

Residual-based failure prognostic in dynamic systems

K. Medjaher, N. Zerhouni

*FEMTO-ST Institute, UMR CNRS 6174 - UFC / ENSMM / UTBM
Automatic Control and Micro-Mechatronic Systems Department
24, rue Alain Savary, 25000 Besançon, France.
(e-mail: kamal.medjaher@ens2m.fr)*

Abstract: this paper deals with failure prognostic in dynamic systems. The system's remaining useful life is estimated based on residual signals. This supposes the possibility to build a dynamic model of the system by using the bond graph tool, and the existence of a degradation model in order to predict its future health state. The choice of bond graph is motivated by the fact that it is well suited for modeling physical systems where several types of energies are involved. In addition, it allows to generate residuals for fault diagnostic and prognostic. The proposed method is then applied on a simple dynamic model of a hydraulic system to show its feasibility.

Keywords: Prognostic, Degradation, Failure, Residual, Remaining Useful Life, Monitoring.

1. INTRODUCTION

Nowadays, industrial systems are more and more complex due, in part, to their growing size and to the integration of new technologies. With ageing, these systems become more vulnerable to failures and their maintenance difficult and expensive. This situation combined with requirements of productivity, profit growth, operational availability and safety pushes industrials and researchers to look for innovative tools and methods allowing them to satisfy these requirements and reach their objectives. To do this, one of the possible levers consists in maintenance activity. By maintaining the system, one can reduce its global life cycle costs, increase its availability, improve the safety of operators and reduce the environmental incidents. Maintenance tasks can be curative or preventive. In curative maintenance framework, the components are replaced only when they are not able to fulfill the task for which they are designed. The main drawback of this solution is that the machine undergoes the fault, which is sometimes simply a non desired situation (explosion, chemical and poisoning materials, etc.). To overcome this, it is possible to monitor some significant parameters of the system and then, by setting some threshold values, one can proceed to component changes when the monitored parameters exceed their corresponding defined thresholds. This is what can be done in the framework of condition based maintenance (CBM) (Jardine et al. [2006]). But, this is still not sufficient, because it happens that, at the time of fault occurrence, the spare parts are not available or not sufficient, or simply that the needed resources (maintainers) are busy. A "best" maintenance could be then a proactive one which can be achieved in the prognostic framework (Vachtsevanos et al. [2006]). It means that, one first tries to predict the health state of the system, and then plans appropriate actions according to what simulations return.

Contrary to fault diagnostic which is relatively mature (Isermann [2005]), well developed and spread within re-

search and industrial communities, failure prognostic is a new research activity with few publications and applications. However, some research works (Muller et al. [2008], Vachtsevanos et al. [2006], Kacprzyński et al. [2004], W.Q. Wang et al. [2004], Provan [2003], Byington et al. [2002], Chelidze et al. [2002]), have been achieved in the literature and can be grouped in three main approaches, namely: model-based prognostic, data-driven prognostic and experience-based prognostic. In the present contribution, model-based prognostic in dynamic systems is addressed. The mathematical model is derived from the physical knowledge of the system by using the bond graph tool. This graphical formalism allows to generate residual signals, in which degradation model can be injected to predict the Remaining Useful Life (RUL) of the system. The paper is organized as follows: section 2 gives a brief presentation of the main prognostic approaches, and section 3 details the proposed prognostic method. Section 4 is dedicated to the application of the method on a small hydraulic system, and finally, a conclusion with some future works is given at section 5.

2. FAILURE PROGNOSTIC

2.1 Definitions and terminologies

The term prognostic finds its origin in the Greek word "*progignōskein*" which means "to know in advance". Prognostic is well used in medical domain, where doctors try to make predictions about the health of a patient by taking into account the actual diagnosis of a disease and its evolution compared with other similar observed cases. This reasoning can be transposed into the industrial domain at a condition to replace the patient by a machine, an industrial system or a component.

