CHEMICAL ENGINEERING TRANSACTIONS
VOL. 33, 2013
Guest Editors: Enrico Zio, Piero Baraldi

ISBN 978-88-95608-24-2; **ISSN** 1974-9791

A publication of
AIDIC
The Italian Association
of Chemical Engineering
Online at: www.aidic.it/cet

DOI: 10.3303/CET1333028

Prognostics under Different Available Information

Piero Baraldi*[a], Francesco Cadini[a], Francesca Mangili[a] and Enrico Zio[a,b]

[a]Dipartimento di energia, Politecnico di Milano, Italy
[b]Chair on Systems Science and the Energetic challenge, European Foundation for New Energy-EDF, Ecole Centrale Paris and Supelec, Paris, France
piero.baraldi@polimi.it

In this work, we consider two practical situations with different information available, concerning the prediction of the Remaining Useful Life (RUL) of a creeping turbine blade for which a sequence of observations of the creep strain level is available. In the first case considered, we have available a stochastic model of the creep growth process and we know the value of the failure threshold, i.e., the maximum creep strain level beyond which the blade cracks. On this basis, a Monte Carlo-based filtering technique, called particle filtering, is set-up to predict the distribution of the system RUL and online-update it when new observations are collected. In the second case considered, the only available information is the sequence of observations of the creep strain of the blade of interest and the value of the failure threshold. On this basis, a data-driven method, based on an ensemble of bootstrap models, has been developed to estimate the turbine blade RUL and the uncertainty of the estimate caused by the uncertainty in the data, the variability of the blades behaviour and the imprecision of the empirical model. The two approaches are evaluated in terms of the assumptions they require and the accuracy of the RUL predictions they provide. The ability of providing measures of confidence in the outcomes is also considered.

1. Introduction

Different forms of information and data may be available for prognostics; depending on that, different prognostic methods may be applied. In this work, we consider two practical situations about prognostics of the creep growth process in the turbine blades of a Gas Turbine Modular Helium nuclear Reactor (GT-MHR) simulated by a stochastic model based on the Norton law (Baraldi et al., 2013). For each situation considered, an accurate and robust prognostic method is proposed.

In general, prognostic methods can be classified in model-based and data-driven (Brotherton et al., 2000). Model-based methods use an explicit mathematical model of the degradation process to predict the future evolution of the degradation state and, thus, the RUL of the system. In practice, even when the model of the degradation process is known, the RUL estimate may be difficult to obtain, since the degradation state of the system may not be directly observable and/or the measurements may be affected by noise and disturbances. In these cases, model-based estimation methods aim at inferring the dynamic degradation state and provide a reliable quantification of the estimation uncertainty on the basis of the sequence of available noisy measurements. In the first case considered in this work, hereafter referred to as case 1, we have available a stochastic model of the creep growth process based on the Norton law and we know the value of the failure threshold, i.e., the maximum level of deformation beyond which the blade cracks. Also, a sequence of direct observations of the degradation of the blade, measured by its creep strain, is available. On this basis, model-based approaches are suggested. Many approaches rely on Bayesian methods: due to its flexibility and ease of design, a numerical approximation of the Bayesian estimate based on the Monte Carlo sampling technique and called particle filtering, is set-up in this work to predict the distribution of the system RUL and online-update it when new observations are collected (Baraldi et al., 2013).

On the other side, data-driven methods are used when an explicit model of the degradation process is not available, but sufficient historical data have been collected. These methods are based on statistical models

Please cite this article as: Baraldi P., Cadini F., Mangili F., Zio E., 2013, Prognostics under different available information, Chemical Engineering Transactions, 33, 163-168 DOI: 10.3303/CET1333028

that ‘learn’ trends from the data. Recently, 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 and the possibility of estimating the uncertainty of the predictions (Baraldi et al., 2012), (Baraldi et al., 2013). In this work, we consider a situation, hereafter referred to as case 2, in which only the sequence of direct creep strain measures for the blade of interest is available and the value of the failure threshold is known. On this basis, an ensemble of bootstrap models is built to estimate the system RUL and the uncertainty of the estimate, caused by the uncertainty in the data, the variability of the system behaviour and imprecision of the empirical model.
The remainder of the paper is organized as follows: in Section 2, the problem of blade creeping in high temperature turbines is illustrated, the sources of information for prognostics discussed, and the two cases considered as well as the objectives of the prognostic activity are presented; in Section 3, the prognostic methods developed to tackle the two prognostic cases are described; in Section 4, results are discussed; finally, in Section 5 some conclusions are drawn and potential for future work suggested.

