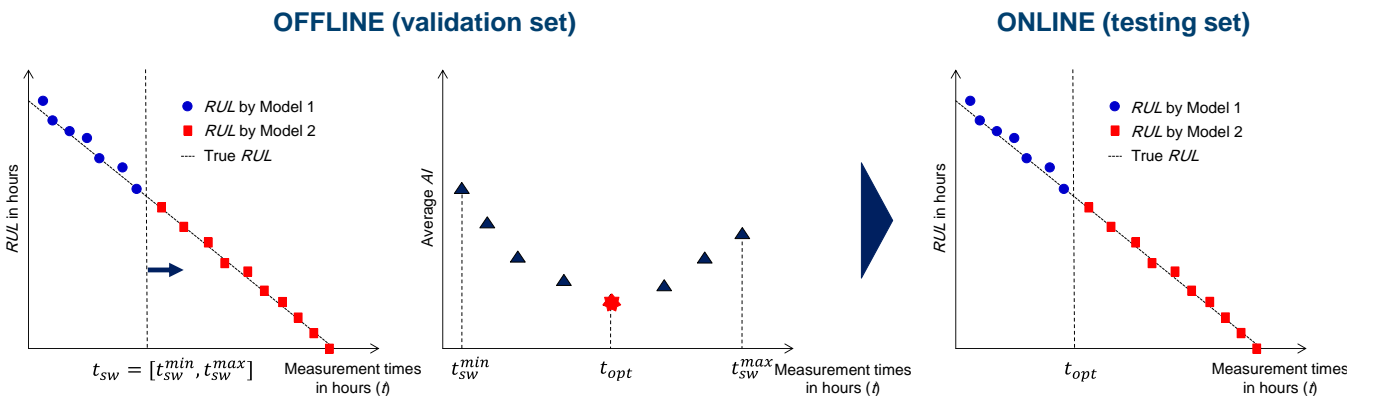


Figure 1. Sketch of the proposed ensemble approach.

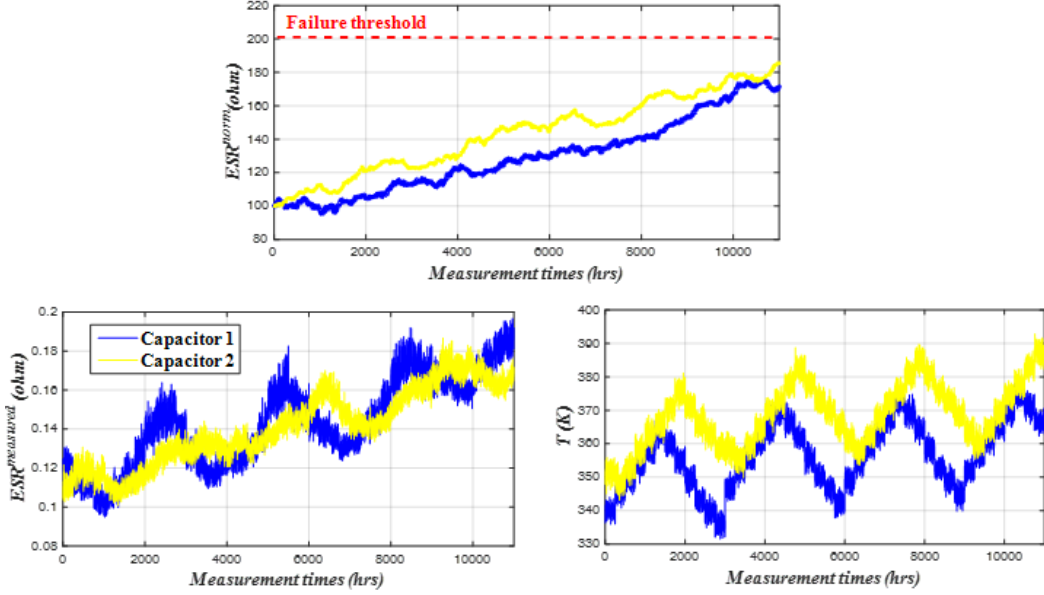


Figure 2. The true degradation process (ESR^{norm}) (top), the ESR measurements ($ESR^{measured}$) (left bottom) and the temperature profiles experienced by the capacitors (T) (right bottom).