Many definitions have been given in the literature about industrial prognostic (see Muller [2005], W.Q. Wang et al. [2004], Byington et al. [2002], Lebold and Thurston [2001]

for more details). Three main points are highlighted, namely: the system's actual state, the projection (or extrapolation) of this latter, and the estimation of the remaining time before failure. These definitions are then normalized by that one given by the standard (ISO, 13381-1 [2004]) in which prognostic is defined as the estimation of the operating time before failure and the risk of future existence or appearance of one or several failure modes. This standard defines the outlines of prognostic, identifies the data needed to perform prognostic and sets the alarm thresholds and the limits of system's reset (total shut-down). The main steps defined in this standard are summarized in Fig. 1.



Fig. 1. Summary of the ISO 13381-1: 2004 standard main steps

The first step consists in monitoring the system by a set of sensors or inspections achieved by operators. The monitored data are then pre-processed in order to be used by the Diagnostic module. The output of this module identifies the actual operating mode. This state is then projected in the future, by using adequate tools, in order to predict the system's future state. The intersection point between the value of each projected parameter or feature and its corresponding alarm threshold leads to what is known as RUL (Remaining Useful Life) of the system (Fig. 2). Finally, appropriate maintenance actions can be taken depending on the estimated RUL. These actions may aim at eliminating the origin of a failure which can lead the system to evolve to any critical failure mode, delaying the instant of a failure by some maintenance actions or simply stopping the system if this is judged necessary.

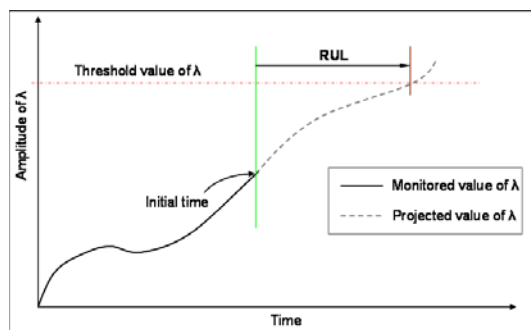


Fig. 2. Estimation of the value of the RUL

As in any prediction work, a prediction error should be associated to the estimated value of the RUL (Fig. 3). The sources of the prediction error may be multiple: modeling hypotheses, non-significant data, used prediction tools, uncertainty in the thresholds' values, etc.. In addition, uncertainty is intrinsic to any prognostic work as mentioned by Provan [2003]: *"uncertainty is central to any definition of prognosis. This is because a prognosis involves a projection into the future, and we argue that all such future projections must contain some uncertainty, since the future can not be predicted with certainty"*.

The error associated to any RUL estimation should decrease as the time of the real failure approaches. This is

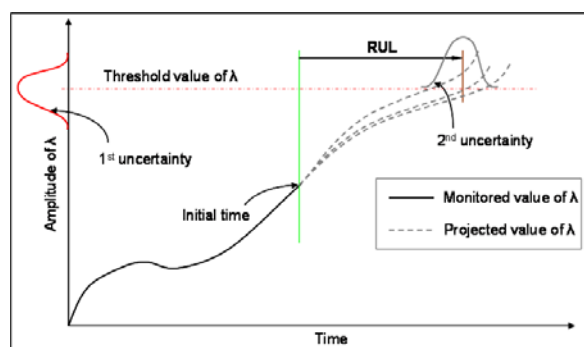


Fig. 3. Uncertainty associated to RUL estimation

exactly what happens in the case of weather forecast: the predictions given at the beginning of a week for the next Sunday, for example, are less precise than those given for the same day (next Sunday) but at one or two days before. This is because the predictions are adjusted each time new data are acquired.

Similarly to weather forecast, a confidence degree should be associated with any industrial prognostic work to render its conclusions more credible. Indeed, instead of telling an industrial that his/her machine will fail in x unites of time, it would be more realistic to give an estimated RUL with a confidence value. By including the uncertainty and confidence degree, the prognostic steps shown in Fig. 1 become more detailed as shown in Fig. 4.

