Estimating Gas Turbine Compressor Discharge Temperature using Bayesian Neuro-fuzzy Modelling

Yu Zhang, Miguel Martinez-Garcia
School of Engineering, University of Lincoln
Lincoln, U.K.
{yzhang; mmartinez}@lincoln.ac.uk

Anthony Latimer
Siemens Industrial Turbomachinery LTD
Lincoln, U.K.
anthony.latimer@siemens.com

Abstract — **The objective of this paper is to estimate the compressor discharge temperature measurements on an industrial gas turbine that is undergoing commissioning at site, using a data-driven model which is built using the test bed measurements of the engine. This paper proposes a Bayesian neuro-fuzzy modelling (BNFM) approach, which combines the adaptive neuro-fuzzy inference system (ANFIS) and variational Bayesian Gaussian mixture model (VBGMM) techniques. A data-driven compressor model is built using ANFIS, and VBGMM is applied in the set-up stage to automatically select the number of input membership functions in the fuzzy system. The efficacy of the proposed BFNM approach is established through experimental trials of a sub-15MW gas turbine, and the results, from the model that is built using test bed data, are shown to be promising for estimating the compressor discharge temperatures on the gas turbine during commissioning.**

Keywords— Bayesian neuro-fuzzy modeling; adaptive neuro-fuzzy inference system; variational Bayesian Gaussian mixture model; compressor discharge temperature; industrial gas turbine.

I. INTRODUCTION

Predictive modelling is of great importance in the field of condition monitoring, fault diagnosis and measurement estimation for industrial gas turbines (IGTs) [1-3]. This study focusses on the development of a model (for a compressor, in this case) with IGT test bed data, which is then used to estimate an unknown measurement (compressor discharge temperature, in this case) on the engine during commissioning.

Pure dynamic models of IGTs are commonly difficult to achieve due to the complexity in the structure, the auxiliaries and the control system. Panov [4] has developed a Simulink model, which can be used to simulate start-up operation, change of load, control system, power-system stabilities and real-time modelling of IGTs. However, the modelled IGT behaviours rely heavily on precise component maps delivered by specific engine tests. On the other hand, this paper aims to build an IGT model with data-driven techniques, i.e. in the absence of the specifically pre-defined component maps, which are normally incomplete due to the restrictions of the testing facilities, environments and control limits, etc.

In contrast to the physical models (white-box models), black-box modelling approaches, such as the use of artificial neural networks (ANNs), have been broadly applied because of their high learning abilities and the non-linear and non-parametric properties [5-7]. It has been shown with some success by ANNs, however, these models do not provide insights on the dynamical or physical properties of the system, and consequently are not adequate at dealing with uncertainties [8]. Alternatively, fuzzy inference systems (FISs) can use fuzzy rules to display and interpret the relationship between the inputs and outputs of the system, which are then more effective with uncertainties. As a result, a hybridization of ANN and FIS is sought, with a noticeable example being an adaptive neuro-fuzzy inference system (ANFIS) [9]. ANFIS builds a FIS with a collection of fuzzy rules, and an ANN is used in each rule for parameter tuning. This takes the advantages of both FIS and ANN, e.g. with the former providing interpretation of the system inputs and outputs, and with the latter for non-parametric estimation in non-linear systems. ANFIS has been a popular tool since its creation. For instance, Salahshoor et al. [10] have used three ANFIS classifiers, based on the most dominant parameters, for data-driven fault detection on an industrial steam turbine. Chen et al. [11] have proposed a-priori knowledge-based ANFIS for automated flagging of significant pitch faults on wind turbines. Moreover, Zhang et al. [12] have applied ANFIS to extract start-up vibration signatures for novelty detection on IGTs.

