Fatigue Crack Growth Prognostics by Particle Filtering and Ensemble Neural Networks

Piero Baraldi[1], Michele Compare[1], Sergio Sauco[1] and Enrico Zio[1, 2]

[1]*Politecnico di Milano, Milano, Italy*
piero.baraldi@polimi.it
michele.compare@mail.polimi.it
enrico.zio@polimi.it

[2]*Chair on Systems Science and the Energetic Challenge, European Foundation for New Energy-Electricité de France, Ecole Centrale Paris and Supelec, France*
enrico.zio@ecp.fr
enrico.zio@supelec.fr

ABSTRACT

Particle Filtering (PF) is a model-driven approach widely used in prognostics, which requires models of both the degradation process and the measurement acquisition system. In many practical cases, analytical models are not available, but a dataset containing a number of pairs component state - corresponding measurement may be available.

In this work, a data-driven approach based on a bagged ensemble of Artificial Neural Networks (ANNs) is adopted to build an empirical measurement model of a Particle Filter for the prediction of the Residual Useful Life (RUL) of a structure whose degradation process is described by a stochastic fatigue crack growth model of literature. The work focuses on the investigation of the capability of the proposed approach to cope with the uncertainty affecting the RUL prediction.

1. INTRODUCTION

The prediction of the Remaining Useful Life (RUL) of a degrading equipment is affected by several sources of uncertainty such as the randomness in the future degradation of the equipment, the inaccuracy of the prognostic model used to perform the prediction and the noise in the sensor data used by the prognostic model to obtain the RUL prediction. Thus, any RUL prediction provided by a prognostic model should be accompanied by an estimate of its uncertainty (Tang et al. 2009; Liu et al. 2011) in order to confidently plan maintenance actions, taking into account the degree of mismatch between the RUL predicted by the prognostic model and the real RUL of the equipment (Coble 2010; Zio 2012). In this respect, a method able to estimate a probability density function of the degrading equipment RUL is PF, which is a model-based approach successfully used in prognostics applications (e.g., Vachtsevanos et al. 2006, Orchard et al. 2005, Orchard & Vachtsevanos 2009, Cadini et al. 2009). PF is a Bayesian tool for non-linear state estimation, which requires (e.g., Gustaffson & Saha 2010, Doucet et al. 2001, Arulampalam et al. 2002):

1) The knowledge of the degradation model describing the stochastic evolution in time of the equipment degradation $\underline{x}$ (in general a multi-dimensional vector):

$$\underline{x}(t+1) = \underline{g}(\underline{x}(t), \underline{\omega}(t)) \quad (1)$$

where $\underline{g}$ is a possibly non-linear vector function and $\underline{\omega}(t)$ is a possibly non-Gaussian noise.

2) A set of measures $\underline{z}(1),...,\underline{z}(t)$ of past and present values of some physical quantities $\underline{z}$ related to the equipment degradation $\underline{x}$. Although z in general is a multi-dimensional vector, in this work it is considered as a mono-dimensional variable; then, the underline notation is omitted.

3) A probabilistic measurement model which links the measure z with the equipment degradation $\underline{x}$:

$$z(t) = h(\underline{x}(t), \underline{\nu}(\underline{x}(t))) \quad (2)$$

where h is a possibly non-linear vector function and $\nu(\underline{x})$ is the measurement noise vector.

In practical cases, the measurement model h may not be available in analytical form but a dataset $T = \{(\underline{x}_n, \underline{z}_n), n = 1,..., N_{training}\}$ containing a number $N_{training}$ of pairs of state $\underline{x}_n$ and corresponding measurement z_n may be available. This is the case, for example, of the piping of deep water offshore well drilling plants, which degrades due to a process of scale deposition. This may cause a decrease, or even a plug, of the cross sections of the tubular. Giving the inaccessibility of the piping, it is usually impossible to acquire a direct, on line, measure of the scale deposition thickness. On the other side, research efforts are devoted to perform laboratory tests to investigate the relationships between the scale deposition thickness and other parameters which can be more easily measured during plant operation, such as pressures, temperatures and brine concentrations. By this way, one can populate a dataset with the values of the measurable parameters for different scale deposition thicknesses, and use the data to build data-driven models for predicting the scale deposition thickness (Moura et al. 2011).

In this work we have developed an ensemble of ANNs (e.g., Baraldi et al. 2012) as model of the measurement equation in a PF scheme. The proposed prognostic approach is applied to a literature case study (Orchard & Vachtsevanos 2009) concerning crack propagation. The obtained results are compared to those which would be obtained by direct using the measurement equation in the PF model, considering the accuracy of the RUL prediction and the capability of the method of providing an estimate of its uncertainty.

