Hydro-generators fault diagnosis with short-time-wavelet-entropy and variational auto-encoder

Ryad Zemouri¹,*, Simon Bernier¹, Olivier Kokoko¹, Arezki Merkhoufi

¹Institut de Recherche d’Hydro-Québec (IREQ), Varennes, QC, J3X1S1 Canada

*Corresponding author: zemouri.ryad@hydroquebec.com

Abstract. The prognosis and health management (PHM) of hydroelectric plants are full of difficulties caused by the complexity of the hydro-generators where each machine is different and almost unique. At industrial level, several tools are used to monitor the generator condition. Among these tools, the measurement of magnetic stray flux is one which is gaining interest. This measurement is generally based on an inductive sensor and mainly mounted near the stator. The main advantages of the magnetic stray flux are the non-invasive nature and the simplicity of its implementation. In this work, the discrete wavelet transform (DWT) is used to decompose the stray flux signal. Short-Time-Wavelet-Entropy (STWE) is then applied to extract the features from the sub-bands. Finally, a variational auto-encoder (VAE) is used in an unsupervised learning process to structure the STWE signatures of more than 400 stray flux measurement collected on real hydroelectric plants. The obtained results show that the VAE has well captured the features from the wavelet entropy (WE) signatures. An analysis of the resulting latent space shows a strong correlation between a given trajectory in the reduced space and an increase of the WE.

1. Introduction

In a global reluctance to build new hydroelectric power plants and in the context of increasing demand for electricity, Hydro-Québec must maximize the production with the highest possible reliability and minimum downtime. Hydro-Québec operates more than 350 hydro-generators, with most of them subjected to normal and slow degradation mechanisms. However, in some cases, degradation is more accelerated and can lead to premature failure and significant economic losses. To avoid this, Hydro-Québec put significant efforts on functional safety condition-based maintenance and prognosis of their equipment. A PHM platform has been developed at Hydro-Québec’s research institute (IREQ), in order to detect and prevent incipient faults in a real-time or at least at early stage of the apparition [1]. Diagnosis is based on online measurement of signals such as partial discharges [2-3], vibrations and acoustics signals [4], temperature [5], magnetic stray flux or magnetic flux in the air-gap [6]. In this work, we study the health of the hydro-generators by measuring the magnetic stray flux signal collected at the stator. This kind of signal provides useful information that can be used to monitor the health of the machine such as the static or dynamic eccentricity of the rotor, or the inter-turn short circuit [6-9]. The measurement of the magnetic stray flux is particularly interesting as it is non-intrusive.

The database processed in this paper contains more than 400 magnetic stray flux signals from 8 hydro-electric power plants operated by Hydro-Québec. The conventional approaches use Fourier
transform (FT) as a signal processing tool to assess any specific changes in frequency amplitude according to the fault under study. However, it is necessary to have an a-priori knowledge of the frequencies contained in the signals under study [10]. The FT needs a high stationarity of the signal and does not provide the time evolution of the frequency patterns. This disadvantage can be partially resolved in order to follow the evolution of the frequencies by using a short time Fourier transform (STFT) where the FT is applied to a time-evolving windows. In this case, a compromise between frequency versus time resolution must be found. Narrow windows lead to poor frequency resolution, in the opposite, too wide windows lead to poor time resolution. This limitation becomes significant when the signal has a series of widely dispersed transient components. To overcome these limitations, the discrete wavelet transform (DWT) represents an interesting solution [11-12]. The DWT makes no assumptions about the stationarity of the signal; the only information needed is the time evolution. In this case, the tracking of the temporal evolution of the frequency’s patterns contained in the signal becomes possible with an appropriate time-frequency representation. The wavelet analysis is a method that relies on the introduction of an appropriate base. It allows to characterize a temporal signal given by the distribution of the amplitude in this base. If the wavelet is necessary to form an appropriate orthogonal basis, it has the advantage of being able to uniquely decompose an arbitrary time series and to invert the decomposition [11-12]. DWT is then based on the hypothesis that any signal can be approximated as the arithmetic sum of several functions called wavelets [13]. In PHM application, the wavelet theory has been widely used for stationary and non-stationary signal analysis such as power systems [14][7][8][15], electrical machines [10], cutting machinery [16], partial discharge in power transmission [17], mechanical transmission system of shearer ranging arm [18], online fuel cell remaining useful lifetime prediction [19]. Thus, the DWT as a time-frequency representation of a signal, provides an effective tool for feature extraction.