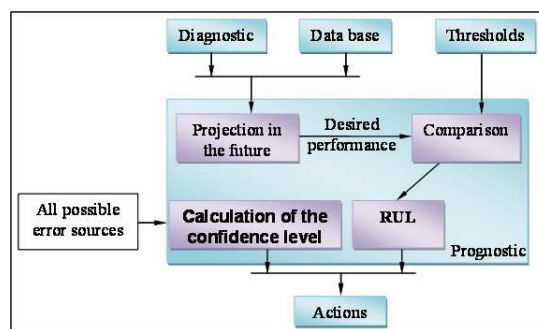


Fig. 4. Prognostic and confidence degree

As mentioned previously, the value of the estimated RUL is the output of some comparison between the projected state of the system and the predetermined threshold values (these values can be determined by using learning algorithms like those of neuro-fuzzy systems (Chinnam and Pundarikaksha [2004])). Note that, at the projection step, what is needed is not necessarily a value of a physical parameter but can be a desired performance, an achieved function or an availability of a service, depending on the kind of system on which prognostic is performed.

After having given some definitions and terminologies used in the prognostic framework, the following section deals with the existing approaches, methods and techniques allowing to quantify the indicators previously introduced.

2.2 Prognostic main approaches

In the literature, there exists three main prognostic approaches summarized in Fig. 5 (Vachtsevanos et al. [2006], Lebold and Thurston [2001]). A survey (but non exhaustive) of the methods used in each approach can be found in

(Jardine et al. [2006]) and in the second chapter of Muller's thesis (Muller [2005]).

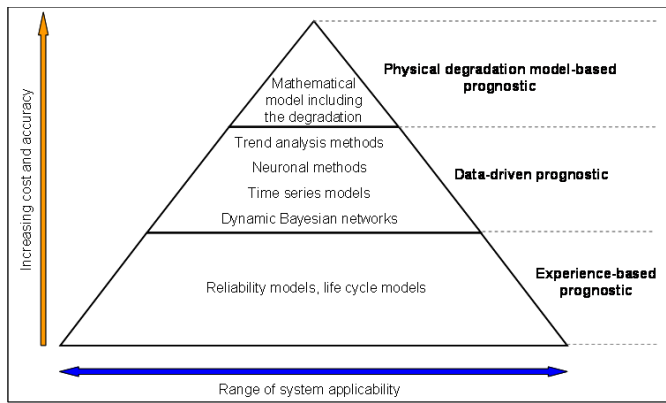


Fig. 5. Prognostic main approaches

Experience-based prognostic: it consists in using probabilist or stochastic models of the degradation phenomenon, or of the life cycle of the components, by taking into account the data and knowledge accumulated by experience during the whole exploitation period of the industrial system. The probabilist model can be a simple probability function or a modeling in the form of stochastic process. In this framework, the most used probability functions are: Weibull law, exponential law when the failure rate is supposed to be constant, and normal, log-normal and Poisson laws. The parameters of each law are estimated from the data gathered during the whole exploitation period of time (experience feedback, maintenance data, etc.). Stochastic process models can be Markovian or semi-Markovian.

The advantage of the methods of this approach is that it is not necessary to have complex mathematical models to do prognostic. Moreover, this approach is easy to apply on systems for which significant data are stored in a same standard that facilitates their use. For example, a company which has conserved during a long period of time a production and maintenance database with some minor rules and standards for data storing, can easily get the estimation of the parameters of the probability laws.

However, the main drawback of this approach dwells in the amount of data needed to estimate the parameters of the used laws. Indeed, huge and significant amount of exploitation data are needed in order to determine parameters that model faithfully the degradation phenomenon or the life cycle of the concerned system. Consequently, this approach can not be applied in the case of new systems for which data from experience feedback do not exist. The other kind of problem is that in most of cases, it is necessary to filter and pre-process the data to extract the useful ones, because the stored data are not always directly exploitable (for example, in the same company, two maintenance operators may enter two different informations or appreciations for the same resolved problem).