2. PARTICLE FILTERING

In PF, a set of N_s weighted particles, which evolve independently on each other according to the probabilistic degradation model of Eq. 1, is considered. The basic idea is that such set of weighted random samples constitutes a discrete approximation of the true probability density function (pdf) of the system state $\underline{x}$ at time t. When a new measurement is collected, it is used to adjust the predicted pdf through the modification of the weights of the particles in a Bayesian perspective. This requires the knowledge of the probabilistic law which links the state of the component to the gathered measure (Eq. 1). From this model, the probability distribution $P(z|\underline{x})$ of observing the sensors output z given the true degradation state $\underline{x}$ is derived (measurement distribution). This distribution is then used to update the weights of the particles upon a new measurement collection. Roughly speaking, the smaller the probability of encountering the acquired measurement value, when the actual component state is that of the particle, the larger the reduction of the particle's weight. On the contrary, a good match between the acquired measure and the particle state results in an increase of the particle importance (for further details, see Arulampalam et al. 2002 and Doucet et al. 2001).

3. BAGGED ENSEMBLE OF ANNS FOR BUILDING THE MEASUREMENT MODEL

A method to estimate the pdf $P(z|\underline{x})$ of the measurement z in correspondence of a give equipment degradation state $\underline{x}$ is proposed in this Section. It is derived from Carney et al. (1999) and Nix &Weigend (1994), and requires the availability of a dataset made of $N_{training}$ couples $(\underline{x}_n, z_n)$.

The underlying hypothesis of this approach is that the measurement model, which is unknown, can be written in the form:

$$z(\underline{x}) = f(\underline{x}) + \nu(\underline{x}) \quad (3)$$

where $f(\underline{x})$ is a biunivocal mathematical function and the measurement noise $\nu(\underline{x})$ is a zero mean Gaussian noise.

The method of Carney et al. (1999) is based on the use of a bagged ensemble of ANNs, which are employed to build an interpolator $\varphi(\underline{x})$ of the available training patterns $T = \{(\underline{x}_n, z_n), n = 1.,..., N_{training}\}$.

The key idea of bagging (Breiman 1999) is to treat the available dataset T as if it were the entire population, and then create alternative versions of the training set, by randomly sampling from it with replacement. This allows providing more stable estimations. In details, a number B of alternative versions $\{T_b^*\}_{b=1}^B$ of T are created by randomly sampling from it with replacement. Using these training sets, the networks $\{\varphi_b(\underline{x}; T_b^*)\}_{b=1}^B$ are built and the output $\varphi_{avg}(\underline{x})$ of the bagged ensemble in correspondence of the generic test state $\underline{x}$ is obtained by averaging the single ANN output according to:

$$\varphi_{avg}(\underline{x}) = \frac{1}{B}\sum_{b=1}^{B}\varphi(\underline{x}; T_b^*) \quad (4)$$

On the other hand, since PF requires the knowledge of the pdf $P(z|\underline{x})$, the estimate of $f(\underline{x})$ does not suffice to apply PF. In this respect, the procedure proposed in Carney et al. (1999) allows to estimate the pdf $P(z|f(\underline{x}))$ from which the pdf $P(z|\underline{x})$ can be obtained, being the function f invertible for hypothesis. The procedure is based on the subtraction of the random quantity $\varphi_{avg}(\underline{x})$ to both sides of Eq. 3:

$$z(\underline{x}) - \varphi_{avg}(\underline{x}) = [f(\underline{x}) - \varphi_{avg}(\underline{x})] + \nu(\underline{x}) \quad (5)$$

The left-hand side of Eq. 5 is a random variable which represents the error of the ensemble output $\varphi_{avg}(\underline{x})$ with respect to the measurement $z(\underline{x})$.

This random error is made up of two contributions (right hand side of Eq. 5):

1. The random difference $f(\underline{x})-\varphi_{avg}(\underline{x})$ between the unknown deterministic quantity $f(\underline{x})$ and the ensemble output $\varphi_{avg}(\underline{x})$. This quantity is a random variable distributed according to $P(f(\underline{x})\,|\,\varphi_{avg}(\underline{x}))$, being $\varphi_{avg}(\underline{x})$ dependent on the random training set T_b, b=1,…, B; i.e., different training sets would lead to different ensemble models and thus to different output $\varphi_{avg}(\underline{x})$. Since $f(\underline{x})-\varphi_{avg}(\underline{x})$ can be seen as the model $\varphi_{avg}(\underline{x})$ error, its variance will be referred to as model error variance and indicated by $\sigma_m^2(\underline{x})$.
2. The intrinsic noise $\nu(\underline{x})$ of the measurement process, whose variance is indicated by $\alpha^2(\underline{x})$.

These two contributions are estimated by means of the procedures described in the two following Sections.