2. Creep Growth Case Study

Creep is an irreversible deformation process affecting materials exposed to a load below their elastic limit for a protracted length of time and at high temperature. A turbine undergoing this degradation process can experience the loss of its blades, one of the most feared failure modes of turbomachinery since it is accompanied by abrupt changes in the power conversion system and in the reactor flow conditions (Saez et al. 2006).
The uniaxial creep deformation consists of an augmentation of the original length, and a reduction of the diameter. In this work, the dimensionless quantity ε, defined as the percentage of elongation of the turbine blade in the longitudinal direction with respect to its original length, is considered as measure of the creep strain. Traditional methods for predicting creep life, based on the extrapolation of some creep constitutive equations from accelerated tensile creep tests, are very useful during the design phase, but are not enough accurate for blade health monitoring during operations (Penny and Marriott, 1995).

2.1 Information and data for prognostics

The main sources of degradation-related information for the creep growth process, listed in Table 1, are further detailed in this Section (Baraldi et al., 2013).
Information a: C*reep growth model.* Creeping in turbine blades is a stochastic degradation process which can be modeled using the Norton Law assuming that the dependence from the temperature follows the Arrhenius law. Moreover, for Δt sufficiently small (here $\Delta t = 5$ days, with respect to the time horizon of several thousand), the Norton Law can be discretized to give (Baraldi et al., 2013)

$$\varepsilon_{j+1} = \varepsilon_j + A\exp\left(-\frac{Q}{RT_j}\right)\left(K\omega_j^2 + \delta\varphi_j\right)^n \Delta t \quad \varepsilon_0 = 0 \tag{1}$$

where ε_j is the creep strain at time t_j, Q is the activation energy, A and n are material inherent characteristics varying from one blade to another, K is a constant relating the load to the rotational speed ω_j, R is the ideal gas constant, T_j is the blade operating temperature, and $\delta\varphi_j$ is a random variable modeling the fluctuations in the stress applied to a specific blade, which are due to fabrication defects, aging and corrosion of the blade, vibrations of the system or turbulences of the gas flow. The values of the parameters T_j, ω_j, and K have been set as in Baraldi et al. (2013) with reference to the helium gas turbine of a Gas Turbine Modular Helium nuclear Reactor (GT-MHR).
Information b: creep strain measurements. This source of information consists in a sequence of observations $z_{1:i}$ of creep strain, hereafter called ‘test trajectory’, performed on the blade of interest. Given the unavailability of real experimental data, in this work the creep growth trajectory is simulated using Eq. (1) and following the procedure described in Baraldi et al. (2013). A total number of 87 creep strain measurements have been obtained for a turbine blade by adding a white Gaussian noise υ_j with standard deviation $\sigma_\upsilon = 0.02$ to the simulate creep strain value $\varepsilon_{j.}$
Information c: Failure threshold. The failure threshold for creep strain ε_{th} is set equal to the value of 1.5%.
Information d: Measurement equation. Since it has been assumed that the value of the creep strain is directly measured with sensor noise $\upsilon_{j,}$ the observation equation is:

$$z_j = \varepsilon_j + \upsilon_j \tag{2}$$

Table 1. Main sources of information for creep growth prognostics

Source	Description of the source of information	Mathematical representation
a	The creep growth model and the distributions of the model parameters	Eq. (1)
b	Measurements of the creep strain of the currently creeping blade taken at i different time instants t_j	$z_{1:i} = \varepsilon_{1:i} + \upsilon_{1:i}$
c	The value of the failure threshold	ε_{th}
d	The measurement equation and the noise distribution	Eq. (2) and σ_v

2.2 Two prognostic cases with different sources of information

By differentiating the source of information available, two cases are considered. In both cases a set of measurements $z_{1:i}$ collected during the life of the turbine blade of interest (source of information *b*) is available in combination with other different sources of information (Table 2).

In case 1, the physical model of the creep growth process (Eq. 1) is known, as well as the distribution and evolution in time of all its characteristic and external parameters (Table 1, source *a*). Other sources of information available are the value of the failure threshold ε_{th} (source *c*) and the measurement equation (source *d*), linking the observations with the creep strain (Eq. 2). In case 2, the information available is limited to the value of the failure threshold ε_{th} (source *c*).

