

Remaining Useful Life Prediction Using Elliptical Basis Function Network and Markov Chain

Yi Yu, Lin Ma, Yong Sun, Yuantong Gu

Abstract—This paper presents a novel method for remaining useful life prediction using the Elliptical Basis Function (EBF) network and a Markov chain. The EBF structure is trained by a modified Expectation-Maximization (EM) algorithm in order to take into account the missing covariate set. No explicit extrapolation is needed for internal covariates while a Markov chain is constructed to represent the evolution of external covariates in the study. The estimated external and the unknown internal covariates constitute an incomplete covariate set which are then used and analyzed by the EBF network to provide survival information of the asset. It is shown in the case study that the method slightly underestimates the remaining useful life of an asset which is a desirable result for early maintenance decision and resource planning.

Keywords—Elliptical Basis Function Network, Markov Chain, Missing Covariates, Remaining Useful Life

I. INTRODUCTION

MODERN Condition Based Maintenance (CBM) requires up to date knowledge about the health status of an engineering asset under specific operating conditions. It is desirable that the asset's lifetime information can also be predicted. Hazard models [1-5] provide detailed information about the current and historical asset health as well as the imminent risk of asset failure. However, hazard models themselves generally cannot predict the lifetime health status of an asset because of the lack of future covariates. Ancillary methods such as polynomial fitting [1] are often needed for covariate extrapolation. Polynomial fitting can produce satisfactory results in certain situations. However, valuable information could be lost where variation of covariates is

Y. Yu is with the Cooperative Research Centre for Infrastructure and Engineering Asset Management (CIEAM) and School of Engineering Systems, Queensland University of Technology, Brisbane, Australia (phone: +61 7 3138 9158; fax: +61 7 3138 1469; e-mail: y.yu@qut.edu.au).

L. Ma is with the Cooperative Research Centre for Infrastructure and Engineering Asset Management (CIEAM) and School of Engineering Systems, Queensland University of Technology, Brisbane, Australia (e-mail: l.ma@qut.edu.au).

Y. Sun is with the Cooperative Research Centre for Infrastructure and Engineering Asset Management (CIEAM) and School of Engineering Systems, Queensland University of Technology, Brisbane, Australia (e-mail: y3.sun@qut.edu.au).

Y. Gu is with the Cooperative Research Centre for Infrastructure and Engineering Asset Management (CIEAM) and School of Engineering Systems, Queensland University of Technology, Brisbane, Australia (e-mail: yuantong.gu@qut.edu.au).

smoothed out in the polynomial fitting. Alternatively, a non-linear model such as neural network can be employed to extrapolate the covariates into the future [6, 7]. This method, however, is often limited due to the error accumulated in the process.

Covariates, in general, could be categorized into two classes: external covariates and internal covariates [8]. External covariates, including environmental settings, stress and load, influence but are not directly related to the asset failure mechanism. Internal covariates reflect and are often generated by the asset failure mechanism, e.g. diagnostic factors or degradation measurements. Internal covariates are usually time dependent and have a large scale of variations. It is thus often difficult to model this type of data. On the contrary, external covariates in practical environments could be constant or vary between a finite number of states or values during an asset operating period. This kind of data, therefore, can be described and modeled with reasonable accuracy. Furthermore, internal covariates are generally influenced by or correlated with external covariates. Hence it would be easier to predict internal covariates when information on external covariates is available.

If future covariates are treated as missing data rather than unknown upcoming variables, they could be estimated by missing covariate handling approaches within the survival analysis procedure. In this paper, such an approach is utilized for modeling future survival status without explicitly extrapolating all future covariates. A Markov chain is constructed to represent the fluctuation of external covariates. The future internal covariates are regarded as missing values and would be estimated implicitly by the EBF network model.

The remainder of this paper is organized as follows: Section 2 briefly reviews the hazard models and proposes a non-linear hazard model based on the EBF network. This EBF hazard model is further modified to consider the missing covariate problem. Section 3 establishes a Markov chain model to represent the external covariate evolution. The constructed Markov chain is utilized together with the EBF hazard model in Section 4 to predict multi-step survival information. A case study is presented in Section 5 with data obtained from a typical engineering system.