3.1. Distribution of the model error variance

The procedure here used to estimate the distribution $P(\varphi_{avg}(\underline{x})\,|\,f(\underline{x}))$ of the ensemble output $\varphi_{avg}(\underline{x})$ given the true value of $f(\underline{x})$ (i.e., the 'inverse' of $P(f(\underline{x})\,|\,\varphi_{avg}(\underline{x}))$), is based on the assumption that the random variable $f(\underline{x})-\varphi_{avg}(\underline{x})$ is Gaussian with zero mean and standard deviation $\sigma_m(\underline{x})$, which entails that $P(\varphi_{avg}(\underline{x})\,|\,f(\underline{x}))$ is Gaussian with mean $f(\underline{x})$, and that all we need to know is $\sigma_m(\underline{x})$. Notice that residual errors in the output of the ANN are usually not caused by variance alone; rather, there may be biases in the output of the ANN, which invalidate the assumption that the mean of the distribution is zero. However, it is generally accepted that the contribution of the variance in the residual error of the ANN dominates that of the bias (see Stuart et al 1992 for further details on this). Furthermore, the bias in the output of an ensemble of NNs is expected to be smaller than that of the single ANN.

In order to estimate the model error variance $\sigma_m^2(\underline{x})$, the technique in Carney et al. (1999) requires to divide the B networks of the ensemble $\varphi_{avg}(\underline{x})$ into M smaller sub-ensembles, each one containing K networks, and to consider the output $\varphi_{com}^m(\underline{x})$, $m=1,..,M$ of each sub-ensemble as:

$$\varphi_{com}^m(\underline{x}) = \frac{1}{K}\sum_{k=1}^{K}\varphi_k(\underline{x}) \tag{6}$$

The set $\zeta = \{\varphi_{com}^m(\underline{x})\}_{m=1}^M$ constitutes a sampling of M values from the distribution $P(\varphi_{com}(\underline{x})\,|\,\varphi_{avg}(\underline{x}))$ and its sample variance $\hat{\sigma}_m^2(\underline{x})$ could be used to approximate the unknown variance $\sigma_m{}^2(\underline{x})$ of the ensemble output.

Notice that the idea behind this procedure is that by estimating $f(\underline{x})$ with $\varphi_{avg}(\underline{x})$, one can approximate $P(\varphi_{avg}(\underline{x})\,|\,f(\underline{x}))$ by $P(\varphi_{com}(\underline{x})\,|\,\varphi_{avg}(\underline{x}))$. In order to improve the reliability and stability of $\hat{\sigma}_m^2(\underline{x})$, bagging is also performed on the values of ζ. Thus, P bagging re-sampled sets of ζ are gathered:

$$\Gamma = \{\zeta_p^*\}_{p=1}^P \tag{7}$$

where ζ_p^* is the p-th subset containing M values of $\varphi_{com}(\underline{x})$, sampled with replacement from ζ. For any subset ζ_p^*, $p=1,...,P$, the corresponding variance $\sigma_p^{2*}(\underline{x})$ is computed; then, the estimate $\hat{\sigma}_m^2(\underline{x})$ of the variance $\sigma_m^2(\underline{x})$ is calculated as their average value:

$$\hat{\sigma}_m{}^2(\underline{x}) = \frac{1}{P}\sum_{p=1}^{P}\sigma_p^{2*}(\underline{x}) \tag{8}$$

Finally, the estimate of the regression distribution $P(\varphi_{avg}(\underline{x})\,|\,f(\underline{x}))$ proposed by the method is:

$$P(\varphi_{avg}(\underline{x})\,|\,f(\underline{x})) \approx N(\varphi_{avg}(\underline{x}), \hat{\sigma}_m{}^2(\underline{x})) \tag{9}$$

3.2. Distribution of the measurement noise

In this Section, the technique proposed in Nix & Weigend, (1994) is applied to estimate the variance $\alpha^2(\underline{x})$ of the Gaussian zero mean noise $\nu(\underline{x})$ affecting the measurement equation (Eq. 3).

From Eq. 5, one can derive:

$$\begin{aligned} &Var[z-\varphi_{avg}(\underline{x})] = \\ &Var[f(\underline{x})-\varphi_{avg}(\underline{x})]+Var[\nu(\underline{x})]+2E\{[f(\underline{x})-\varphi_{avg}(\underline{x})]\nu(\underline{x})\} = \\ &\sigma_m{}^2(\underline{x})+\alpha^2(\underline{x}) \end{aligned} \tag{10}$$

The last equality is due to the independence of the error $[f(\underline{x})-\varphi_{avg}(\underline{x})]$ from the measurement noise $\nu(\underline{x})$. To explain this, notice that $[f(\underline{x})-\varphi_{avg}(\underline{x})]$ depends on the noise values ν_n affecting the measures $z_n = f(\underline{x}_n)+\nu_n$, $n=1.,...,N_{training}$, in the training data $T=\{(\underline{x}_n, z_n), n=1.,...,N_{training}\}$, which are used to build the

ensemble model $\varphi_{com}^{m}(\underline{x})$, whereas $\nu(\underline{x})$ is the value of the noise affecting the measure of the test data $\underline{x}$, not used for training the model. Thus, $\nu_n \quad n=1.,...,N_{training}$, and the values sampled from $\nu(\underline{x})$ in the test data are different, independent realizations of the same random variable.

Notice also that $\nu^2(\underline{x})$ obeys a Chi-square $\chi^2(\underline{x})$ distribution with 1 degree of freedom.