Table 2. Information available in each prognostic case considered

Case	Source of information			
	a	*b*	*c*	*d*
1	X	X	X	X
2		X	X	

Aim of the work

The aim of this work is to use the information available for estimating the RUL of the degrading turbine blade, i.e., the time left from the current time t_i before its creep strain crosses the failure threshold ε_{th}. Since the evolution of the creep is intrinsically random, the blade RUL at time t_i is a random variable, RUL_i, and, thus, the objective of applying a prognostic method to a blade whose current creep strain is ε_i is to estimate the probability distribution $\Pr(RUL_i|\varepsilon_i)$ of RUL_i. Notice that the uncertainty described by this distribution regards the future stochastic evolution of creep degradation and thus is irreducible.

In practical cases, the current degradation level ε_i and the degradation model can be not exactly known. Thus, due to the limited information available, one is interested in:

- The expected value $\mu_{RUL_i|\varepsilon_i}$ of RUL_i;
- The variance of the prediction error $\sigma^2_{r\hat{u}l_i} = E[(r\hat{u}l_i - RUL_i)^2]$, as a measure of the accuracy with which the estimated expected value, $r\hat{u}l_i$, predicts the actual RUL value, rul_i.

The contributors to the uncertainty of the prediction $r\hat{u}l_i$ can be classified in the following three categories (Baraldi et al., 2013):

A. Variability in the future creep strain rate, mainly due to the unforeseen future values of the blade load $K\omega^2 + \delta\varphi$, and temperature T;
B. Inaccuracy of the prognostic model used to perform the prediction;
C. Sensor noise affecting the measurement z_i of the current degradation level, which is fed in input to the prognostic model.

It can then be useful to decompose the error variance $\sigma^2_{r\hat{u}l_i}$ of the model prediction $r\hat{u}l_i$ in three terms: the process uncertainty σ^2_A, the model uncertainty σ^2_B, and the noise uncertainty σ^2_C (Baraldi et al., 2013). To this aim, introducing the quantity $\mu_{RUL_i|z_i}$ which represents the expected value of RUL_i for a degrading equipment for which at time t_i we have the observation z_i, and assuming that the prognostic model is an unbiased estimator of $\mu_{RUL_i|z_i}$, we obtain:

$$\begin{aligned}\sigma^2_{r\hat{u}l_i} &= E[(RUL_i - r\hat{u}l_i)^2] = \sigma^2_A + \sigma^2_B + \sigma^2_C = E\left[\left(RUL_i - \mu_{RUL_i|\varepsilon_i} + \mu_{RUL_i|\varepsilon_i} - \mu_{RUL_i|z_i} + \mu_{RUL_i|z_i} - r\hat{u}l_i\right)^2\right] \\ &= E\left[\left(RUL_i - \mu_{RUL_i|\varepsilon_i}\right)^2\right] + E\left[\left(\mu_{RUL_i|z_i} - r\hat{u}l_i\right)^2\right] + E\left[\left(\mu_{RUL_i|z_i} - \mu_{RUL_i|\varepsilon_i}\right)^2\right]\end{aligned} \quad (3)$$

3. Modeling approaches

This Section illustrates the two modeling approaches undertaken to cope with the two prognostic cases outlined in Section 2.2 (Table 2).

3.1 Particle Filtering (Case 1)

In case 1, at time t_i, the stochastic creep dynamic model (Eq. (1)), the observation equation (Eq. (2)), the sequence of i measures $z_{1:i}$ of creep strain and the value of the failure threshold ε_{th} (i.e., sources of information *a*, *b*, *c* and *d*) are available.

In this setting, it is possible to predict, within a Bayesian framework, the filtered posterior distribution $\Pr(RUL_i|z_{1:i})$ of the blade RUL at time t_i given the observations $z_{1:i}$ collected up to time t_i by a recursive computational procedure divided into successive prediction and update stages. In the creep growth process, since the combination of speed, temperature and stress fluctuations described by Eq. (1) entails a non-Gaussian noise, the exact computation of $\Pr(RUL_i|z_{1:i})$ involves the solution of an integral which does not have, in general, a closed-form solution. Then, a numerical approximation based on Monte Carlo sampling, the particle filtering method, has been applied in this work for its flexibility and ease of design. The particle filtering technique provides a solution to the prognostic problem by approximating the integrals in the Bayesian recursive procedure with weighted summations over a high number of samples called particles (Cadini et al., 2009).