II. MODIFIED EBF NETWORK FOR HAZARD PREDICTION WITH MISSING COVARIATES

Several modeling approaches are available for survival

analysis. The most commonly used approaches are parametric and nonparametric statistical methods. The parametric statistical approach [1-3] assumes specified families of failure time distributions. This assumption, however, simplifies the failure behavior and limits the application of the method. Non-parametric models [3-5] relax these assumptions, but require different ones such as the linear influence of system features on asset failure behavior. The assumption of linear feature's influence on failure can also be relaxed when a neural network is employed instead of statistical survival models because of its highly non-linear characteristics [9]. Moreover, a neural network could provide personalized survival predictions which are valuable for subsequent maintenances and the whole engineering asset management process.

In our previous work [10], a Multi-Layer Perceptron (MLP) neural network was adopted to model the relationship between hazard and covariates. To prevent the over-fitting problem of MLP network, Bayesian regularization is employed in the study. The EBF network is a useful alternative to the MLP structure, which has good generalization ability and easy to train [9, 11]. Another reason for adopting the EBF network is the possibility to modify its training session to handle the missing covariate problem. This is described in detail below.

A. EBF structure for hazard prediction

An EBF network is the extension of Radial Basis Function (RBF) networks. It uses a full covariance matrix instead of the spherical or diagonal covariance matrix as in RBF structure. It has been demonstrated that a smaller size EBF network could outperform a large size RBF network [12]. The EBF network adopted for hazard prediction in this paper can be described as:

$$h_k = \beta_o + \sum_{j=1}^M w_{oj} \phi_j(x_{ik}) \quad (1)$$

The output of this EBF network is hazard h_k at the time interval k . x_{ik} represents the covariates (including a time variable).

$$\phi_j = (2\pi)^{-d/2} |\Sigma_j|^{-1/2} \cdot \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{x} - \boldsymbol{\mu}_j)\right) \quad (2)$$

Equation (2) is a basis function which has a mean vector $\boldsymbol{\mu}_j$ and a full covariance matrix Σ_j .

The parameters of the hidden layer $(\boldsymbol{\mu}, \Sigma)$ and the output weight parameters (β_o, w_o) in this structure can be determined in two separate steps without the non-linear optimization procedure. This thus results in a faster training process. The sum of basis functions is usually chosen to represent the unconditional density of the input data. Therefore, determination of the mean and covariance of the basis function becomes essentially the Gaussian mixture density estimation problem [9]. The mixture density estimation is typically carried out by using the well-known Expectation-Maximization (EM)

algorithm [13]. The application of EM algorithm in EBF network training indicates that parameters of basis functions are determined in an iterative fashion.

At each iteration, the EM algorithm increases the log-likelihood as:

$$L(\theta | x) = \sum_{k=1}^N \sum_{j=1}^M z_j^k \left[\frac{k}{2} \log 2\pi + \frac{1}{2} \log |\Sigma_j| - \frac{1}{2} (\mathbf{x}_k - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{x}_k - \boldsymbol{\mu}_j) \right] \quad (3)$$

where z_j^k represents the probability that data \mathbf{x}_k is generated by a mixture density j .

By maximizing (3), the parameters are shifted closer to the optimal estimates according to the following equation [9]:

$$\boldsymbol{\mu}_j^{new} = \frac{\sum_k p(j | \mathbf{x}_k) \mathbf{x}_k}{\sum_k p(j | \mathbf{x}_k)} \quad (4)$$

$$\Sigma_j^{new} = \frac{\sum_k p(j | \mathbf{x}_k) (\mathbf{x}_k - \boldsymbol{\mu}_j^{new})(\mathbf{x}_k - \boldsymbol{\mu}_j^{new})^T}{\sum_k p(j | \mathbf{x}_k)} \quad (5)$$

The weighting part of the above equations are the posterior probability of the j -th basis which can be obtained from the Baye's theorem [14].