The term $\sigma_m{}^2(\underline{x})$ can be estimated according to the procedure illustrated in the previous Section 3.1 whereas, being $z(\underline{x})-\varphi_{avg}(\underline{x})$ a zero mean random variable, its variance is given by:

$$Var[z(\underline{x})-\varphi_{avg}(\underline{x})] = E[(z(\underline{x})-\varphi_{avg}(\underline{x}))^2] \tag{11}$$

Thus, in correspondence of the training couples $(\underline{x}_n, z_n)$, $n=1,...,N_{training}$, one can approximate $E\left[\left(z(\underline{x})-\varphi_{avg}(\underline{x})\right)^2\right]$ by $\left(z(\underline{x})-\varphi_{avg}(\underline{x})\right)^2$ and obtain, according to Eq. 10, a dataset formed by the pairs $(\underline{x}_n, \hat{\alpha}_n^2)$, $n=1,...,N_{training}$, where:

$$\hat{\alpha}_n^2 = max\{(z_n-\varphi_{avg}(\underline{x}_n))^2-\hat{\sigma}^2(\underline{x}_n),0\} \tag{12}$$

Finally, in order to estimate $\alpha^2(\underline{x})$ for a generic $\underline{x}$, a single ANN is trained using the dataset $(\underline{x}_n, \hat{\alpha}_n^2)$, $n=1,...,N_{training}$.

3.3. Estimate of the measurement distribution P(z|x)

Being $\varphi_{avg}(\underline{x})$ an estimate of $f(\underline{x})$, the measurement distribution $P(z\,|\,f(\underline{x}))$ can be approximated by the distribution $P(z\,|\,\varphi_{avg}(\underline{x}))$ which can be derived from the distribution $P(\varphi_{avg}(x)\,|\,f(\underline{x}))$ and the distribution of the measurement noise $\nu(\underline{x})$, according to Eq. 5. Since these two distributions are both Gaussian, with means and variances estimated as shown in Sections 3.1 and 3.2, $P(z\,|\,f(\underline{x}))$ is approximated by a Gaussian distribution with mean $\varphi_{avg}(\underline{x})$ and variance $\hat{\sigma}_m^2(\underline{x})+\hat{\alpha}^2(\underline{x})$. Finally, being $f(\underline{x})$ invertible, the distribution of the measurement z in correspondence of a given state $\underline{x}$, $P(z\,|\,\underline{x})$ is given by:

$$P(z\,|\,\underline{x}) \approx P(z\,|\,f(\underline{x})) \approx N(\varphi_{avg}(\underline{x}), \hat{\sigma}_m^2(\underline{x})+\hat{\alpha}^2(\underline{x})) \tag{13}$$

4. CASE STUDY

In this Section, the technique previously described for estimating the measurement distribution $P(z\,|\,\underline{x})$ is applied to a case study derived from Orchard & Vatchsevanos (2009), which deals with the crack propagation phenomenon in a component subject to fatigue load. The system state is described by the vector $\underline{x}(t)=(x_1(t),x_2(t))$, whose first element, $x_1(t)$, indicates the crack depth whereas the second element, $x_2(t)$, represents a time-varying model parameter that directly affects the crack growth rate. The evolution of this degradation process is described by the following two equations, which form a Markovian system of order one:

$$x_1(t+1) = x_1(t)+3\cdot 10^{-4}(0.05+0.1\cdot x_2(t))^3+\omega_1(t) \tag{14}$$

$$x_2(t+1) = x_2(t)+\omega_2(t) \tag{15}$$

where $\omega_1(t)$ is a Gaussian noise with mean 0.045 and standard deviation 0.116, and $\omega_2(t)$ is a zero mean Gaussian noise with standard deviation 0.010.

In the present case study, the measurement equation is assumed to be unknown whereas a dataset formed by the $N_{training}$ pairs $(x_{1,n}, z_n)$ $n=1,...,N_{training}$, is available, where the subscript 1 refers to the first component of vector $x(t)$.

In practice, given the purpose of the present work of showing the feasibility of the proposed approach, the dataset $T=\left\{(x_{1,n},z_n), n=1,...,N_{training}\right\}$ has actually been artificially obtained by simulating the behavior of the degradation process $\underline{x}(t)$, and sampling from the probabilistic measurement model (Orchard & Vachtsevanos 2009):

$$z(t) = f(x_1)+\nu(x_1) = x_1(t)+0.25+\nu(x_1) \tag{16}$$

where $\nu(x_1)$ is a zero mean Gaussian noise, whose standard deviation depends on x_1:

$$Std[\nu(x_1)] = -\frac{1}{120}x_1^2+\frac{1}{10}x_1+\frac{1}{2} \tag{17}$$

According to Eq. 16, the function $f(\underline{x})=f(x_1)$ is given by x_1+0.25, which is, as required by the method, an invertible function.