In this article, ANFIS is employed to model the thermo-dynamic behaviours of a compressor on an IGT based on the engine test data, which is then utilised to estimate the compressor discharge temperatures on this engine at site. Still, ANFIS needs a pre-defined number of input membership functions (MFs) or rules – less rules result in loss of precision, whilst too many rules cause redundancy and degradation of interpretability. In this sense, Bayesian method is a popular solution, when one aims to "take the human out of the loop" [13]. For example, a variational Bayesian Gaussian mixture model (VBGMM) has been applied to discriminate the operational data of IGTs into steady-state and transient responses automatically [14].

The main contribution of the methodology in this paper, compared with the original ANFIS is that, VBGMM is applied for clustering the initial rules of the FIS automatically, so that the adequate number of input MFs can be selected, which then provides an adequate trade-off between the model specificity and sensitivity. The proposed Bayesian neuro-fuzzy modelling (BNFM) approach uses VBGMM to automatically select the number of input MFs, and ANFIS to set up the FIS and tune parameters using ANN. The efficacy of the BNFM is then demonstrated from the experimental trials of an IGT using both test bed data and field data.

II. METHODOLOGY

In Table 1, a concise description of the methodology is presented. An overview of the supporting principles for the used approach is explained in the following.

TABLE I. METHODOLOGY FRAMEWORK

Phases	Inputs	Methods	Outcomes
Set up FIS	Test bed data	VBGMM	Automatically selected cluster number
Build BNFM	Cluster number, Gaussian MFs for inputs	ANFIS	Compressor model built from test bed data
Measurement estimation	Field input measurements	Built BNFM	Output sensor estimation based on field data

A. Adaptive Neuro-fuzzy Inference System

ANFIS is a potent tool for predictive modelling, which combines the benefits from both FIS and ANN [9]. ANFIS applies fuzzy rules, where the parameters related to the MFs are estimated using ANN – usually with a backpropagation learning algorithm from input and output data.

FIS utilises MFs to describe input characteristics and applies rules to map the input MFs to the output characteristics. Here, ANFIS uses a Sugeno-type FIS, where a typical rule is: If Input 1 is x_1 and Input 2 is x_2, then Output is $y = ax_1 + bx_2 + c$, and the final output is the weighted average of the outputs from all rules [16]. Similar to ANNs, an ANFIS provides MF parameter training to choose the best associated FIS for the given input/output data [9,15].

Fig. 1 shows a typical ANFIS structure which consists of 2 inputs, 2 rules and 5 layers – Layer 1: fuzzy layer, Layer 2: product layer, Layer 3: normalisation layer, Layer 4: defuzzify layer and Layer 5: output layer. There are 2 adaption layers (Layer 1 and Layer 4), where Layer 1 has tuneable premise parameters related to the input MFs, and Layer 4 has tuneable parameters related to the consequence part of the fuzzy rules. The objective is to optimise all these parameters to best match the train data. [11]
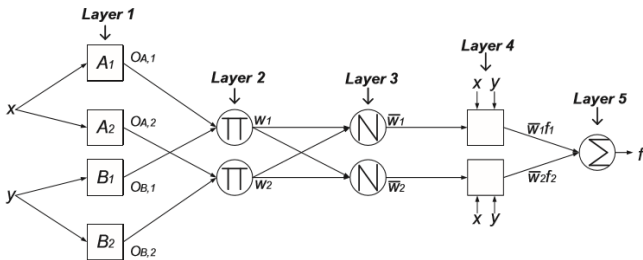

Fig. 1. Typical ANFIS structure [11]

B. Variational Bayesian Gaussian Mixture Model

A Gaussian mixture model (GMM) uses a linear combination of Gaussian distributions to estimate the probability density function of the sample data [17]. In this manner, the probabilistic distribution of sample data can be written as a sum of K variables normally distributed, with corresponding mean and standard deviation μ_k and σ_k respectively. A GMM including K mixture components is then expressed as

$$p(x \mid \theta, \mu, \sigma) = \sum_{k=1}^{K} \theta_k N(x \mid \mu_k, \sigma_k), \qquad (1)$$

where θ_k are the mixing coefficients, and $N(x \mid \mu, \sigma)$ denotes a normal probability density function of a multi-dimensional variable x, with mean $\mu \equiv \{\mu_k\}$ and standard deviation $\sigma \equiv \{\sigma_k\}$.