An emerging technique which allows to evaluate the degree of order/disorder of a signal is to compute the Shannon entropy to the relative energies of each wavelet decomposition [20]. Therefore, the obtained wavelet entropy (WE) provides a useful insight into the inherent dynamic process related to the signal, and has been employed in specific tasks as an indicator of faults. In fact, a highly ordered process can be thought of as a single-frequency periodic signal. For such a signal, all relative wavelet energies will be nearly zero, except for the level of wavelet resolution that includes the frequency of the representative signal. For this level, the relative wavelet energy will be almost equal to one and, therefore, the total WE will be close to zero. Conversely, a random signal can be considered to represent a very disordered behavior. This type of signal will have a wavelet mapping with all frequency bands contributing significantly. The WE will reach their maximum values since the relative energy of the wavelets will be mostly equal for all the levels of resolution.

WE has been used in specific assignments as an indicator of faults and features extraction in several PHM domains. The entropies of the wavelet decompositions have been leveraged for power transmission line fault classification and location [14][21]. In [10], a stray flux signal was first decomposed by DWT, then WE computed to feed an artificial neural network (ANN) in order to estimate the winding insulation healthiness in induction motors. WE has been used by [22] to eliminate white noise in partial discharge signals by computing a new improved threshold function. For fault diagnosis of gearbox, WE has been used by [23] to extract statistical features from wavelet coefficients of the vibration signals. The obtained features were then used to feed an ANN for fault classification. In other domains such as biomedical applications, WE has been widely used for the dynamic analysis of electroencephalogram signals [12], diagnosing congestive heart failures [24], or for brain computer interface [25].

In this study, a WE is utilized to extract the features from a stray flux signals for hydro power generators fault detection and diagnosis. As illustrated in the figure 1, the stray flux signals were first decomposed by a DWT in order to have the different time-frequency sub-bands. The energy of the Shannon WE distribution will strongly depend on the various operating and healthy conditions. In order to analyze the various WE signatures and to identify their likelihoods, we have exploited the properties of the latent space obtained by an unsupervised training of a VAE. This approach based on an
unsupervised analysis is particularly adapted to our industrial context where most of the stray flux signals acquired from power plants are not labeled. This whole approach is presented in the next section. Experimental results and latent space analysis are presented in section 3. The conclusion and future works are summarized in the last section of the paper.

2. Signal Processing

2.1. Discrete Wavelet Transform

The DWT is used to decompose the stray flux signal with high-pass (HPF) and low-pass (LPF) filters, where the bandwidth of both filters is the same. After each decomposition step, the sampling frequency of the signal is divided by two. The output of the low-pass filter is recursively split to produce the next sub-band of the wavelet. The original signal $S(k)$ is then decomposed into details and approximations as follow:

$$ S(k) = \sum_{i=1}^{j} D_i(k) + A_j(k) = \sum_{i=1}^{j+1} S_{b_i}(k) $$

(1)

Where $j$ is the number of decomposition level, $D_i(k)$ and $A_j(k)$ are details and approximations of wavelet. The stray flux signal is then decomposed into several sub-bands $S_{b_i}(k)$. In this paper, as shown by the figure 2, a total of 10 sub-bands are defined and the Daubechies-8 ($n = 8$) is used as a mother wavelet [11]. According to Nyquist's theorem and based on sampling frequency of 50 kHz, the highest frequency contained on the signal is 25 kHz. The several frequency ranges of each sub-band are also given by the figure 2. Figure 3 illustrates an example of the different sub-bands $S_{b_i}(k)$ obtained by decomposing an original signal $S(k)$ with a DWT.

Figure 1. Flowchart of the stray flux signals analysis.

Figure 2. DWT with the corresponding frequency bands obtained from a stray flux signal with a sampling frequency of 50 kHz.
2.2. Short-Time-Wavelet-Entropy

The Shannon entropy is a practical index for evaluating and assessing the probability distribution [7] [20]. In order to measure the uncertainty and disorder of wavelet sub-bands, we define the moving energy $E_i(k)$ of each sub-band $S_{bi}(k)$ within a slide window $[k, k + \Delta k]$ as follow:

$$E_i(k) = \sum_{k}^{k+\Delta k} |S_{bi}(k)|^2$$

(2)

The total energy $E_{total}(k)$ of the signal $S(k)$ within the moving window $[k, k + \Delta k]$ is given by:

$$E_{total}(k) = \sum_{i=1}^{j+1} E_i(k)$$

(3)