If we consider an extra basis function ϕ_0 with the fixed activation value of 1, the bias in (1) could then be absorbed into the weights. As a result, the hazard in (1) can be described in a matrix form, $\mathbf{h} = \boldsymbol{\omega} \Phi$, in which $\boldsymbol{\omega} = (\beta_o, w_o)$ and $\Phi = (\phi_j, \phi_0)$. Once the basis function is determined, the output weights $\boldsymbol{\omega}$ can be optimized through the minimization of a quadratic error function using the pseudo-inverse of the basis activation matrix.

B. Modification of EBF network's training process to consider the missing covariate problem

In engineering applications, covariates are often found to be incomplete or missing due to sensor failure, multiple measurement scales as well as the storage limit. Several missing covariates handling approaches have been studied and evaluated by [10]. Gaussian Mixture Model (GMM) was found to have better performance than the others. As discussed previously, determination of EBF basis function is equivalent to the estimation of mixture density models. Therefore, the procedure of GMM method in dealing with missing covariates can be employed to train EBF network for the missing covariate problem.

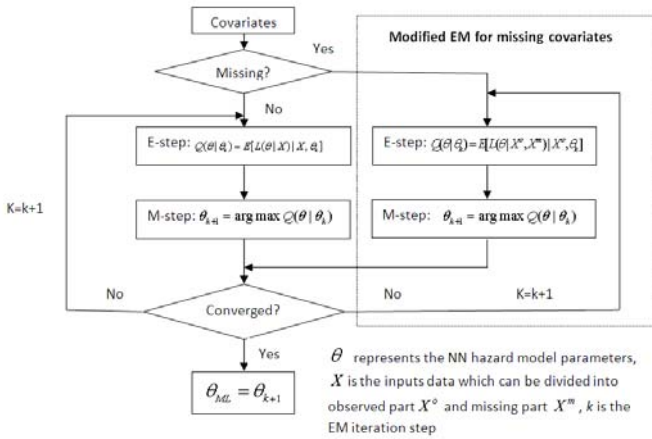


Fig. 1 Modified EM algorithm for incorporation of missing covariates

The modification of EM algorithm for EBF network training is illustrated in Fig. 1. The covariates \mathbf{x}_k are divided into two parts: the observed part \mathbf{x}_k^o and the missing part \mathbf{x}_k^m . Correspondingly, the mean and the inverse covariance matrix of density component j are also divided [15]:

$$\boldsymbol{\mu}_j = (\boldsymbol{\mu}_j^m, \boldsymbol{\mu}_j^o) \quad (6)$$

and

$$\boldsymbol{\Sigma}_j^{-1} = \begin{bmatrix} \boldsymbol{\Sigma}_j^{-1,mm} & \boldsymbol{\Sigma}_j^{-1,mo} \\ \boldsymbol{\Sigma}_j^{-1,om} & \boldsymbol{\Sigma}_j^{-1,oo} \end{bmatrix} \quad (7)$$

The likelihood function in (3) is rewritten according to the separation of observed and missing covariates as:

$$L(\theta | \mathbf{x}^o, \mathbf{x}^m) = \sum_{k=1}^N \sum_{j=1}^M \left\{ z_{kj}^k \left[\frac{k}{2} \log 2\pi + \frac{1}{2} \log |\boldsymbol{\Sigma}_j| \right] - \frac{1}{2} (\mathbf{x}_k^o - \boldsymbol{\mu}_j^o)^T \boldsymbol{\Sigma}_j^{-1,oo} (\mathbf{x}_k^o - \boldsymbol{\mu}_j^o) - \frac{1}{2} (\mathbf{x}_k^o - \boldsymbol{\mu}_j^o)^T \boldsymbol{\Sigma}_j^{-1,om} (\mathbf{x}_k^m - \boldsymbol{\mu}_j^m) - \frac{1}{2} (\mathbf{x}_k^m - \boldsymbol{\mu}_j^m)^T \boldsymbol{\Sigma}_j^{-1,mm} (\mathbf{x}_k^m - \boldsymbol{\mu}_j^m) \right\} \quad (8)$$

On each E-step of the EM algorithm, the expectation of likelihood is calculated conditional on the observed covariates rather than the complete covariates.