A variational Bayesian (VB) technique can be applied for selecting the necessary number of the mixture components. A set of binary latent variables $z_{nk} \in \{0,1\}$ is introduced, which specifies the associations among the given N data points and the K mixture components. The matrix Z identifies which data sample x_n belongs to which mixture component k.

The joint probability distribution for all the variables of the GMM combined can be written as

$$p(X, Z, \theta, \mu, \sigma) = p(X \mid Z, \mu, \sigma) p(Z \mid \theta) p(\theta) p(\mu \mid \sigma) p(\sigma) \ (2)$$

where $p(\theta)$, $p(\mu \mid \sigma)$, and $p(\sigma)$ are Dirichlet, Gaussian and Wishart probability distributions separately [18].

By using a VB technique, a lower bound on $p(X \mid \theta) = p(X \mid Z, \mu, \sigma) p(Z \mid \theta)$ can be obtained. By representing $\Psi \equiv \{Z, \mu, \sigma\}$ and introducing a variational posterior distribution $q(\Psi) \approx p(\Psi \mid \theta)$, for the marginal log-likelihood $\ln p(X \mid \theta)$, the relation holds that:

$$\ln p(X \mid \theta) = D_{KL}(q \| p) + L(q), \qquad (3)$$

where $L(q)$ is the lower bound and $D_{KL}(q \| p)$ is the Kullback-Leibler divergence. The variational posterior distribution can be reflected in terms of factorization over the subsets $\{\Psi_i\} = \{Z, \mu, \sigma\}$, which writes

$$q(Z, \mu, \sigma) = q_Z(Z) q_\mu(\mu) q_\sigma(\sigma), \qquad (4)$$

The optimal distribution for each of the factors can be established through a free-form minimization over q_i:

$$q_i^*(\Psi_i) = \frac{\exp\left(\mathbb{E}_{j \neq i}\left[\ln p(X, \Psi)\right]\right)}{\int \exp\left(\mathbb{E}_{j \neq i}\left[\ln p(X, \Psi)\right]\right) d\Psi_i}, \qquad (5)$$

where $\mathbb{E}_{j \neq i}[\cdot]$ denotes the expectation of the distributions $q_j(\Psi_j)$ for all $j \neq i$. After calculating the variational factors in (4) through the use of (5), the lower bound $L(q)$ can be

evaluated. The maximization of $L(q)$ minimizes $\ln p(X \mid \theta)$, then the mixing coefficients θ can be found. And the optimal distributions q_i are achieved by iteratively updating the variational factors until meeting the convergence criterion.

For a more in-depth explanation on the application of the VB technique to GMMs, the reader is directed to Corduneanu and Bishop [18].

C. Bayesian Neuro-fuzzy Modelling

BNFM generates an initial FIS by using VBGMM to extract a set of rules that represents the data behaviour. The rule extraction method first applies the VB technique to determine the number of rules automatically, and the MFs for the premises are set to be Gaussian functions [19]. This initial FIS is then fed into the ANFIS for training, which estimates the system parameters through a neural network training algorithm which combines the use least-squares minimization and the back-propagation gradient descent methodology to match the training data set [15]. The scheme of the methodology can be directed again to Table 1.

III. CASE STUDY

System modelling of the compressor via BNFM is applied using the test bed measurements of a sub-15MW IGT unit. The measurements include T1 (ºC) – compressor inlet temperature (ambient temperature), T2 (ºC) – compressor discharge temperature, P2 (bar) – compressor discharge pressure and Speed (%) – compressor rotational speed, as shown in Fig. 2. In this case, P1 (bar) – compressor inlet pressure (ambient pressure), is not measured, which is assumed $\approx$ 1 bar. The sampling rate varies from per second to per minute, thereafter we address the data "sample points" instead.