According to design experts (Lahyani et al. 1998), temperature variations experienced by the capacitors during life are mainly caused by i) the seasonality of the environmental external temperature and by ii) the aging (barely up to 10% of its initial temperature value). Therefore, the simulated temperature profiles follow an arbitrary sinusoidal function that justifies seasonality, by adding to this a shift sigmoidal function accounting for aging.

The heterogeneity among the capacitors that belong to the fleet is guaranteed by considering arbitrary parameters of the sinusoidal and the sigmoidal functions.

For clarification purposes, Figure 2 shows the simulated data of two capacitors (capacitor 1 and capacitor 2 – dark and light shade of color, respectively): Figure 2 (top) shows the capacitors degradation process (ESR^{norm}), Figure 2 (left bottom) shows the ESR measurements ($ESR^{measured}$), whereas Figure 2 (right bottom) shows the temperature profiles experienced by the capacitors (T).

It is worth noticing that the higher the temperature (capacitor 2 – light shade of color), the faster the vaporization process due to the increase of the self-heating effects and, hence, the faster the failure process, as shown in Figure 2 (top) (Rigamonti et al. 2016; Wolfgang 2007).

The $P=150$ trajectories generated under different temperature conditions are divided into $P_{train} = 100$ training, $P_{valid} = 25$ validation and $P_{test} = 25$ testing subsets. Among the $P_{train} = 100$ capacitors, $P_{train}^c = 20$ are assumed to show degradation trajectories that last all the way to the failure threshold, whereas $P_{train}^{ic} = 80$ show incomplete run-to failure trajectories.

For clarification purposes, Figure 3 shows examples of the ESR^{norm} signals of the complete and incomplete run-to-failure degradation trajectories (in dark and light shade of color, respectively). The objective is to accurately predict the RUL in the testing data set.

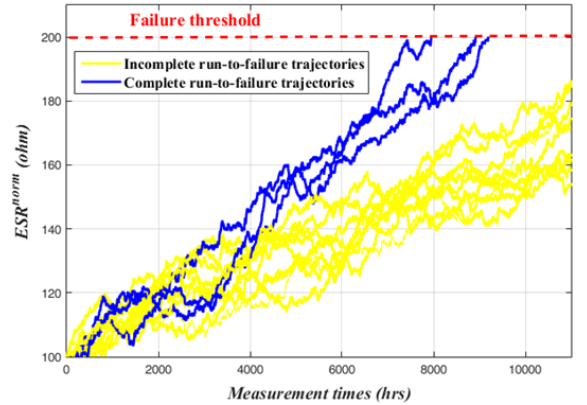


Figure 3. Examples of simulated complete and incomplete run-to-failure degradation trajectories.

4.2 Implementation of the Ensemble Approach

To implement the ensemble approach, firstly the individual models are built by using the $P_{train} = 100$ capacitors.

Concerning the $HDTFSSMM$, the whole train trajectories are used to build the degradation model, estimate its parameters, and then, estimate the RUL of the $P_{test} = 25$ capacitors within a direct MC simulation scheme with $N_{max} = 1000$ MC trials (Al-Dahidi et al. 2016b).

Concerning the FSB model, only the $P_{train}^c = 20$ complete run-to-failure training trajectories can be used to build a reference library for estimating the RUL of the $P_{test} = 25$ capacitors.

4.3 Results of Ensemble Approach

We rely on the $P_{valid} = 25$ trajectories to identify the optimum switching time t_{opt} among possible switching times $t_{sw} = [1000, 12000]$ hours, where the lower bound (1000) is the first available measurement time, whereas the upper bound (12000) is the longest end-of-life among the $P_{valid} = 25$ capacitors. The optimum switching time is the value at which the AI is minimized, and is therefore, set equal to $t_{opt} = 9000$ hours (star in Figure 4).