The likelihood in (8) is linear in three values, namely, z_{kj} , $z_{kj} \mathbf{x}_k^m$ and $z_{kj} \mathbf{x}_k^m (\mathbf{x}_k^m)^T$. Given the observed data \mathbf{x}_k^o , the conditional expectation of these values are [15]:

$$E(z_{kj} | \mathbf{x}_k^o, \boldsymbol{\mu}_j^{old}, \boldsymbol{\Sigma}_j^{old}) = p^{old}(j | \mathbf{x}_k^o) \quad (9)$$

$$E(z_{kj} \mathbf{x}_k^m | \mathbf{x}_k^o, \boldsymbol{\mu}_j^{old}, \boldsymbol{\Sigma}_j^{old}) = p^{old}(j | \mathbf{x}_k^o) \hat{\mathbf{x}}_{kj}^m \quad (10)$$

and

$$E(z_{kj} \mathbf{x}_k^m (\mathbf{x}_k^m)^T | \mathbf{x}_k^o, \boldsymbol{\mu}_j^{old}, \boldsymbol{\Sigma}_j^{old}) = p^{old}(j | \mathbf{x}_k^o) \cdot [\boldsymbol{\Sigma}_j^{mmm} - \boldsymbol{\Sigma}_j^{mo} (\boldsymbol{\Sigma}_j^{oo})^{-1} (\boldsymbol{\Sigma}_j^{mo})^T + \hat{\mathbf{x}}_{kj}^m (\hat{\mathbf{x}}_{kj}^m)^T] \quad (11)$$

$\hat{\mathbf{x}}_{kj}^m$ is the least square linear regression between \mathbf{x}_k^m and \mathbf{x}_k^o for $z_{kj} = 1$:

$$\hat{\mathbf{x}}_{kj}^m = E(\mathbf{x}_k^m | \mathbf{x}_k^o, z_{kj} = 1, \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) = \boldsymbol{\mu}_j^m + \boldsymbol{\Sigma}_j^{mo} (\boldsymbol{\Sigma}_j^{oo})^{-1} (\mathbf{x}_k^o - \boldsymbol{\mu}_j^o) \quad (12)$$

Substituting these conditional expectations into (4) and (5), parameters of the EBF basis function can be determined.

III. MARKOV CHAIN FOR EXTERNAL COVARIATE EXTRAPOLATION

Internal covariates such as degradation observations reflect the instantaneous asset health. This data is highly influenced by external covariates of the asset and has a large scale of variation, especially when the asset is under abnormal health condition. Although this information is valuable, it often imposes difficulties for prediction or extrapolation. Fitting them with standard statistical functions such as polynomial would smooth out the useful variation information. Using non-linear extrapolation model is subject to large error accumulation. Stochastic approaches are also proposed to model the evolution of internal covariates [16]. The application of the method, however, requires an explicit definition of the failure threshold which is usually unknown or vaguely defined in survival analysis. Complicated numerical simulation might also be involved if these covariates are integrated into survival models.

Unlike internal covariates, external covariates reflect the influence of external variables to the degradation of an engineering asset. The values of external covariates such as stress or load do not vary as dramatically as interval covariates. They typically switch between several finite states or values. Modeling external covariates is much easier than modeling internal covariates in general. A Markov chain can be employed to model the evolution of external covariates and is thus utilized in this study. Elements of the transition probability matrix of the Markov process are calculated by:

$$\hat{P}_{ij}(k) = \frac{n_{ij}(k)}{\sum_j n_{ij}(k)} \quad (13)$$

in which $n_{ij}(k)$ is the number of one step transitions from state or value i towards j within period k . When covariate values are missing in certain steps, the transition probability can still be established by following methods presented in [17] and [18].