In order to generate an interpretable surface for the compressor model, the input and output parameters are selected as, Input 1: Speed, Input 2: P2 (also presenting the pressure ratio PR = P2/P1), and Output: temperature ratio TR = T2/T1 (here, T1 and T2 are converted to degrees Kelvin according to the thermodynamic laws). Then the final estimated T2 can be simply converted back to Celsius from the output TR by

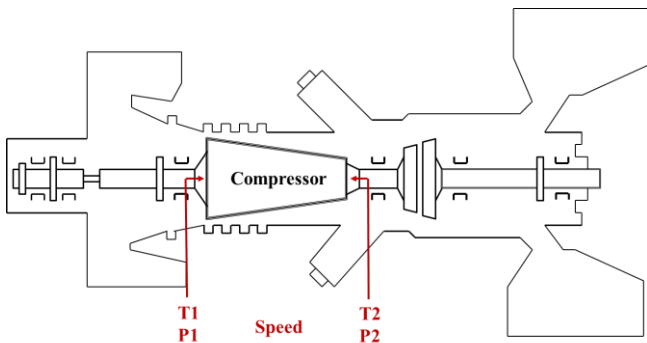$$T2 = TR \cdot (T1 + 273.15) - 273.15 \ . \tag{6}$$

Fig. 2. Simplified diagram of the components of a prototypal industrial gas turbine. The input and output sensor positions for measuring the compressor characteristics are displayed in red colour.

VBGMM is applied to the 3-D test bed data of Speed, PR and TR. In this case study, 5 clusters are identified by using the VB technique, as shown in Fig. 3. Consequently, the associated Gaussian MFs for the inputs Speed and PR are plotted in Fig. 4. It can be seen from the test bed data in Fig. 3 that, a lot of the operational regimes are not tested due to the restrictions of unit testing control limits, whilst these regimes of Speed and PR are estimated over the whole range using the Gaussian MFs as shown in Fig. 4.

Corresponding to the ANFIS structure in Fig. 1, now there are 5 MFs for each input, then the BNFM structure can be represented as shown in Fig. 5 in this case [20].
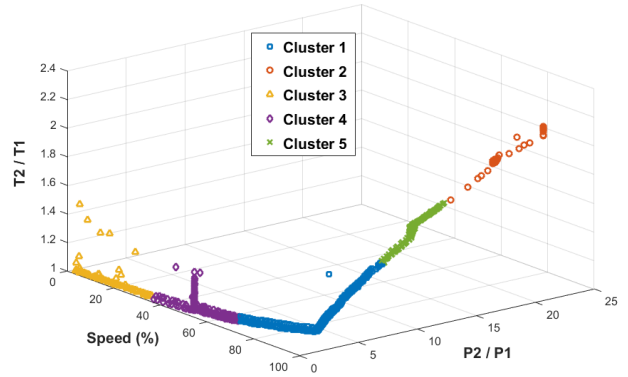
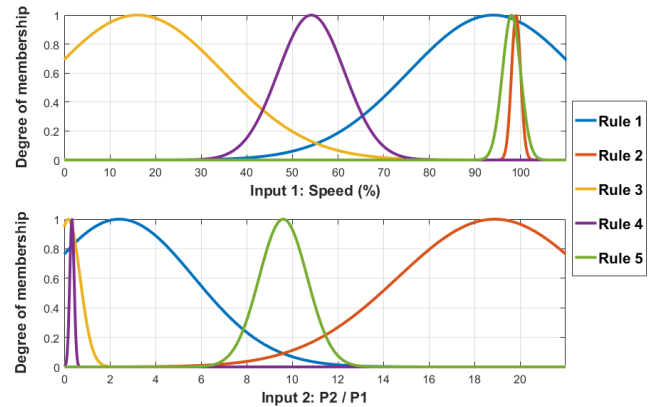Fig. 3. VBGMM clustering result of the inputs (speed and P2/P1) vs. output (T2/T1).