Data-driven prognostic: the principle of this approach consists in collecting information and data from the system and projecting them in order to predict the future evolution of some parameters, descriptors or features, and thus, predict the possible probable faults. Without being exhaustive, mathematical tools used in this approach

are mainly those used by the artificial intelligence community, namely: temporal prediction series, trend analysis techniques, neuronal networks under all their facets, neuro-fuzzy systems, hidden Markov models and dynamic bayesian networks.

The advantage of this approach is that, for a well monitored system, it is possible to predict the future evolution of a degradation without any need of prior mathematical model of the degradation.

However, the results obtained by this approach suffer from precision, and are sometimes considered as local ones (for the case of neural networks and neuro-fuzzy methods). In addition, the monitoring system must be well designed to insure acceptable prognostic results.

Model-based prognostic: this consists in studying each component or sub-system in order to establish for each one of them a mathematical model of the degradation phenomenon. The derived model is then used to predict the future evolution of the degradation (Luo et al. [2003], Chelidze et al. [2002]). In this case, the prognostic consists in evolving the degradation model till a determined future instant from the actual deterioration state and by considering the future use conditions of the corresponding component. Three main steps are needed in the framework of model-based prognostic. The first step is related to the construction of an analytical dynamic model including the degradation mechanism or phenomenon, and to the determination of failure thresholds. Follows, in the second step, a setup of a monitoring/diagnostic system which allows to evaluate the actual value of the degradation. Finally, a development or a selection of an adequate technique to solve the derived dynamic model (prediction step) is necessary.

The main advantage of this approach dwells in the precision of the obtained results, as the predictions are achieved based on a mathematical model of the degradation.

However, the derived degradation model is specific to a particular kind of component or material, and thus, can not be generalized to all the system components. In addition to that, getting a mathematical model of a degradation is not an easy task and needs well instrumented test-benches which can be expensive.

3. RESIDUAL-BASED PROGNOSTIC

The present contribution can be classified within the model-based approach. It aims at performing failure prognostic by generalizing the residual technique which is still used in Fault Detection and Isolation (FDI) framework.

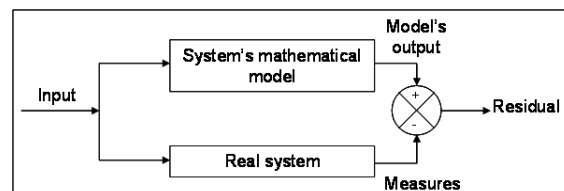


Fig. 6. The principle of residuals

Residuals are signals which are used to verify the coherence between the nominal and the actual operating modes of a system as shown in Fig. 6. When the system operates under its nominal operating mode, the residuals' values

should be theoretically equal to zero, and increases (or decreases) as the system leaves its nominal mode, which may be a consequence of fault occurrence. Note that, even if the concept of the residual introduced in this paper can be applied in both fault detection and isolation approaches, namely quantitative and qualitative, in the following work only the quantitative application is considered. More particularly, in this contribution, it is considered that the residual signals are generated from a dynamic model derived from the physical system by using a unified multi-disciplinary tool: the bond graph (Samantaray and Bouamama [2008], Karnopp et al. [2006], Dauphin-Tanguy [2001]).

The bond graph tool is a graphical representation of power transfer within a physical system. A bond graph model is situated between the physical model and the mathematical model. It is used in modeling to derive mathematical models in forms of state space and transfer function, in structural analysis of the system's properties like controllability, observability, model reduction, actuator and sensor placement, and in FDI.

The proposed prognostic method is based on residual signals derived from a bond graph model of the system. It is supposed to be applied on dynamic systems for which there exists a mathematical model of the degradation. The main steps of this method are summarized in Fig. 7.