A number $P = 1000$ of particles, representing degradation trajectories, are built by sampling for each of them the characteristic parameters A and n from their distributions and then recursively sampling the particle degradation level ε_i^p at time t_i according to Eq. (1), until the failure threshold ε_{th} is exceeded and the duration of life L_p of the particle recorded. When an observation z_i is collected, each particle is assigned a weight w_i^p proportional to the likelihood $\Pr(z_i|\varepsilon_i^p)$ of observing z_i given the degradation level ε_i^p the particle has reached. The weighted average and the weighted standard deviation of the particle RULs at time t_i represent the prediction $r\hat{u}l_i$ of the expected value $\mu_{RUL_i|\varepsilon_i}$ of RUL_i and the estimate $\hat{\sigma}^2_{r\hat{u}l_i}$ of the prediction error variance, $\sigma^2_{r\hat{u}l_i}$, respectively. For further details on this method, the interested reader is referred to Baraldi et al. (2013).

3.2 Data-driven ensemble of bootstrapped models (Case 2)

This case is characterized by the availability of the sequence of creep strain measurements $z_{1:i}$ for the blade of interest and the value of the failure threshold ε_{th} (i.e., sources of information *b* and *c*). In this case, the prognostic model has been developed only after time t_{30} in order to have available a dataset $\mathbf{D} = \{z_{1:i}\}$ of at least $i = 30$ direct creep strain measurements.

The solution proposed, in this context of available information, is to develop an empirical model of the time evolution of the creep strain $\hat{\varepsilon} = \eta(t)$ on the basis of the sequence of past creep strain measurements $z_{1:i}$. The predicted failure time is then given by the time at which the degradation will exceed the known degradation threshold ε_{th}. In principle, linear regression methods could be applied to derive a model $\hat{\varepsilon} = \eta(t)$. However, they usually do not allow estimating the process uncertainty σ_A^2. Furthermore, the bootstrapped ensemble approach is here preferred for estimating the model uncertainty σ_B^2 due to the fact that it can be extended to more complex non-linear problems. To this purpose, the bootstrap method described in Baraldi et al. (2013) is applied.

The general idea is to partition the set of creep strain measurements $\mathbf{D}$ into two non overlapping datasets of consecutive measurements $\mathbf{D}^{trn}$ and $\mathbf{D}^{val}$. According to the bootstrap method, *H* different linear models are built using the training dataset $\mathbf{D}^{trn}$ and aggregated by averaging their outcomes to obtain the single prediction $r\hat{u}l_i$ of the blade RUL at time t_i. With respect to the estimate of the error variance $\sigma^2_{r\hat{u}l_i} = \sigma_A^2 + \sigma_B^2 + \sigma_C^2$, the variance of the outcomes of the ensemble provides the estimate of σ_B^2, whereas the remaining terms $\sigma_A^2 + \sigma_C^2$ are estimated by building a model χ of the errors made by the bootstrap ensemble model when used to predict the time $\Delta t_{j,j'}$ needed to obtained the creep strain increments $\Delta z_{j,j'} = z_j - z_{j'}$ observed in the validation dataset $\mathbf{D}^{val}$. The output of model χ, in correspondence of the creep strain increment $\Delta z_{i,th} = \varepsilon_{th} - z_i$ needed for the blade to reach the failure threshold, provides the estimate $\hat{\sigma}^2_{A+C}$ of the variance component σ^2_{A+C} of the error made in predicting rul_i.

Notice that the data used for training model χ concern creep strain increments which for a large part of the creep growth trajectory are smaller than the increment $\Delta z_{i,th}$ considered for obtaining the prognostic results, so that the empirical model χ is used in an input region not described by the training data. This represents a limit to the quality of the estimate $\hat{\sigma}^2_{A+C}$, since, in general, the performance of empirical models are best when they are applied to input regions well described by the training data, and degrade away from these regions.