Fig. 4. Associated membership functions of BNFM for Input 1: Speed (%) and Input 2: P2/P1.
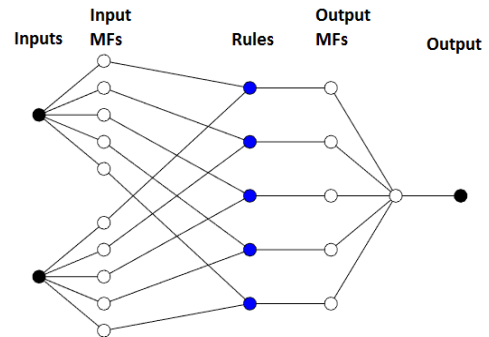
Fig. 5. BNFM / ANFIS structure for the case study

After parameter tuning of the model through ANFIS, the input-output relations can be represented through a 3-D surface as shown in Fig. 6, where the output TR is plotted in a log-scale to reduce the large value effect of the top areas. Comparing with the test bed data in Fig. 3, it is shown that the plot across the whole range of speeds and PRs is estimated by using the data-driven fuzzy modelling, although some of the regimes are infeasible to reach in a real testing situation according to the compressor performance. For instance, in the low speed vs. high PR regime, the model gives unreasonably large values of TR output due to the lack of training data in this regime. However, for normal IGT running regimes within the training data ranges, the modelled results are overall satisfactory.

The estimated T2 measurements, calculated by (6) from the TR output using BNFM based on the test bed data, are plotted in Fig. 7. Estimation of T2 based on the training (test bed) data is shown in Fig. 7 (a), compared with the real training data of T2. The built BNFM model based on the test bed data is then applied to the field input measurements of Speed and PR, and the estimated T2 measurements, as shown in Fig. 7(b), are calculated using (6) from the modelled output TR and the known field measurements of T1. The estimation measurements are compared with the real field data of T2. From the results it can be seen that, the estimated T2 measurements from BNFM based on test bed data can provide reliable estimation for field T2 measurements, and therefore could be used to replace the missing sensor or erroneous measurements during sensor malfunctions, and can also be used as a virtual sensor to provide additional evidence for IGT condition monitoring in additional to the existing physical sensors.

Regarding the advantage of BNFM over ANFIS, the root mean squared errors (RMSEs) are checked for the ANFIS output TR by using a different number of rules in the FIS, e.g. varying from 2 to 10 rules. The output RMSEs of TR are plotted for training (test bed) data and for field estimation separately, as shown in Fig. 8. From the training set, it is shown that a selection of 7 rules/clusters gives the least RMSE, however, in the field estimation by using 7 rules, it gives a larger RMSE for estimated TR. This is due to the fact that, an optimised model is more specific towards the training data, and
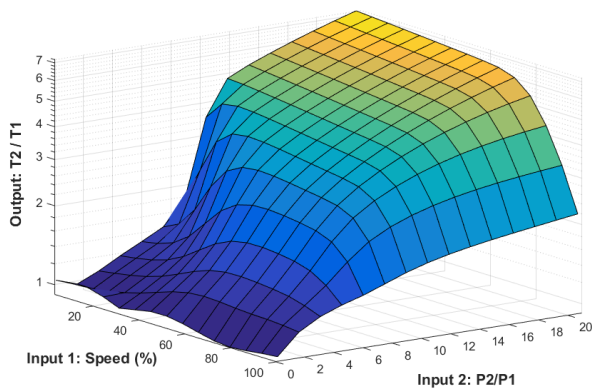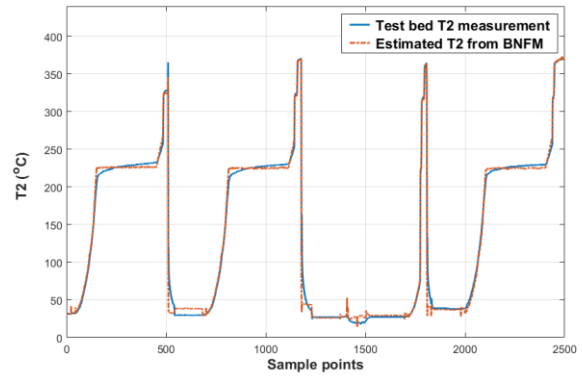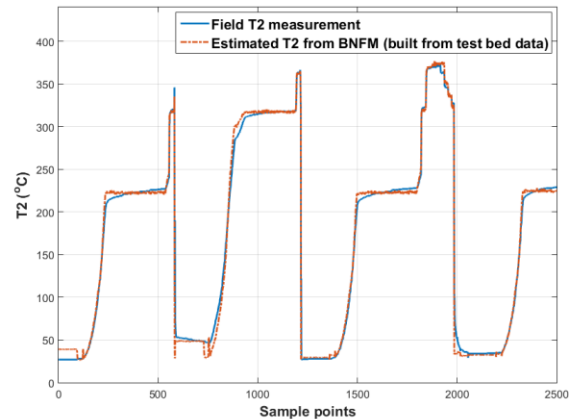