To conclude this Section, notice that the probabilistic measurement model in Eq.(9) has been intentionally taken simple, being the main interest of this work the quantification of the uncertainty in the RUL prediction and not the ensemble ability in reproducing the measurement equation. In this respect, the knowledge of the variance of the measurement noise is fundamental, as it determines the amplitude of the prediction intervals of the RUL estimates. Thus, the capability of correctly reconstructing the variance behavior plays a key role in the assessment of the potential of the proposed technique.

4.1. Estimate of the measurement distribution

According to the technique illustrated in Section 3, an ensemble of $B=200$ ANNs has been built using the

available dataset $T=\left\{(x_{1,n},z_n), n=1,...,N_{training}\right\}$, where $N_{training}=1000$. Every ANN has 5 tan-sigmoidal hidden neurons and one linear output neuron. To estimate $\sigma_m^2(\underline{x})=\sigma_m^2(x_1)$, the ensemble has been divided into $M=20$ sub-ensembles and $P=1000$ bagging resamples of the sub-ensemble outputs $\varphi_{com}^m(\underline{x})=\varphi_{com}^m(x_1)$, $m=1,...,M$, have been considered.

The results are evaluated in terms of the following performance indicators, which are computed by considering a set of $N_{test}=1000$ pairs $(x_{1,i},z_i)$, $i=1,...,N_{test}$ which have been obtained from Eq. 16 and 17:

1. The square bias $\bar{b}^2$; i.e., the average quadratic difference between the true value of $f(x_1)$ and the ensemble estimate of this quantity $\varphi_{avg}(x_1)$:

$$\bar{b}^2=\frac{1}{N_{test}}\sum_{i=1}^{N_{test}}(f(x_{1,i})-\varphi_{avg}(x_{1,i}))^2 \quad (18)$$

This value gives information on the accuracy of the estimate of $f(\underline{x})=f(x_1)$ provided by the ensemble. Notice that the computation of this indicator requires the knowledge of the function $f(x_1)$, which is not available if the measurement equation (Eq. 16) is not known. Thus, in general one can only compute:

$$MSE=\frac{1}{N_{test}}\sum_{i=1}^{N_{test}}(\varphi(x_{1,i})-z_i)^2 \quad (19)$$

Small values of *MSE* indicate satisfactory performance of the ensemble.

2. The coverage of the Prediction Interval (PI) with confidence 0.68. This indicator is used to verify the accuracy of the estimate of the distribution $P(z|\underline{x})=P(z|x_1)$. A PI with a confidence level γ_p is defined as a random interval in which the observation $z(\underline{x})=z(x_1)$ will fall with probability γ_p (Carney et al. 1999, Heskes 1997):

$$P(z(x_1)\in PI_{\gamma_p}(x_1))=\gamma_p \quad (20)$$

Being the estimate of $P(z|x_1)$ a Gaussian distribution with mean $\varphi_{avg}(x_1)$ and variance $\hat{\sigma}_m^2(x_1)+\hat{\alpha}^2(x_1)$, the PI with $\gamma_p=0.68$ is given by:

$$\begin{aligned}&\varphi_{avg}(x_1)-\sqrt{\hat{\sigma}_m^2(x_1)+\hat{\alpha}^2(x_1)}\le z(x_1)\\ &z(x_1)\le\varphi_{avg}(x_1)+\sqrt{\hat{\sigma}_m^2(x_1)+\hat{\alpha}^2(x_1)}\end{aligned} \quad (21)$$

In order to verify whether the estimate of $P(z|x_1)$ provides a satisfactory approximation of the true pdf, we will consider how many times the measurement z_i falls within the $PI_{\gamma_p=0.68}(x_{1,i})$. The closer to γ_p the portion of points hitting the γ_p-confidence interval, the more accurate the estimation of the parameters of the Gaussian pdf.

In practice, for every $x_{1,i}$, $i=1,...,N_{test}$, a counter C_i is set to 1 or 0 depending on whether the z_i belongs or not to the estimated $PI_{\gamma_p=0.68}(x_{1,i})$. The closer the average of C_i, $i=1,...,N_{test}$ to 0.68, the better the approximation.

Cross-validation of the results has been done by repeating the computations with $N_{set}=25$ different, randomly generated training and test sets. This avoids over/under estimations of the performance indicators $\bar{b}^2$ and coverage.

Table 1 reports means and standard deviations (std) of the performance indicators over the 25 cross-validations.

model	Ensemble	1 ANN
$\bar{b}^2$	0.0040 ± 0.0015	0.0097 ± 0.0060
PI coverage	0.6758 ± 0.0366	-

Table 1: Performance indicators over 25 cross-validations; the mean ± std is reported

Notice that the ensemble output $\varphi_{avg}(x_1)$ is very accurate in the prediction of the function $f(x_1)$, the bias being very small. Furthermore, notice that the ensemble outperforms a single ANN trained with all the 1000 training patterns. With respect to the estimate of the distribution $P(z|x_1)$, the proposed method provides a satisfactory approximation, being the coverage very close to 0.68.