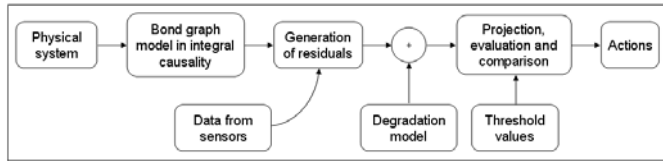


Fig. 7. The main steps of the proposed method

The bond graph model is constructed from the physical knowledge of the dynamic system. In the present work, this model is put in preferred integral causality form in order to deduce the mathematical model.

The mathematical model of the degradation is chosen according to the type of degradation mechanism affecting the physical system. Degradation phenomena can be classified into three main categories (Muller [2005]): linear, concave and convex profiles, as shown in Fig. 8.

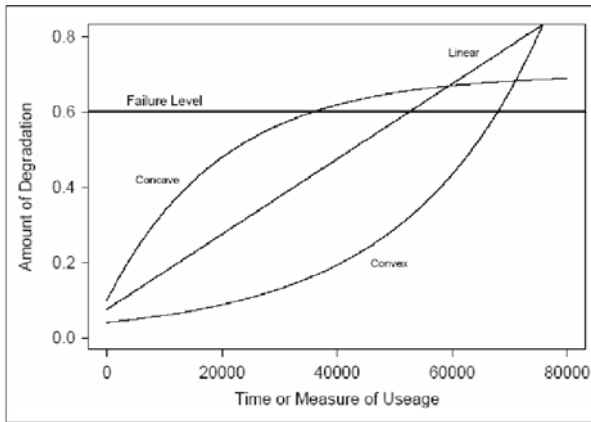


Fig. 8. The main degradation models

Pneumatic wear and hydraulic resistance variation, for example, can be modeled by linear degradation models,

where the analytical expression of the degradation is given by the following linear equation:

$$D(t) = D(0) + Ct, \quad (1)$$

where, D is the wear rate and C is the material parameter. In the category of convex models, Paris-Erdogan (Luo et al. [2003]) law, given by (2), is the most used model to represent degradation mechanisms like crack growth by fatigue for example.

$$\frac{d\theta}{dn} = C(\Delta K)^\gamma. \quad (2)$$

In equation (2), n stands for the number of constraint's cycles, θ is the crack length, ΔK is the stress intensity, and C and γ are material parameters.

For concave models, one can cite corrosion and erosion mechanisms in printed circuits which may lead to short-circuits. These degradation phenomena can be modeled by the following two differential equations:

$$\frac{dA_1}{dt} = -k_1 A_1, \quad (3)$$

$$\frac{dA_2}{dt} = k_1 A_2, \quad (4)$$

where, A_1 and A_2 are quantity of chlorine (Cl) and chlorine of copper ($ClCu$) respectively, and k_1 stands for the transformation rate of chlorine to chlorine of copper ($ClCu$).

The choice of a degradation model depends on the kind of phenomenon one wants to take into account in the real system. The considered degradation should then be represented in the bond graph model, and consequently, in the generated residuals. Indeed, any degradation mechanism can be characterized by a change in one or more parameters of the system, and this can be directly modeled by one of the bond graph passive elements: R , C and I . For example, deposit of sediments in a hydraulic pipe can be modeled as a variation of a resistance element, fluid leakage from a reservoir as a change of its cross section (and consequently, of its hydraulic capacity), etc..

After having constructed the bond graph model, residual equations can be generated. The obtained residuals are numerical evaluation of some coherence relations called Analytical Redundancy Relations (ARRs), and thus have a physical meaning as they include the parameters of the real system. For reminder, in FDI domain the residuals are generated from a bond graph model in derivative causality. This is motivated by the fact that in FDI framework, integral causality has to be avoided since the time origin of the failure is considered as an unknown parameter. This can also be explained by the fact that in FDI domain, one observes the effects of the failure and tries to go back to the cause of the abnormal observed situation, and this corresponds exactly to derivative causality on the bond graph model. However, in the case of fault prognostic, the residuals may be derived from a bond graph model in preferred integral causality. Indeed, in this case, the problem of initial conditions encountered in FDI is simply solved since we are interested in failure prediction from an initial state given by the monitoring system. Thus, in

the integral form, the initial conditions become a known parameters (note that in the integral causality, the cause is a known information which is used to calculate the unknown information: the effect).