(a)

(b)

Fig. 7. BNFM estimation results of T2 for (a) test bed data – the training data and (b) field data – with supposedly unknown T2 measurement.

therefore tends to have over-fitting issues when encountering new input data. In this case, 9 rules will give the least errors for both training data and field estimation, however, in the absence of real field measurement which is supposed to be unknown, it is difficult to determine 9 rules by training errors alone. Therefore, it can be seen that the choice of 5 clusters/rules is reasonable and presents a decent compromise between model specificity and sensitivity.
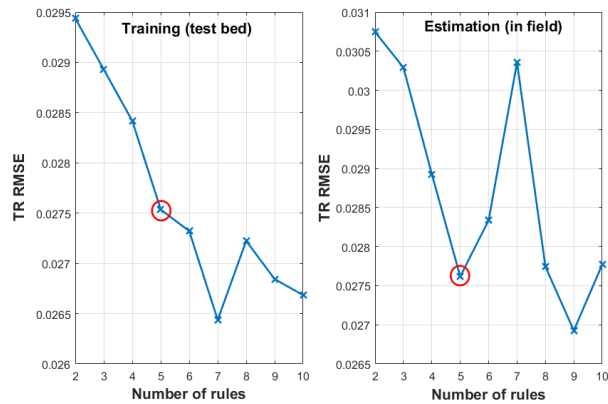

Fig. 6. BNFM 3-D surface plot of inputs (speed and PR) vs. output (TR).

Fig. 8. Predictive performance vs. number of rules in ANFIS, for training data and for estimated measurement at site.

IV. Conclusion

A BNFM methodology is proposed in this paper, where VBGMM is used to identify the number of rules/ input MFs of an ANFIS in order to present more interpretability of the system modelling. The effectiveness of the proposed approach is then validated by using an experimental trial of an IGT compressor. It is shown that the built BNFM model based on test bed data can reliably estimate the T2 measurement on an IGT compressor at site, which can be used for both condition monitoring and missing sensor measurement reconstruction purposes. Moreover, this technique can be readily transferable for system modelling of other components to estimate different measurements.

Nomenclature

ANFIS — Adaptive Neuro-Fuzzy Inference System
ANN — Artificial Neural Network
BNFM — Bayesian Neuro-Fuzzy Modelling
FIS — Fuzzy Inference System
GMM — Gaussian Mixture Model
IGT — Industrial Gas Turbine
MF — Membership Function
P1 — Compressor inlet pressure (ambient pressure)
P2 — Compressor discharge pressure
PR — Pressure Ratio
RMSE —Root Mean Squared Error
Speed – Compressor rotational speed
T1 — Compressor inlet temperature (ambient temperature)
T2 — Compressor discharge temperature
TR — Temperature Ratio
VB — Variational Bayesian
VBGMM — Variational Bayesian Gaussian Mixture Model

Acknowledgment

References

[1] Y. Zhang, C. Bingham, Z. Yang, B.W.K. Ling, M. Gallimore. "Machine fault detection by signal denoising—with application to industrial gas turbines." *Measurement*, vol. 58, pp. 230-240, 2014.