The normalised energy value $P_i(k)$ of each wavelet sub-band $S_{bi}(k)$, relative to the total energy $E_{total}(k)$ of the signal is given by the following equation:

$$P_i(k) = \frac{E_i(k)}{E_{total}(k)}$$

(4)

Where $\sum_i P_i(k) = 1$. The Short-Time-Wavelet-Entropy (STWE) of each sub-band $S_{bi}(k)$ through each moving window $[k, k + \Delta k]$ is then given by:

$$W_i(k) = -P_i(k) \log(P_i(k))$$

(5)

The figure 4 shows some examples of different stray flux signals with their three majors corresponding STWE signatures $W_i(k)$ obtained by the equation 5. The length $\Delta k$ of the sliding window used to compute the moving energy $E_i(k)$ is 1 second.
2.3. Variational Autoencoder

The VAE is composed by two main structures: an encoder $\phi$ and a decoder $\theta$. In a first step, the input data $x$ is transformed into a latent representation $z$ by the encoder function $z = f_\phi(x)$, whereas in the second step, the original data is reconstructed by the decoder $\hat{x} = h_\theta(z)$. The popularity of the VAE is due to the combination of the Bayesian inference and the efficiency of the ANNs to obtain a non-linear low-dimensional latent space. Recently, the VAEs are increasingly used in several PHM applications [26-39]. The Bayesian inference is obtained by a sampling additional layer $z$ with a prior specified standard Gaussian distribution $p(z) \equiv N(O, I)$, where $I$ is the identity matrix. Each element $z_i$ of the sampling layer is obtained as follows:

$$z_i = \mu_i + \sigma_i \epsilon$$

(6)

Where $\mu_i$ and $\sigma_i$ are the $i^{th}$ components of the mean and standard deviation vectors, $\epsilon$ is a random variable following a standard Normal distribution $\epsilon \sim N(0,1)$. Unlike the AE, the VAE generates vector of means $\mu_i$ and standard deviations $\sigma_i$ which provides more continuity in the latent space than the original AE. In VAE, the loss function is composed of two terms:

$$L = L_{rec} + L_{kl}$$

(7)

The first term $L_{rec} = - \mathbb{E}_{q_\phi(z|x)} [\log p_\theta(x|z)]$ is the reconstruction loss function, which allows to minimize the difference between the input and output instances. Both the negative expected log-likelihood (e.g., the cross-entropy function) and the mean squared error (MSE) can be used. The second term $L_{kl} = D_{kl}(q_\phi(z|x) \parallel p(z))$ corresponds to the Kullback–Leibler ($D_{kl}$) divergence loss term that forces the generation of a latent vector with the specified Normal distribution [40]. The $D_{kl}$ divergence is a theoretical measure of proximity between two densities $q(x)$ and $p(x)$ and measures how close the conditional distribution density $q_\phi(z|x)$ of the encoded latent vectors is from the desired Normal distribution $p(z)$.

The figure 5 gives the detailed architecture of the VAE used to analyse the STWE signatures. A slide window of $10 \times 10$ is used as input of the VAE. Each input signal $S(k)$ will then generate a cluster of features in the reduced latent space. Two stray flux signals with two signatures having maximum likelihood will get two feature clusters projected into the same area of the latent space. On the contrary, two different signatures will have two opposite projections.
3. Experimental results

3.1. Latent space analysis

Figure 6 gives the projection through the encoder of the whole database, where ($z_1, z_2$) are the output of the sampling layer of the VAE. The statistical distribution of the latent space is given by the figure 7. We can see that some areas of this latent space are more highly concentrated than others. These regions correspond to the most frequently encountered signatures in the database.

Once the VAE learning stage is complete, the second step is to analyze the distribution of data in the reduced space in order to have an interpretation of the health status that corresponds to each area. Since the main part of the data set is unlabeled, this identification step of the latent space distribution is crucial for the stray flux signal recognition and fault detection. We have taken some samples from this latent space and analyzed their different $W_i(k)$ signatures. As shown in Figure 6, a total of eight areas were analyzed. The shape and the value of each entropy $W_i(k)$ corresponding to each sub-band $Sb_i(k)$ of each area is provided in the Figure 8. Table 1 and 2 display respectively the mean value (in %) of the relative normalised energy $E_i(k)$ and the mean value of the $W_i(k)$ for each sub-band $Sb_i(k)$ corresponding to each area.