4. Results

Table 3 reports the true information about the RUL of the simulated turbine blade under test, at two time instants $t_{50} = 1475$ days and $t_{80} = 2375$ days; row 3 reports the true RUL value, rul_i, observed for the turbine blade under test, whereas rows 4 and 5 report the expected value $\mu_{RUL_i|\varepsilon_i}$ and the prediction error variance component σ_A^2 which corresponds to the variance of the distribution $Pr(RUL_i|\varepsilon_i)$ and represents the irreducible uncertainty of the RUL prediction caused by the stochastic future evolution of the creep strain. $Pr(RUL_i|\varepsilon_i)$ has been obtained by simulating $P = 1000$ degradation trajectories all characterized by the values A and n of the blade under test and by the creep strain level ε_i at time t_i.
Table 4 reports the corresponding estimates $r\hat{u}l_i$ of the RUL expected value and $\sigma_{r\hat{u}l_i}$ of the prediction error standard deviation, obtained by applying the two prognostic approaches to the same degrading blade. Columns 3 and 4 refer to the RUL predictions performed at time t_{50} days on the basis of the measurements $z_{1:50}$ of the test trajectory, and at time t_{80} days on the basis of the measurements $z_{1:80}$, respectively.

Table 3. Creep propagation in the simulated turbine blade under test

Parameter	Description	Values [d]	
t_i	Time of the prediction	1475	2375
rul_i	RUL value at time t_i	1110	210
$\mu_{RUL_i|\varepsilon_i}$	Expected value of *RUL*$_i$ given a creep strain value ε_i at time t_i	1092	264
σ_A	Standard deviation of *RUL*$_i$	90	42

Table 4. Creep propagation estimates for the turbine blade under test

Case	Parameter	Description	Values [d]	
1	$r\hat{u}l_i$	Estimate of the expected value of *RUL*$_i$	1085	247
	$\sigma_{r\hat{u}l_i}$	Estimate of the prediction error standard deviation	107	45
2	$r\hat{u}l_i$	Estimate of the expected value of *RUL*$_i$	1075	248
	$\sigma_{r\hat{u}l_i}$	Estimate of the prediction error standard deviation	238	57

Notice that the $r\hat{u}l_i$ predictions of the two approaches provide satisfactory estimates of $\mu_{RUL_i|\varepsilon_i}$, whereas in both cases the prediction error variances $\hat{\sigma}^2_{r\hat{u}l_i}$ is larger than σ_A^2. This is due to the fact that, as explained in Section 2.2, $\hat{\sigma}^2_{r\hat{u}l_i}$ takes into consideration both the uncertainty σ_A^2 due to the future stochastic evolution of the test trajectory, the uncertainty σ_B^2 due to the regression error of the prognostic model, and the uncertainty σ_C^2 due to noise on the input data. It is interesting to observe that an analyst who has to decide the maintenance policy to be applied to the turbine blade would like to have the least uncertain prediction of the RUL. Thus, in the case in which the analyst were in the position to choose one of the two prognostic approaches, he/she would prefer the one whose prediction error variance is smaller.
In correspondence of each prediction $r\hat{u}l_i$, it is also possible to estimate the prediction interval $[C_i^{\text{inf}}(\alpha); C_i^{\text{sup}}(\alpha)]$, i.e., the interval expected to contain the true RUL value rul_i with a probability of $1-\alpha$. According to the two approaches, this interval can be obtained as follows:

- In case 1, $C_i^{\text{inf}}(\alpha)$ and $C_i^{\text{sup}}(\alpha)$ are the $\alpha/2$ and $1-\alpha/2$ percentiles, respectively, of the RUL distribution estimated with the particle filtering method.
- In case 2, assuming that the prediction error has a Gaussian distribution, the value of $C_i^{\text{inf}}(\alpha)$ and $C_i^{\text{sup}}(\alpha)$ can be computed as (Baraldi et al., 2013):

$$C_i^{\text{inf/sup}}(\alpha) = r\hat{u}l_i \pm c^{\alpha}_{conf}\hat{\sigma}_{r\hat{u}l_i} \quad (4)$$

where c^{α}_{conf} is the $1-\alpha/2$ percentile of a Student's t-distribution with number of degrees of freedom equal to the number $H = 25$ of bootstrap models.
Figure 1 shows the evolution of the true value of the blade RUL (continuous line), its estimated value $r\hat{u}l_i$ (dots) and the corresponding prediction interval for $\alpha = 0.32$ (dashed line) obtained during the turbine blade life at times t_i, $i = 1, \ldots, 87$ for case 1 and $i = 30, \ldots, 87$ for case 2. Notice that in the latter case the

prediction intervals are characterized by large oscillations and low accuracy, especially at the beginning of the trajectory, i.e., when few training data are available. Furthermore, the RUL prediction itself is very noisy.