[2] Y. Zhang, C. Bingham, M. Gallimore. "Fault detection and diagonosis based on extensions of PCA." *Advances in Military Technology*, vol. 8, no. 2, pp. 27-41, 2013.

[3] Y. Zhang, C. Bingham, M. Gallimore. "Applied sensor fault detection, identification and data reconstruction." *Advances in Military Technology*, vol. 8, no. 2, pp.13-26, 2013.

[4] V. Panov. "GasTurboLib – Simulink library for gas turbine engine modelling." *Proc. of ASME Turbo Expo*, Florida, USA, 2009.

[5] H. Asgari, X. Chen, M.B. Menhaj, R. Sainudiin. "Artificial neural network–based system identification for a single-shaft gas turbine." *ASME. J. Eng. Gas Turbines Power*, vol. 135, no. 9, 092601-7, 2013.

[6] Y. Zhang, C. Bingham, M. Gallimore, Z. Yang, J. Chen. "Applied sensor fault detection and validation using transposed input data PCA and ANNs." *Proc. of IEEE Conf. on MFI*, Hamburg, Germany, Sep. 2012.

[7] Y. Zhang, C. Bingham, M. Gallimore, Z. Yang, J. Chen. "Sensor fault detection and diagnosis based on SOMNNs for steady-state and transient operation." *Proc. of IEEE Conf. on CIVEMSA*, Milan, Italy, July 2013.

[8] Y. Zhang, J. Chen, C. Bingham, M. Mahfouf. "A new adaptive Mamdani-type fuzzy modeling strategy for industrial gas turbines." *Proc. of IEEE Conf. on Fuzzy Systems*, Beijing, China, July 2014.

[9] J.S.R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no. 3, pp. 665-685, 1993.

[10] K. Salahshoor, M.S. Khoshro, M. Kordestani, "Fault detection and diagnosis of an industrial steam turbine using a distributed configuration of adaptive neuro-fuzzy inference systems." *Simulation Modelling Practice and Theory*, vol. 19, no. 5, pp. 1280-1293, 2011.

[11] B. Chen, P.C. Matthews, P.J. Tavner, "Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS." *Expert Systems with Applications*, vol. 40, no. 17, pp. 6863-6876, 2013.

[12] Y. Zhang, S. Cruz-Manzo, A. Latimer. "Start-up vibration analysis for novelty detection on industrial gas turbines." *Proc. of XI Int. Symp. on Industrial Electronics*, Banja Luka, Bosnia and Herzegovina, Nov. 2016.

[13] B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, N. de Freitas. "Taking the human out of the loop: A review of Bayesian optimization." *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016.

[14] Y. Zhang, C. Bingham, M. Gallimore, J. Chen. "Steady-state and transient operation discrimination by variational Bayesian Gaussian mixture models." *Proc. of IEEE Int. Workshop on MLSP*, Southampton, UK, Sep. 2013.

[15] P. Tahmasebi, A. Hezarkhani. "A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation." Computers & Geosciences, vol. 42, pp. 18–27, 2012.

[16] M. Sugeno. "Industrial applications of fuzzy control." Elsevier Science Pub. Co., 1985.

[17] C.M. Bishop. "Chapter 9: Mixture Models and EM." Pattern Recognition and Machine Learning. Series: Information Science and Statistics. Springer, Heidelberg, 2006.

[18] A. Corduneanu, C.M. Bishop. "Variational Bayesian model selection for mixture distributions." *Artificial Intelligent Statistics*, pp. 27–34. 2001.

[19] Y. Zhang, J. Chen, C. Bingham, T. Gordon. "Variational Bayesian modified adaptive Mamdani Fuzzy Modelling for use in condition monitoring." *Proc. of IEEE Conf. on CIVEMSA*, Shenzhen, China, June 2015.

[20] MATLAB®2016a, The MathWorks, Inc., Natick, Massachusetts, United States.