What we can say from this first analysis of the latent space, is that each area, among those selected, has its own STWE signature $W_i(k)$. For example, area #1 concerns the signals with a low entropy for all the sub-bands. This is due to the mono-frequency character of the original signal: the sub-band 10 dominates the total energy of the signal with $E_{10} = 97\%$. On the other hand, the area #5 covers signals with a completely different STWE signature compared to the area #1, 90% of the signal energy is distributed among the sub-bands #7,8,9. Other types of signals are also projected in other regions, such as region #8 where 93% of the total energy of the signal is dominated by the sub-bands #1,2,3,4.
3.2. Trend analysis

In order to investigate whether a particular trajectory in the latent space correlates with a particular trend in the evolution of the STWE $W_i(k)$, we have selected 38 samples of stray flux signals with a projection into latent space located on the trajectory shown by the figure 9. For each of these 38 signals, we have calculated the mean value $\bar{W}_i$ of the STWE $W_i(k)$ for each sub-band $S_{bi}(k)$. The figure 10 illustrates the evolution of $\bar{W}_i$ according to each selected area for each sub-band $S_{bi}(k)$. The total entropy is also given for each of the 38 signals by the top plot of the figure. One can see that the top graph in the figure 10 shows an overall increasing trend in WE from projected signal in area #1 to projected signal in area #38. This increasing trend is also visible for the sub-bands #7 #8 and #9. This result shows that the VAE has captured this upward trend of the WE and has organized part of its latent space according to the evolution of this criteria. This is especially important since, if we refer to some recent studies conducted on the diagnosis of hydropower generators [7-8], there is a strong correlation between the increase of the WE and some failures like static or dynamic eccentricity. The next step of our study would be to analyze the failure mechanisms according to the operating mode and/or to the machines categories. Then, a correlation with the various levels of the WE should be conducted. Therefore, the latent space of the VAE could be exploited as a visual PHM support.

Figure 8. The $W_i(k)$ corresponding to each sub-band $S_{bi}(k)$ of each area selected from the latent space.

Figure 9. Selected Area for the trend analysis.

Figure 10. Trend analysis of each sub-band.
### Table 1. The mean value (in %) of the relative normalised energy $E_i(k)$ of each wavelet sub-band $S_{bi}(k)$ corresponding to each selected area of the figure 6.

| #1   | #2    | #3    | #4    | #5    | #6    | #7    | #8   |
|------|-------|-------|-------|-------|-------|-------|------|
| $S_{b1}$ | 0.005643 | 0.004632 | 0.003587 | 0.00345 | 0.006461 | 0.00525 | **4.443** | **25.15** |
| $S_{b2}$ | 0.02645 | 0.06347 | 0.02334 | 0.02402 | 0.01906 | 0.006002 | **2.025** | **11.84** |
| $S_{b3}$ | 0.1963 | 0.301 | 0.07621 | 0.08365 | 0.06513 | 0.012 | 0.8874 | **48.43** |
| $S_{b4}$ | 0.04469 | 0.04693 | 0.1152 | 0.1671 | 0.03352 | **0.8874** | **48.43** |
| $S_{b5}$ | 0.1149 | 0.1351 | 1.142 | 0.6156 | 0.7754 | 0.02364 | **0.3569** |
| $S_{b6}$ | 0.02741 | 1.179 | **5.429** | **3.219** | **4.696** | 0.03235 | **0.4587** |
| $S_{b7}$ | 0.5429 | 1.705 | 2.75 | **12.22** | **20.72** | 0.297 | 0.7512 | **4.696** |
| $S_{b8}$ | 1.118 | **3.566** | 2.37 | 1.239 | **33.89** | 0.7882 | 1.089 | **0.781** |
| $S_{b9}$ | 0.4392 | 86.01 | 81.25 | 76.06 | 36.38 | 18.77 | **5.108** |
| $S_{b10}$ | **97.48** | **6.986** | **81.25** | **76.06** | **36.38** | **18.77** | **5.108** |

### Table 2. Mean value of the $W_i(k)$ for each sub-band $S_{bi}(k)$ corresponding to each selected area of the figure 6.