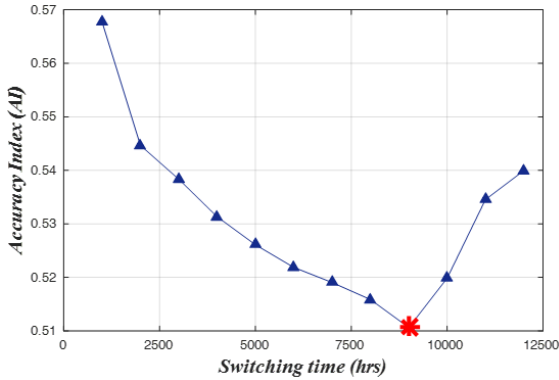


Figure 4. AI values calculated on the $P_{valid}=25$ validation trajectories vs. switching times of the proposed ensemble approach.

Table 1 reports the AI over the $P_{test} = 25$ test trajectories of the proposed ensemble approach, compared to each individual model. It can be observed that the ensemble approach outperforms any of the individual model in terms of the AI with a 20.31% improvement over the best individual member model, FSB , whose AI is 0.64.

Table 1. AI values on the $P_{test}=25$ test trajectories obtained by the proposed ensemble approach and the individual models.

	<i>HDTFSSMM</i>	<i>FSB</i>	Switching approach
AI	1.24	0.64	0.51

The estimates of the RUL obtained by the proposed ensemble approach for two capacitors are plotted in Figure 5, in solid line, before and after the optimal switching time (together with those obtained by individual *HDTFSSMM* and *FSB* models in circles and squares markers, respectively).

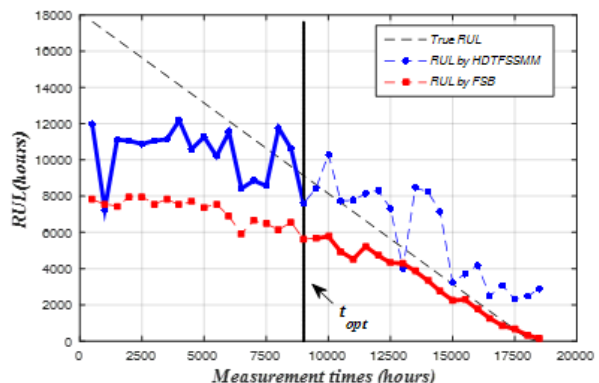
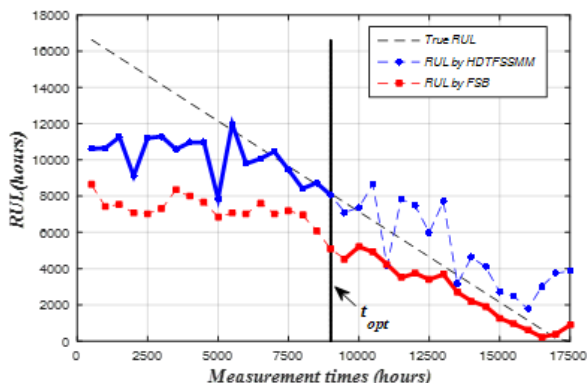


Figure 5. Comparison of the RUL predictions for two capacitors provided by the proposed ensemble approach and each individual *HDTFSSMM* and *FSB* models.

It can be easily noticed that the predictions provided by the two independent models are comparable in terms of overall accuracy: while the *HDTFSSMM* tends to provide more accurate RUL predictions with respect to the *FSB* model at the early stage of its life, the *FSB* model tends to provide more accurate RUL predictions when approaching the end-of life. This has been justified by the fact that the complete run-to-failure trajectories used for training the *FSB* model are characterized by short lives (for the sake of clarity refer to Figure 3) and, thus, the *FSB* model tends, on average, to underestimate the component RUL at the beginning of its degradation trajectory. When the component is close to failure, the *FSB* model overcomes that of the *HDTFSSMM* since the incomplete run-to-failure trajectories are not informative on the component behavior just before the failure. In particular, the *FSB* model becomes more satisfactory than that of the *HDTFSSMM* (Al-Dahidi et al. 2016b).