IV. REMAINING USEFUL LIFE PREDICTION USING THE EBF NETWORK AND THE DISCRETE MARKOV PROCESS

We can view the future unknown internal covariates from another way. If external covariates are available, either

pre-specified or estimated, the whole covariate set would be an incomplete one. Therefore, this will pose a missing covariate problem and can be solved by using the missing covariate handling approaches [10]. Recalling the EBF structure for hazard prediction in Section 2, the prediction of hazard is actually a classification task with two classes: in current time interval the asset will fail (class C_f) or not fail (class C_s) given the asset has survived the operation history. Therefore, hazard at k -interval is $h_k = \Pr(C_f | \mathbf{x}_k)$. Since the EBF network is trained by using the modified EM algorithm to consider the missing covariates problem, when \mathbf{x}_k involves missing values, the hazard is predicted as follows [19]:

$$\Pr(C_f | \mathbf{x}_k^o) = \frac{\sum_{j=1}^M w_{oj} \phi_j(\mathbf{x}_k^o) p(j)}{\sum_{j=1}^M \phi_j(\mathbf{x}_k^o) p(j)} \quad (14)$$

Combining the data set of unknown internal and estimated external covariates, the EBF hazard model performs the survival analysis for future period without explicitly estimating the internal covariates.

With the predicted hazard, remaining useful life of an asset at time i (m_i) is calculated as:

$$m_i = \sum_{j=i+1}^{\infty} \prod_{k < j} (1 - h_k) / \prod_{k < i} (1 - h_k) \quad (15)$$

V. CASE STUDY

Covariates data from a typical engineering system [20] are employed to demonstrate the proposed method. These data were recorded until failure occurs for each asset unit. They varied from case to case because of changes in environmental settings and asset health deterioration.

After examining the external covariates, it is found that the external covariates switch only between six discrete states. The transition matrix for the constructed Markov chain is given as follows according to (13):

$$A = \begin{pmatrix} 0.1475 & 0.1449 & 0.2626 & 0.1487 & 0.1459 & 0.1503 \\ 0.1417 & 0.1524 & 0.2530 & 0.1492 & 0.1449 & 0.1588 \\ 0.1495 & 0.1417 & 0.2555 & 0.1537 & 0.1486 & 0.1509 \\ 0.1497 & 0.1499 & 0.2554 & 0.1456 & 0.1515 & 0.1478 \\ 0.1472 & 0.1506 & 0.2422 & 0.1477 & 0.1538 & 0.1586 \\ 0.1477 & 0.1649 & 0.2413 & 0.1483 & 0.1552 & 0.1427 \end{pmatrix}$$

The external covariates after each prediction point are estimated based on the transition matrix and its current state. The estimated values then form an incomplete covariate set with future internal covariates missing. This incomplete set is then presented into the EBF hazard model trained by the modified EM algorithm. The calculated remaining useful life for seven testing units is illustrated in Fig. 2.

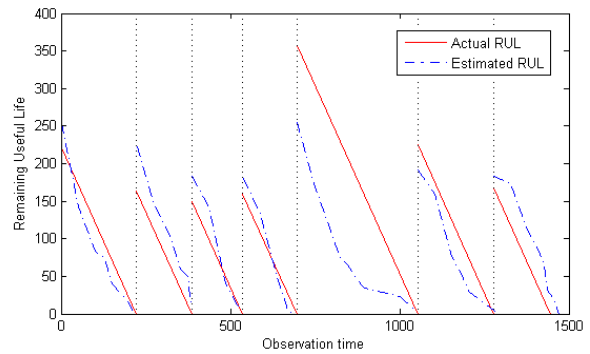


Fig. 2 Remaining useful life for seven units

It is shown in Fig. 2 that the estimated remaining useful life values follow the actual values except for the fifth unit. This is because that unit has a much longer life than the others. The EBF network tries to learn the general survival information rather than over-fit to this particular unit. It is observed that the remaining useful life predicted by the proposed method tends to underestimate the actual residual life. The result is preferable in practical situations since a conservative prediction is better than an overestimated one, especially in critical situations.

VI. CONCLUSION

This paper presents a novel method for remaining useful life prediction. The EM algorithm for EBF network training is modified to enable the EBF network to take into account of missing covariates for survival analysis. By constructing a Markov chain to model external covariate evolution and by treating internal covariates as missing values, extrapolation of all covariates and definition of the failure threshold are no longer required. The case study has demonstrated that the remaining useful life predicted by the proposed method closely follows the actual one, and the estimated values are in general underestimated. This conservative estimate enables maintenance decisions and scheduling to be carried out earlier, which is desirable to minimize the down time of an engineering asset.