| #1 | #2 | #3 | #4 | #5 | #6 | #7 | #8 |
|----|----|----|----|----|----|----|----|
| $S_{b1}$ | 0.00049 | 0.00046 | 0.00037 | 0.00037 | 0.00062 | 0.00043 | **0.14** | **0.35** |
| $S_{b2}$ | 0.0022 | 0.0047 | 0.002 | 0.002 | 0.0016 | 0.00049 | **0.079** | **0.25** |
| $S_{b3}$ | 0.012 | 0.017 | 0.0055 | 0.006 | 0.0048 | 0.00092 | **0.042** | **0.35** |
| $S_{b4}$ | 0.0035 | 0.0036 | 0.0078 | 0.011 | 0.0089 | 0.0021 | 0.026 | **0.22** |
| $S_{b5}$ | 0.0078 | 0.0089 | 0.051 | 0.031 | 0.037 | 0.0017 | 0.02 | 0.038 |
| $S_{b6}$ | 0.0022 | 0.052 | **0.16** | **0.11** | **0.14** | 0.0025 | 0.025 | **0.025** |
| $S_{b7}$ | 0.028 | 0.069 | 0.099 | **0.26** | **0.33** | 0.017 | 0.037 | **0.017** |
| $S_{b8}$ | 0.05 | **0.12** | 0.089 | 0.054 | **0.37** | 0.038 | 0.049 | 0.014 |
| $S_{b9}$ | 0.024 | **0.13** | **0.17** | **0.21** | **0.37** | **0.31** | **0.15** | 0.0093 |
| $S_{b10}$ | 0.025 | **0.19** | **0.18** | **0.18** | **0.11** | **0.18** | **0.14** | **0.11** |

### 4. Conclusions
In this paper, Discrete Wavelet Transform has been used to decompose stray flux signals for hydro-generators monitoring. A Short-Time-Wavelet-Entropy was applied to extract the features from the sub-bands. Finally, a Variational Auto-Encoder has been exploited for the properties of its latent space to structure the different signatures thus obtained. The methodology has been tested on a real stray flux database signal obtained from several power plants operated by Hydro-Québec. The experimental results show that the Variational Autoencoder has well captured the features from the Wavelet Entropy signatures. A trend analysis showed a strong correlation between a given trajectory in the latent space with an upward trend in the Wavelet Entropy. This result provides us with a good basis for the next step, which will aim at a more detailed analysis of the failure mechanisms according to the machine categories operated in the various modes. The identification of the different regions of the latent space would then allow to consider a classifier at the output of the encoder with a supervised or semi-supervised learning.

### References
[1] Blancke O, Tahan A, Komljenovic D, Amyot N, Lévesque M and Hudon C 2018 Reliability Engineering & System Safety 180 136 – 151 ISSN 0951-8320.
[2] Lévesque M, Amyot N, Hudon C, Bélec M and Blancke O 2017 Annual Conference of the Prognostics and Health Management Society 2017 vol 8 p 7.

[3] Zemouri R, Lévesque M, Amyot N, Hudon C, Kokoko O and Tahan S A 2020 IEEE Access 8 5438–5454 ISSN 2169-3536.

[4] Vu VH, Thomas M and Lafleur F 2014 Mechanics & Industry 15 29–37.

[5] Kahlwati G, Tahan A, Merkhouf A, Hudon C and Al-Haddad K 2020 2020 International Conference on Electrical Machines (ICEM) vol 1 pp 1418–1423 ISSN 2381-4802.

[6] Kedjar B, Merkhouf A and Al-Haddad K 2020 2020 International Conference on Electrical Machines (ICEM) vol 1 pp 1384–1389 ISSN 2381-4802.

[7] Ehya H, Nysveen A, Nilssen R and Liu Y 2021 IET Electric Power Applications 15 890–902.

[8] Ehya H and Nysveen A 2021 IEEE Transactions on Industrial Electronics 1–1 ISSN 1557-9948.

[9] Ehya H and Nysveen A 2021 IEEE Transactions on Industry Applications 57 3573–3581 ISSN 1939-9367.

[10] Zamudio-Ramirez I, Osornio-Rios R A, Romero-Troncoso R d J and Antonino-Daviu J A 2019 IECON 2019 45th Annual Conference of the IEEE Industrial Electronics Society vol 1 pp 3716–3722 ISSN 2577-1647.

[11] Daubechies I 1992 CBMS-NSF Regional Conference Series in Applied Mathematics.

[12] Rosso O A, Blanco S, Yordanova J, Kolev V, Figliola A, Schürmann M and BarE 2001 Journal of Neuroscience Methods 105 65–75.