Finally, it can be easily noticed that proposed ensemble approach outperforms each individual model in terms of the accuracy in RUL predictions throughout the entire lives of the capacitors, as expected.

5 CONCLUSIONS

In this work, a switching ensemble approach for data-driven prognostics is proposed for accurately estimating the Remaining Useful Life (RUL) of an electrolytic capacitor based on data coming from a fleet of P capacitors working under different operating conditions. Two data-driven prognostics models are considered, an Homogeneous Discrete-Time Finite-State Semi-Markov Model (*HDTFSSMM*) and a Fuzzy Similarity-Based (*FSB*) model.

The ensemble procedure is based on a switching strategy that selects the most accurate model at each stage of the equipment life. More specifically, the approach selects the optimal switching time t_{opt} among possible switching times, based on the average value of the Accuracy Index (AI) that can be obtained on a validation set.

Once the optimal switching time is identified, the switching ensemble approach is used online for predicting the *RUL* of a new equipment of the fleet.

The proposed approach is able to provide more accurate *RUL* predictions throughout the entire life of the equipment compared to each individual *HDTFSSMM* and *FSB* models.

ACKNOWLEDGEMENTS

The participation of Sameer Al-Dahidi and Piero Baraldi has been supported by the European Union Project INNOVATION through Human Factors in risk analysis and management (INNHF, www.innhf.eu) funded by the 7th framework program FP7-PEOPLE-2011-Initial Training Network: Marie-Curie Action.