The last step in the proposed residual-based failure prognostic deals with the prediction, evaluation and comparison. This aims at projecting in the future the value of the residuals, evaluating their values and comparing these latter to some pre-defined thresholds. The time difference between the initial time of failure occurrence (given by a diagnostic module) and the time corresponding to the intersection point between each residual projection and its corresponding threshold value gives the estimated value of the RUL. Note that in addition to the uncertainty on the threshold values, uncertainty related to the projection model may be considered.

4. APPLICATION EXAMPLE

The proposed methodology presented in the previous section is applied on a small hydraulic system shown in Fig. 9. The tank is supposed to be a cylindric reservoir with a cross-section A . The fluid level in the tank can be controlled by acting either on the flow of the pump or on opening or closing the valve. But, for simplicity, only the open loop structure is considered (the control part is not shown and not taken into account in this example). Two sensors are implemented on the system: the pressure sensor (P) at the bottom of the tank and the flow sensor (F) to measure the rate flow across the valve. The nomenclature of the parameters and variables used in this example is given at table 1.

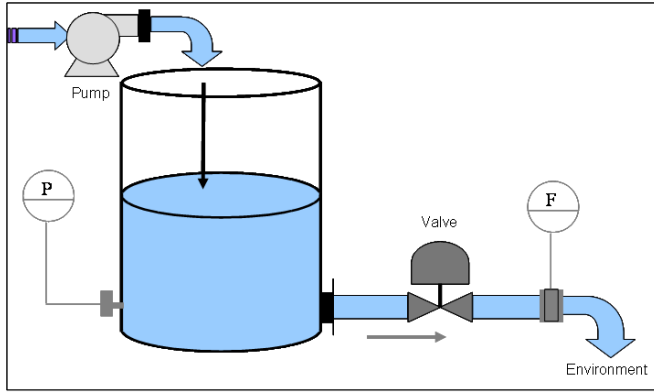


Fig. 9. The hydraulic system

The bond graph model corresponding to this system is given in Fig. 10. This model is built based on the hypothesis that the fluid is incompressible. In this graphical model, the pump is supposed to be a constant flow source and the valve as a non-linear resistance. The bond graph passive elements C and R are used to model the volume conservation in the tank and the flow rate across the valve, respectively.

Several types of physical degradations can be considered: a deposit of sediments in the pipes and the valve, incipient leakage from the tank, cavitation phenomenon, etc.. All these degradations can be modeled as changes or variations of the bond graph elements C , I and R .

In the present application, an incipient (and linear) degradation is simulated. Particularly, a linear increase of the

Table 1. Nomenclature of the used variables and parameters

Symbol	Designation	Value
A	Cross section of the tank (m^2)	1
ρ	Water density (kg/m^3)	1000
g	Gravity constant (m/s^2)	9.81
C_d	Discharge coefficient of the valve ($m^4/s.N^{1/2}$)	$(1/3).10^{-5}$
α	Slope of the linear degradation ($m^4/s^2.N^{1/2}$)	$-(1/36).10^{-7}$
P_{atm}	Atmospheric pressure (N/m^2)	10^5
Q_{in}	Fluid flow at the input of the tank (m^3/s)	$(1/3).10^{-3}$
P	Pressure at the bottom of the tank (N/m^2)	-
F	Fluid flow across the valve (m^3/s)	-