[13] Perez-Macias F J, Perez-Ramirez C A, Valtierra-Rodriguez M, Dominguez-Gonzalez A and AmezquitaSanchez J P 2017 IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society pp 3251–3256.

[14] Zheng-you H, Xiaocing C and Guoming L 2006 2006 International Conference on Power System Technology pp 1–6.

[15] Zamudio-Ramirez I, Antonino-Daviu J A, Osornio-Rios R A, de Jesus Romero-Troncoso R and Razik H 2020 IEEE Transactions on Industrial Electronics 67 5050–5059 ISSN 1557-9948.

[16] Liang P, Deng C, Wu J and Yang Z 2020 Measurement 159 107768 ISSN 0263-2241.

[17] Qu N, Li Z, Zuo J and Chen J 2020 IEEE Access 8 87060–87070 ISSN 2169-3536.

[18] Zheng X, Tang Y Y and Zhou J 2019 IEEE Transactions on Signal Processing 67 1696–1711 ISSN 1941-0476.

[19] Ibrahim M, Steiner N Y, Jemei S and Hissel D 2016 IEEE Transactions on Industrial Electronics 63 5057–5068 ISSN 1557-9948.

[20] Shannon C E 1948 The Bell System Technical Journal, 27 379–423.

[21] Dasgupta A, Nath S and Das A 2012 Electric Power Components and Systems 40 1676–1689.

[22] Guojun L, Yong W, Le L, ShaoFeng G, Haiqing N and Xuemei W 2017 2017 IEEE 11th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED) pp 475–479.

[23] Heidari Bafroui H and Ohadi A 2014 Neurocomputing 133 437–445 ISSN 0925-2312.

[24] Isler` Y and Kuntalp M 2007 Computers in Biology and Medicine 37 1502–1510 ISSN 0168-2214 qT Variability & Heart Rate Variability.

[25] Sankar A S, Nair S S, Dharan V S and Sankaran P 2015 Procedia Computer Science 46 1476–1482 ISSN 1877-0509 proceedings of the International Conference on Information and Communication Technologies, ICICT 2014, 3-5 December 2014 at Bolgatty Palace & Island Resort, Kochi, India.

[26] Proteau A, Tahan A, Zemouri R and Thomas M 2021 Journal of Intelligent Manufacturing.

[27] Proteau A, Zemouri R, Tahan A and Thomas M 2020 The International Journal of Advanced Manufacturing Technology 111 3597–3611.

[28] Huang Y, Chen C and Huang C 2019 IEEE Access 7 139086–139096 ISSN 2169-3536.

[29] Lee S, Kwak M, Tsui K L and Kim S B 2019 Engineering Applications of Artificial Intelligence 83 13 – 27 ISSN 0952-1976.

[30] Cheng F, He Q P and Zhao J 2019 Computers & Chemical Engineering 129 106515 ISSN 0098-
[31] Martin G S, Droguett E L, Meruane V and das Chagas Moura M 2018 Structural Health Monitoring 0 1475921718788299.

[32] Wang K, Forbes M G, Gopaluni B, Chen J and Song Z 2019 IEEE Access 7 22554–22565 ISSN 2169-3536.

[33] Ren Wang Y, Jin Q, dong Sun G and fei Sun C 2019 Knowledge-Based Systems 163 438 – 449 ISSN 0950-7051.

[34] Yoon A S, Lee T, Lim Y, Jung D, Kang P, Kim D, Park K and Choi Y 2017 CoRR abs/1709.00845.

[35] Zhang Z, Jiang T, Zhan C and Yang Y 2019 Journal of Process Control 75 136 – 155 ISSN 0959-1524.

[36] Jiao R, Peng K and Dong J 2020 IEEE Transactions on Instrumentation and Measurement 69 8831–8843 ISSN 1557-9662.

[37] Zhang Z, Zhu J, liu Y and Ge Z 2020 2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS) pp 1370–1376.

[38] Zemouri R, Lévesque M, Hudon C and Kokoko O 2021 PHM Society European Conference, 6 494–503.

[39] Kokoko O, Hudon C, Lévesque M, Amyot N and Zemouri R 2021 2021 IEEE International Conference on Prognostics and Health Management (ICPHM) pp 1–6.

[40] Kingma D 2017 Variational inference & deep learning: A new synthesis Ph.D. thesis Faculty of Science (FNWI), Informatics Institute (IVI), University of Amsterdam.