ACKNOWLEDGMENT

This paper was developed within the CRC for Infrastructure and Engineering Asset Management, established and supported under the Australian Government's Cooperative Research Centre Program.

REFERENCES

- [1] H. Liao, W. Zhao, and H. Guo "Predicting remaining useful life of an individual unit using proportional hazards model and logistic regression model," in *2006 Annual Reliability and Maintainability Symposium. RAMS '06.*, pp. 127-132.
- [2] D. Kumar, B. Westberg, "Some reliability models for analysing the effects of operating conditions," *International Journal of Reliability, Quality and Safety Engineering*, vol. 4, pp. 133-148, 1997.
- [3] D. R. Cox, "Regression models and life tables (with discussion)," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 34, no. 2, pp. 187-220, 1972.
- [4] D. R. Cox, D. Oakes, *Analysis of survival data*. Chapman & Hall/CRC, 1984.

- [5] Y. Sun, L. Ma, J. Mathew, W. Wang, and S. Zhang, "Mechanical systems hazard estimation using condition monitoring," *Mechanical Systems and Signal Processing*, vol. 20, no. 5, pp. 1189-1201, 2006.
- [6] Y. Shao, K. Nezu, "Prognosis of remaining bearing life using neural networks," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 214, no. 3, pp. 217-230, 2000.
- [7] P. W. Tse, D. P. Atherton, "Prediction of machine deterioration using vibration based fault trends and recurrent neural networks," *Journal of Vibration and Acoustics*, vol. 121, no. 3, pp. 355-343, 1999.
- [8] J. D. Kalbeisch, R. L. Prentice, *The statistical analysis of failure time data*. New York: Wiley, 1980.
- [9] C. M. Bishop, *Neural networks for pattern recognition*. Oxford Univ Pr, 2005.
- [10] Y. Yu, L. Ma, Y. Sun, and Y. Gu "Handling Incomplete Data In Survival Analysis With Multiple Covariates," in *2010 Proceedings of the 5rd World Congress on Engineering Asset Management and Intelligent Maintenance Systems.*, to be published.
- [11] M. T. Musavi, W. Ahmed, K. H. Chan, K. B. Faris, and D. M. Hummels, "On the training of radial basis function classifiers," *Neural Networks*, vol. 5, no. 4, pp. 595-603, 1992.
- [12] M. W. Mak, C.K. Li, "Elliptical basis function networks and radial basis function networks for speaker verification: A comparative study," in *1999 International Joint Conference on Neural Networks (IJCNN'99)*, pp. 3034-3039.
- [13] A. P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1-38, 1977.
- [14] G.J. McLachlan, T. Krishnan, *The EM algorithm and extensions*. New York: Wiley, 1997.
- [15] Z. Ghahramani, M. I. Jordan, "Supervised learning from incomplete data via an EM approach," in *Advances in neural information processing systems 6*, J. D. Cowan, G. Tesauro, and J. Alspector, Ed. Morgan Kaufmann, 1995, pp. 120-127.
- [16] W. Wang, "A model to predict the residual life of rolling element bearings given monitored condition information to date," *IMA Journal of Management Mathematics*, vol. 13, no. 1, pp. 3-16, 2002.
- [17] B. Craig, P. Sendi, "Estimation of the transition matrix of a discrete-time Markov chain," *Health Economics*, vol. 11, no. 1, pp. 33-42, 2002.
- [18] H. Yeh, W. Chan, E. Symanski, and B. Davis, "Estimating Transition Probabilities for Ignorable Intermittent Missing Data in a Discrete-Time Markov Chain," *Communications in Statistics-Simulation and Computation*, vol. 39, no. 2, pp. 433-448, 2010.
- [19] R. Dybowski, "Classification of incomplete feature vectors by radial basis function networks," *Pattern Recognition Letters*, vol. 19, no. 14, pp. 1257-1264, 1998.
- [20] PHM, Available from:
<http://www.phmconf.org/jOCS/index.php/phm/2008/challenge.2008>.