Table 2 reports the estimates of the two contributions $\bar{\sigma}_m^2$ and $\bar{\alpha}^2$ of the variance of the estimated measurement distribution $\hat{P}(z|x_1)$. Notice that in this case study, $\bar{\sigma}_m^2$ is negligible with respect to the variance $\bar{\alpha}^2$ of the measurement noise; this entails that the accuracy of the estimate of the PI is more sensible to the estimate of $\bar{\alpha}^2$.

In this respect, Figure 1 shows the estimate of $\alpha^2(x_1)$ and compares it to the true $\bar{\alpha}^2$ value provided by Eq. 17. Notice that this comparison, which is done in this work to assess the performance of the methodology, is not possible in real industrial applications if the measurement model (Eqs. 16 and 17) is not available.

	Estimation	Real value

$\bar{\sigma}_m^2 + \bar{\alpha}^2$	0.4489±0.1359	-
$\bar{\sigma}_m^2$	0.0243±0.0317	-
$\bar{\alpha}^2$	0.4886±0.0276	0.4900

Table 2: Contributions to the $P(z \mid x_1)$ variance

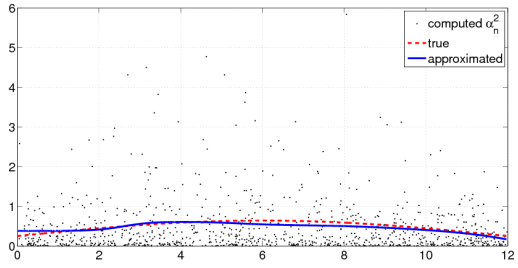

Figure 1: True and approximated measurement noise variance $\alpha^2(x_1)$

4.2. Crack depth prediction

The objective of this Section is to evaluate the performance of the overall scheme in the prediction of the crack depth evolution when the ensemble of ANNs is used to estimate the measurement distribution $P(z \mid \underline{x})$. To this purpose, the problem tackled consists in predicting at $t = 80$ (in arbitrary units) the future crack propagation, on the basis of eight measurements of the crack depth taken at time $t_m = 10 \cdot m$, m=1,…, 8. This prediction phase is performed by considering the evolution of the particles according to the model in Eqs. 14 and 15 (e.g., see Orchard & Vachtsevanos 2009). In particular, we focus on the time instant t=80, when the PF updates via $P(z \mid \underline{x})$ the particles' weights after the last measurement (z=4.6087 in arbitrary units) has been acquired.

Figure 2 shows the prediction of the crack depth evolution performed at t=80, after the acquisition of the last measurement, by using the ensemble model to estimate $P(z \mid \underline{x})$. This prediction has been compared to that which would be obtained by directly using the measurement equation in the PF.

Notice that the linearity of the prediction of the expected value of x_1 can be explained by averaging Eqs. 14 and 15:

$$E[x_2(t+1)\} = E[x_2(t)] + E[\omega_2(t)] = E[x_2(t)] = \text{constant}$$

$$E[x_1(t+1)] - E[x_1(t)] = 3 \cdot 10^{-4}(0.05 + 0.1 \cdot E[x_2(t)])^3 + E[\omega_1(t)]$$

$$E[x_1(t+1)] - E[x_1(t)] = \text{constant}$$

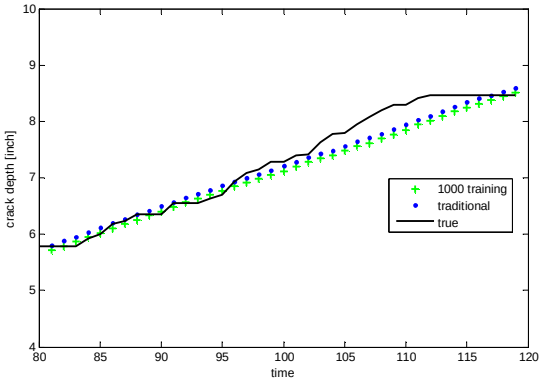

Figure 2: Comparison of the predictions with the true state evolution

To evaluate the impact of replacing the measurement equation with the ensemble of ANNs, $N_{run} = 100$ different degradation trajectories have been simulated and the predictions of the crack depth have been performed.

Also in this case, the prediction provided by the ensemble of ANNs trained with $N_{training} = 1000$ patterns has been compared to that based on the analytical measurement equation $P(z \mid \underline{x})$. Each run is characterized by the same true trajectory, the same acquired measures and the same state noise vector. The following performance indicators have been computed:

1. The coverage of the PI, with confidence 0.68. In particular, the prediction of the crack depth at $t = 120$ has been considered. At each run, the boundaries of the PI are computed by considering the 16th and 84th percentiles of the estimate of the pdf of the crack depth. A counter is set to 1 or 0 if the true trajectory belongs or not to the corresponding interval, in analogy with the coverage verification explained in Section 4.2.