REFERENCES

- (Al-Dahidi et al. 2015) Al-Dahidi, S., Di Maio, F., Baraldi, P., & Zio, E., (2015). Supporting Maintenance Decision with Empirical Models Based on Fleet-Wide Data. *in The 49th ESReDA Seminar on: Innovation through Human Factors in Risk Assessment & Maintenance, Belgium*, 2015, 1–12.
- (Al-Dahidi et al. 2016a) Al-Dahidi, S., Di Maio, F., Baraldi, P., & Zio, E., (2016a). Remaining Useful Life Estimation in Heterogeneous Fleets Working under Variable Operating Conditions, *Reliab. Eng. Syst. Saf.*, under review.
- (Al-Dahidi et al. 2016b) Al-Dahidi, S., Di Maio, F., Baraldi, P., Zio, E., & Seraoui, R., (2016b). A Framework for Reconciling the Consensus Clusterings of a Fleet of Nuclear Power Plants Turbines for Fault Diagnosis. *Appl. Soft Comput.*, Under Review.
- (Baraldi et al. 2013) Baraldi, P., Mangili, F., & Zio, E. (2013). Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data. *Reliability Engineering & System Safety*, 112, 94-108.
- (Baraldi et al. 2015) Baraldi, P., Mangili, F., & Zio, E. (2015). A belief function theory based approach to combining different representation of uncertainty in prognostics. *Information Sciences*, 303, 134-149.
- (Cadini et al. 2009) Cadini, F., Zio, E., & Avram, D. (2009). Monte Carlo-based filtering for fatigue crack growth estimation. *Probabilistic Engineering Mechanics*, 24(3), 367-373.
- (Chiang et al. 2001) Chiang, L. H., Braatz, R. D., & Russell, E. L. (2001). Fault detection and diagnosis in industrial systems. *Springer Science & Business Media*.
- (Di Maio et al. 2013) Di Maio, F., & Zio, E. (2013). Failure prognostics by a data-driven similarity-based approach. *International Journal of Reliability, Quality and Safety Engineering*, 20(01), 1350001.
- (Heng et al. 2009) Heng, A., Zhang, S., Tan, A. C., & Mathew, J. (2009). Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical Systems and Signal Processing*, 23(3), 724-739.
- (Hu et al. 2010) Hu, C., Youn, B. D., & Wang, P. (2010). Ensemble of data-driven prognostic algorithms with weight optimization and k-fold cross validation. *In ASME 2010 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 1023-1032). American Society of Mechanical Engineers.
- (Jardine et al. 2006) 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(7), 1483-1510.
- (Kulkarni et al. 2010) Kulkarni, C., Biswas, G., Koutsoukos, X., Celaya, J., & Goebel, K. (2010). Integrated diagnostic/prognostic experimental setup for capacitor degradation and health monitoring. *In AUTOTESTCON*, 1-7). IEEE.
- (Lahyani et al. 1998) Lahyani, A., Venet, P., Grellet, G., & Viverge, P.-J., (1998). Failure prediction of electrolytic capacitors during operation of a switchmode power supply. *Power Electronics, IEEE Transactions on*, 13(6), 1199–1207.
- (McCall 1965) McCall, J. J. (1965). Maintenance policies for stochastically failing equipment: a survey. *Management science*, 11(5), 493-524.
- (Medina-Oliva et al. 2014) Medina-Oliva, G., Voisin, A., Monnin, M., & Leger, J. B. (2014). Predictive diagnosis based on a fleet-wide ontology approach. *Knowledge-Based Systems*, 68, 40-57.
- (Polikar 2006) Polikar, R. (2006). Ensemble based systems in decision making. *Circuits and systems magazine, IEEE*, 6(3), 21-45.
- (Rigamonti et al. 2016) Rigamonti, M., Baraldi, P., Zio, E., Astigarraga, D., & Galarza, A. (2016). Particle Filter-Based Prognostics for an Electrolytic Capacitor Working in Variable Operating Conditions. *Power Electronics, IEEE Transactions on*, 31(2), 1567-1575.
- (Sankavaram et al. 2009) Sankavaram, C., Pattipati, B., Kodali, A., Pattipati, K., Azam, M., Kumar, S., & Pecht, M. (2009). Model-based and data-driven prognosis of automotive and electronic systems. *In Automation Science and Engineering, 2009. CASE 2009. IEEE International Conference on*, 96-101. IEEE.
- (Saxena et al. 2010) Saxena, A., Celaya, J., Saha, B., Saha, S., & Goebel, K. (2010). Metrics for offline evaluation of prognostic performance. *International Journal of Prognostics and Health Management*, 1(1), 4-23.
- (Tobon-Mejia et al. 2011) Tobon-Mejia, D. A., Medjaher, K., Zerhouni, N., & Tripot, G. (2011). Hidden Markov models for failure diagnostic and prognostic. *In Prognostics and System Health Management Conference (PHM-Shenzhen)*, 1-8. IEEE.
- (Van Tung & Yang 2009) Van Tung, T., & Yang, B. S. (2009). Machine fault diagnosis and prognosis: The state of the art. *International Journal of Fluid Machinery and Systems*, 2(1), 61-71.
- (Venet et al. 1993) Venet, P., Darnand, H., & Grellet, G., (1993). Detection of faults of filter capacitors in a converter. Application to predictive maintenance. *Telecommunications Energy Conference, INTELEC '93. 15th International*, 2, 229–234.
- (Wei et al. 2013) Wei, Z., Tao, T., ZhuoShu, D., & Zio, E. (2013). A dynamic particle filter-support vector regression method for reliability prediction. *Reliability Engineering & System Safety*, 119, 109-116.
- (Wolfgang 2007) Wolfgang, E., (2007). Examples for failures in power electronics systems. *ECPE Tutorial on Reliability of Power Electronic Systems*, Nuremberg, Germany.
- (Zhao et al. 2010) Zhao, Z., Wang, F. L., Jia, M. X., & Wang, S. (2010). Predictive maintenance policy based on process data. *Chemometrics and Intelligent Laboratory Systems*, 103(2), 137-143.
- (Zio 2012) Zio, E. (2012). Prognostics and health management of industrial equipment. *Diagnostics and Prognostics of Engineering Systems: Methods and Techniques*, 333-356.