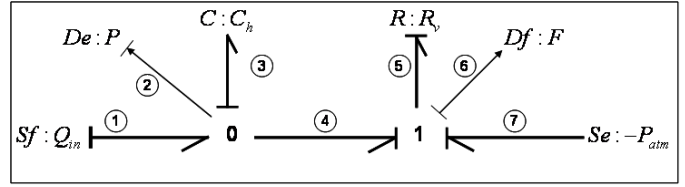


Fig. 10. The bond graph model in integral causality

valve's resistance (which depends on the discharge coefficient C_d) due to deposit of sediments inside it is taken into account. The mathematical model of this physical degradation is given by the following equation:

$$C_d(t) = \begin{cases} C_d & \text{if } t \leq t_0 \\ C_d + \alpha(t - t_0) & \text{if } t > t_0 \end{cases}, \quad (5)$$

where t_0 stands for the initial time at which the degradation begins to rise (which is given by a diagnostic module), and α represents the slope of the linear degradation.

As the system is fully observable and does not present any algebraic or differential loop, the number of residuals that can be generated is equal to two:

$$r_1 = \frac{A}{\rho \cdot g} \frac{dP}{dt} - Q_{in} + C_d \cdot \text{sign}(P - P_{atm}) \sqrt{|P - P_{atm}|}$$

$$r_2 = \frac{F^2}{C_d^2} - P + P_{atm} \quad (6)$$

The last step of the proposed method consists in projection, evaluation and comparison, which allows to estimate the value of the RUL. Indeed, if the threshold values are known, the value of the RUL can be directly deduced as the time difference between an initial time of failure occurrence (which is given by a diagnostic module) and a final time corresponding to the intersection point between the value of the threshold and the value of the corresponding projected residual (Fig. 11). In the given simulations, the threshold values are determined according to the quantity of the fluid flow across the valve, which decreases till a predefined threshold limit (equal to $0.15 \times 10^{-4} m^3/s$) due to the increase of the valve's resistance. Thus, the estimated value of the RUL is about 1100s. Note that the residuals of Fig. 11 might be scaled (or normalized) to

have a same order of magnitude. Indeed, the residual r_1 represents the flow conservation with an order of 10^{-4} and the residual r_2 represents the pressure conservation with an order of 10^3 .