2. The average width over the $N_{run} = 100$ runs of the PI at $t = 120$.

3. The Mean Square Error (MSE) over the $N_{run} = 100$ runs between the prediction of the crack depth provided by using the PF and its true value at $t = 120$. That is:

$$MSE_{120} = \frac{1}{N_{run}} \sum_{n_{run}=1}^{N_{run}} \left(X_{n_{run}} - o_{n_{run}} \right)^2 \qquad (22)$$

where $X_{n_{run}}$ is the true crack depth in the test trajectory at $t=120$ and $o_{n_{run}}$ is the expected value of the crack depth pdf estimated by the PF.

The obtained values are reported in Table 4. It can be noticed that the coverage of the ensemble is very close to 0.68; furthermore, even the other performance indicators are very close to those which would be obtained by considering the measurement equation. This result confirms that the approximation of the distribution $P(z|\underline{x})$ is accurate and therefore it does not remarkably alter the outcome of the PF.

	Traditional	Data-driven
coverage	0.6500	0.7000
PI width	1.3058	1.3226
MSE	0.3421	0.3464

Table 4: Performance indicators at *t=120*

Finally, the performance evaluator *s* proposed by Saxena et al. 2008 has been computed to evaluate the prediction performance:

$$s=\begin{cases}\sum_{i=1}^{n} e^{-\left(\frac{d_i}{a_1}\right)}-1 & if\ d_i<0\\ \sum_{i=1}^{n} e^{-\left(\frac{d_i}{a_2}\right)}-1 & otherwise\end{cases}$$

where a_1= 10, a_2=13, n=100 is the number of simulated histories and d is the difference between the estimated RUL and its true value. To compute the value of this performance metric, the following procedure has been adopted:

1. Set the failure threshold to S_T=7.
2. Simulate the evolution of the degradation process; this allows calculating the true value of the true RUL t_{RUL} at t=80 as the difference between the time instant at which the component achieves ST and 80. Moreover, the set of measures sampled according to the measurement model are collected.
3. Use the PF to estimate the component degradation state at t=80 and predict the RUL $\hat{t}_{RUL}$.
4. Calculate the difference $d=\hat{t}_{RUL}-t_{RUL}$.
5. Perform *n-1* times the steps 2-4 and compute *s*.

The values of the metric *s* obtained in the case in which the RUL is predicted by using the 'traditional' PF approach (*s*=10.30) and the 'data–driven' approach (*s*=10.65) are very close to each other.

5. CONCLUSIONS

PF is often proposed as prognostic technique for estimating the evolution of the degradation state $\underline{x}$ of a system; generally, it resorts to analytical models of both degradation state evolution and measurement. In practice, the measurement model may not be available in an analytical form; rather, there may be available a set of data which allows, through data-mining techniques, building the measurement model. In this work, a technique based on an ensemble of ANNs has been investigated to this aim and applied to a case study derived from the literature. The verification conducted on the results shows that a good approximation of the model may be obtained and its substitution in the PF does not significantly affect its performance. Furthermore, the proposed method has been shown capable of estimating the uncertainty on the RUL prediction.

Additional effort will be dedicated in future works to improve the accuracy of the estimate when only a small training set is available and to extend the applicability of the technique also in those cases in which the measurement equation $f(\underline{x})$ is not biunivocal or has a more complex form. Furthermore, another future objective is the substitution also of the model of the evolution of the system state with a data-driven model, e.g., an ensemble of trained ANNs, in order to allow the usage of PF in those cases where also an analytical model of the evolution of the system is unavailable.

REFERENCES

Arulampalam, M.S., Maskell, S., Gordon, N. and Clapp, T. (2002). A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking, *IEEE Transactions on Signal Processing*, 50 (2), 174-188.

Baraldi, P., Di Maio, F. Zio, E., Sauco, S. Droguett, E., Magno, C. (2012) Ensemble of Neural Networks for Predicting Scale Deposition in Oil Well Plants Equipments, proceedings of PSAM 11 & ESREL 2012.

Breiman, L. (1999) Combining predictors, in Sharkey AJC (Ed.) *Combining artificial neural nets: ensemble and modular multinet systems*. Springer, Berlin Heidelberg New York, pp 31-50.

Cadini, F., Zio, E., Avram, D. (2009) Model-based Monte Carlo state estimation for condition-based component replacement, *Reliability Engineering & System Safety*, Vol. 94 (3), pp. 752-758.

Carney, J., Cunningham, P., & Bhagwan, U. (1999). Confidence and prediction intervals for neural network ensembles. *International Joint Conference on Neural Networks IJCNN*, July 10-16, Washington D.C.

Coble, J.B. (2010) Merging Data Sources to Predict Remaining Useful Life – An Automated Method to Identify Prognostic Parameters. PhD diss., University of Tennessee.

Doucet, A., de Freitas, J.F.G. and Gordon, N.J. (2001) *Sequential Monte Carlo methods in practice*. Springer-Verlag, New York.

Gustaffson, F., & Saha, S. (2010). Particle filtering with dependent noise. In Proceedings of the *13th Conference on Information Fusion (FUSION)*. Edinburgh.