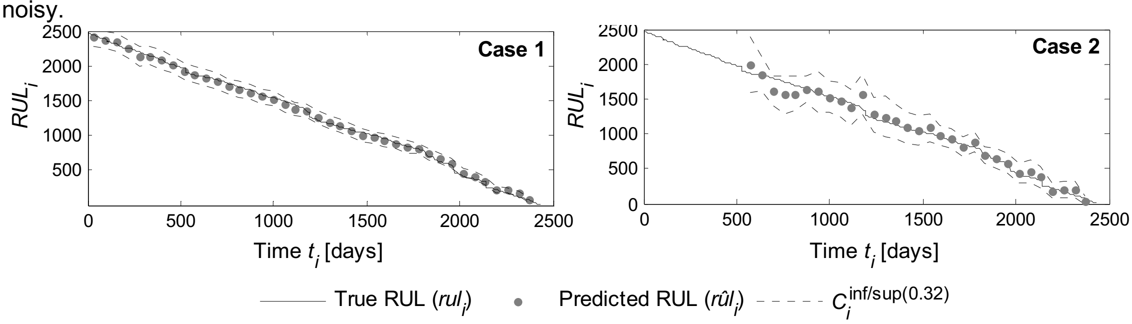

Figure1. true RUL (continuous line) of a turbine blade with its predicted value $r\hat{u}l_i$ (dots) and prediction interval (dashed line) for case 1 (left) and case 2 (right)

5. Conclusions

In this work, we have considered two practical situations with decreasing amount of information available, concerning the prognosis of the RUL of a creeping turbine blade: in the first case the model of the creep growth process is available, in the second case the model is not available but can be empirically derived from a number of direct measurements of the creep strain reached during the life of the blade.
Guidelines about how the prognostic problem should be tackled have been given for both cases, and a bootstrap ensemble-based technique has been proposed and developed to estimate the uncertainty of the RUL prediction in those situations where a priori knowledge of the mechanisms and models of the degradation process is missing (case 2).
The results obtained in the prognostics of turbine blades show that both the particle filter and the bootstrap ensemble methods provide a reliable prediction of the system RUL with a quantification of its uncertainty, although the particle filter provides less uncertain predictions.
With respect to the ensemble of bootstrapped models trained using only a sequence of direct creep strain measurements for the blade of interest, it has been observed that the approach requires building an empirical model for the estimate of the prediction variance which is then used outside the region covered by the training data. Although good extrapolations have been obtained in the linear creep growth case study, the feasibility of the approach on more complex models should be verified.
Finally, future research will consider other possible sources of uncertainty in the information available such as the imperfect knowledge of the degradation model or the uncertain definition of the failure threshold value.

Acknowledgement

The work of Francesca Mangili has been supported by a PhD grant of the Institutt For Energiteknikk (IFE), OECD Halden Reactor Project.

References

Baraldi P., Cadini F., Mangili F., Zio E., 2013. Model-Based and Data-Driven prognostics under different available information. Probabilistic Engineering Mechanics, 32, pp. 66-79.

Baraldi P., Mangili F., Zio E., 2013. Uncertainty treatment in model-based and data-driven prognostics. Reliability Engineering & System Safety, 112C, pp. 94-108.

Baraldi, P., Di Maio, F., Zio, E., Sauco, S., Droguett, E., Magno, C., 2012. Sensitivity analysis of scale deposition on equipment of oil wells plants. Chemical Engineering Transactions, 26, pp. 327-332.

Brotherton T., Jahns G., Jacobs J., Wroblewski D., 2000. Prognosis of faults in gas turbine engines. IEEE Aerospace Conference Proceedings. March 18-25, Big Sky, MT.

Cadini F., Zio E., Avram D., 2009. Monte Carlo-based filtering for fatigue crack growth estimation. Probabilistic Engineering Mechanics, 24, pp. 367-373.

Penny R. K., Marriott D. L., 1995. Design for Creep, 2nd edition. London: Chapman & Hall.

Saez M., Tauveron N., Chataing T., Geffraye G., Briottet L., Alborghetti N., 2006. Analysis of the turbine deblading in an HTGR with the CATHARE code. Nuclear Engineering and Design, vol. 236, pp. 574–586.