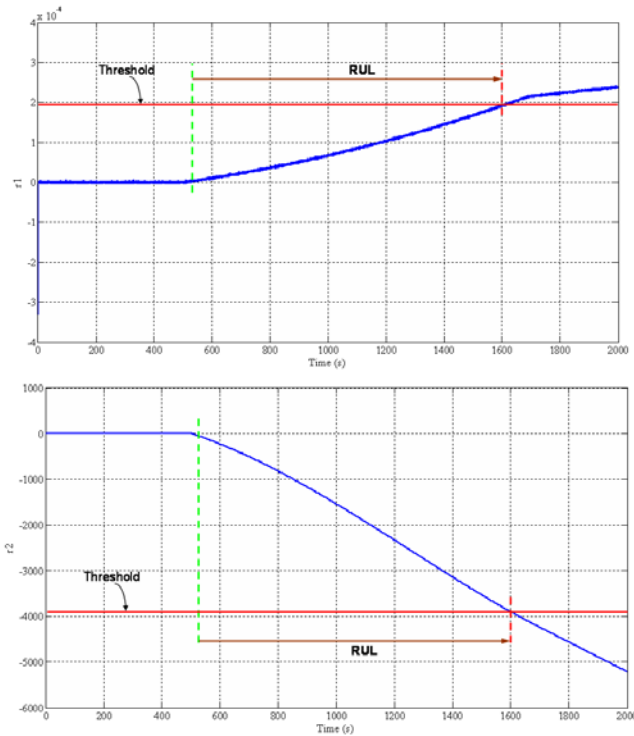


Fig. 11. Time variation of the residuals r_1 and r_2

5. CONCLUSION AND FUTURE WORK

A contribution about failure prognostic has been presented in this paper. The proposed method is based on the use of residuals to estimate the RUL. This method is applicable on systems for which an analytical dynamic model can be derived and where mathematical models of degradation mechanisms are available. The choice of the bond graph is justified by the fact that this graphical tool can be used to easily model and generate analytical redundancy relations and corresponding residuals for multi-physical dynamic systems where different kinds of energies are involved. The proposed method is then applied on a small hydraulic system and significant simulation results are obtained. However, obtaining a mathematical model of a degradation phenomenon is not a trivial task and this needs consequent means. The determination of the threshold values needed in the calculation of the RUL is another point which deserves to be developed. Indeed, in the present contribution, the thresholds are supposed as known parameters, which in practice is not the case. Statistical or neuro-fuzzy methods can then be employed in order to set significant and persistent threshold values that minimize the number of false alarms and avoid the non-detection situations. Finally, uncertainties and confidence limits are not taken into account, in the given simulations, to estimate the RUL. The problem can be handled by introducing uncertainties in the system's parameters and in the degradation model so that confidence values can be calculated.

REFERENCES

- C. Byington, M. Roemer, G. Kacprzynski, and T. Galie. Prognostic enhancements to diagnostic systems for improved condition-based maintenance. In *2002 IEEE Aerospace Conf., Big Sky, USA*, 2002.
- D. Chelidze, J.P. Cusumano, and A. Charterjee. A dynamical systems approach to damage evolution tracking, part 1: The experimental method. *Jour. of Vib. and Acoustics*, 124:250–257, 2002.
- R. Chinnam and B. Pundarikaksha. A neurofuzzy approach for estimating mean residual life in condition-based maintenance systems. *Int. J. Materials and Product Technology*, 20:166–179, 2004.
- G. Dauphin-Tanguy. *Les Bond Graphs*. Hermes, 2001.
- R. Isermann. Model-based fault-detection and diagnosis - status and applications. *Annual Reviews in Control*, 29: 71–85, 2005.
- A.K.S. Jardine, D. Lin, and D. Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. and Sign. Proc.*, 20:1483–1510, 2006.
- G.J. Kacprzynski, A. Sarlashkar, M.J. Roemer, A. Hess, and W. Hardman. Predicting remaining life by fusing the physics of failure modeling with diagnostics. *JOM*, 56(3):29 – 35, 2004.
- D. Karnopp, D.L. Margolis, and R.C. Rosenberg. *Systems Dynamics: Modeling and Simulation of Mechatronic Systems*. John Wiley, New York, 2nd edition, 2006.
- M. Lebold and M. Thurston. Open standards for condition-based maintenance and prognostic systems. 5th annual maintenance and reliability conference (marcon 2001). Gatlinburg, USA, 2001.
- J. Luo, M. Namburu, K. Pattipati, L. Qiao, M. Kawamoto, and S. Chigusa. Model-based prognostic techniques. In *Proc. of IEEE AUTOTESTCON 2003*, pages 330–340, 2003.
- A. Muller. *Contribution à la maintenance prévisionnelle des systèmes de production par la formalisation d'un processus de pronostic*. PhD thesis, Université Henri Poincaré, Nancy 1, 2005.
- A. Muller, M.C. Suhner, and B. Iung. Formalisation of a new prognosis model for supporting proactive maintenance implementation on industrial system. *Reliability Engineering and System Safety*, 93:234–253, 2008.
- G. Provan. Prognosis and condition-based monitoring: an open systems architecture. In *5th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, pages 57–62, 2003.
- A.K. Samantaray and B. Ould Bouamama. *Model-Based Process Supervision: A Bond Graph Approach*. Springer, 2008.
- ISO, 13381-1. *Condition monitoring and diagnostics of machines - prognostics - Part1: General guidelines*. Int. Standard, ISO, 2004.
- W.Q. Wang, M.F. Goldnaraghi, and F. Ismail. Prognosis of machine health condition using neuro-fuzzy systems. *Mechanical Systems and Signal Processing*, 18:813–831, 2004.
- G. Vachtsevanos, F.L. Lewis, M. Roemer, A. Hess, and B. Wu. *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. New Jersey, Hoboken: Wiley & Sons, 2006.