Heskes, T. (1997) Practical Confidence and Prediction Intervals, in M. Mozer, M. Jordan and T. Peskes, editors, *Advances in Neural Information Processing Systems*,vol 9, pages 466-472, Cambridge, 1997, MIT Press.

Hsu, C.W. , Chang, C.C., Lin, C.J. (2003) A Practical Guide to Support Vector Classification. Technical Report, 2003.

Liu, R., Ma, L., Kang, R. and Wang, N. (2011) The Modeling Method on Failure Prognostics Uncertainties in Maintenance Policy Decision Process, *Proc. 9th Int. Conf. on Reliability, Maintainability and Safety (ICRMS 2011)*, pp. 815-820.

Moura, M.C., Lins, I.D., Ferreira, R.J., Droguett, E.L., Jacinto, C.M.C. Predictive maintenance policy for oil well equipment in case of scaling through support vector machines, in Proceedings of the European Safety and Reliability Conference-ESREL 2011, pp. 503-507.

Nix, D. and Weigend, A. (1994). Estimating the mean and the variance of the target probability distribution," in *IEEE world congress on computational intelligence, International Joint Conference on Neural Networks*, June 27-July 2, Orlando, Florida.Vol 1, pp. 55-60.

Orchard, M., Wu, B., Vachtsevanos, G. (2005) A Particle Filter Framework for Failure Prognosis, Proceedings of WTC2005 World Tribology Congress III. Washington D.C., USA, Sept. 12-16, 2005

Orchard, M. and Vachtsevanos, G. (2009) A Particle Filtering Approach for On-Line Fault Diagnosis and Failure Prognosis, *Transactions of the Institute of Measurement and Control*, Vol. 31 (3-4), pp. 221-246.

Papoulis, A. and Pillai, S.U. (2002) *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Higher Education, 4th edition.

Saxena, A., Goebel K, Simon, D., Eklund, N. (2008) Damage Propagation Modeling for Aircraft Engine Run-to-Failure Simulation, in International Conference on Prognostics and Heath Management, PHM2008.

Stuart, G., Bienenstock, E. and Doursat, R. (1992) Neural Networks and the bias/variance dilemma. *Neural Computation*, Vol. 4(1), pp. 1-58.

Tang, L., Kacprzynski, G.J., Goebel, K. and Vachtsevanos, G. (2009) Methodologies for Uncertainty Management in Prognostics. Proc. IEEE Aerospace conference, 2009, pp. 1-12.

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

Zio. E. (2012) Prognostics and health management of industrial equipment, in *Diagnostics and Prognostics of Engineering Systems: Methods and Techniques*, S. Kadry, Eds. IGI-Global, 2012.

BIOGRAPHIES

Piero Baraldi (BS in nuclear engng., Politecnico di Milano, 2002; PhD in nuclear engng., Politecnico di Milano, 2006) is assistant professor of Nuclear Engineering at the department of Energy at the Politecnico di Milano. He is the current chairman of the European Safety and Reliability Association, ESRA, Technical Committee on Fault Diagnosis. His main research efforts are currently devoted to the development of methods and techniques for system health monitoring, fault diagnosis, prognosis and maintenance optimization. He is co-author of 42 papers on international journals and 38 on proceedings of international conferences, and serves as referee of 5 international journals.

Michele Compare (BS in mechanical engng., University of Naples Federico II, 2003, PhD in nuclear engng., Politecnico di Milano, 2011) is currently a post-doc at the Politecnico di Milano. He worked as RAMS engineer and risk manager. His main research efforts are devoted to the development of methods and techniques in support of maintenance of complex systems.

Sergio Sauco BS in Energy engng. 2009, Politecnico di Milano; MS in Nuclear engng., Politecnico di Milano, 2011.

Enrico Zio (BS in nuclear engng., Politecnico di Milano, 1991; MSc in mechanical engng., UCLA, 1995; PhD, in nuclear engng., Politecnico di Milano, 1995; PhD, in nuclear engng., MIT, 1998) is Director of the Chair in Complex Systems and the Energetic Challenge of Ecole Centrale Paris and Supelec, full professor, Rector's delegate for the Alumni Association and past-Director of the Graduate School at Politecnico di Milano, adjunct professor at University of Stavanger. He is the Chairman of the European Safety and Reliability Association ESRA, member of the Korean Nuclear society and China Prognostics and Health Management society, and past-Chairman of the Italian Chapter of the IEEE Reliability Society. He is serving as Associate Editor of IEEE Transactions on Reliability and as editorial board member in various international scientific journals. He has functioned as Scientific Chairman of three International Conferences and as Associate General Chairman of two others. His research topics are: analysis of the reliability, safety and security of complex systems under stationary and dynamic conditions, particularly by Monte Carlo simulation methods; development of soft computing techniques for safety, reliability and maintenance applications, system monitoring, fault diagnosis and prognosis. He is author or co-author of five international books and more than 170 papers on international journals.