Robust switching Kalman filter for diagnostics of long-term condition monitoring data in the presence of non-Gaussian noise

Robust switching Kalman filter for diagnostics of long-term condition monitoring data in the presence of non-Gaussian noise

Hamid Shiri, Jacek Wodecki, Radosław Zimroz*

Faculty of Geoengineering, Mining and Geology, Wroclaw University of Science and Technology, Na Grobli 15, 50-421 Wroclaw, Poland

E-mail: hamid.shiri@pwr.edu.pl

Abstract.

Machinery condition prognosis system uses long-term historical data to predict remaining useful life (RUL). One of the critical steps to reach this purpose is to segment long-term data into two or several degradation stages (Healthy, Unhealthy, and Critic stage). Finding changing points between regimes may be a crucial preliminary task for further predicting the degradation process. However, finding the accurate partition into two or more regimes is a challenging task in the actual application when the noise inherent in the observed process is non-Gaussian. Therefore, this paper introduced a robust methodology based on switching Kalman filters to address the problems mentioned. This approach uses multiple dynamic system models to explain different degradation stages, utilizing robust Bayesian estimation. Also, based on this fact, this approach works based on dynamic behavior; a threshold for diagnostics is no longer needed. Ultimately, the proposed approach is applied for the online diagnosis of simulated data sets in the presence of Gaussian and non-Gaussian noise. The result of the applied methodology on simulated data sets proves the method's efficacy.

Diagnostics, condition monitoring, fault detection, robust methods, non-Gaussian noise, threshold, switching Kalman filter

1. Introduction

Condition-based maintenance(CBM) has become increasingly popular in the industry with the development of condition monitoring systems. CBM programs are developed to assess the machine's condition by collecting massive amounts of data during the operation of machines. Using the long-term condition monitoring data is a crucial element in both diagnostics and prognostics. Many methods published in recent years for CBM can be categorized into four main groups [1]: stochastic-based [2, 3, 4, 5, 6, 7, 8], machine learning-based [9, 10, 11, 12], physics or model-based[13, 14], and hybrid methods [15, 16]. Both machine learning and stochastic approaches such as neural networks [12, 17, 18, 19, 20, 21, 22, 23] and hidden Markov models (HMM) [24, 25, 26, 27, 28, 29] have strong potential to be used for diagnostics and prognostics areas. Nevertheless, they need a large number of data for the model to be trained. Furthermore, by changing the domain of working or changing the machine, they need to retrain again, which is often not possible in actual application. On the other hand, model-based methods use a mathematical representation of the degradation process, which needs less training data.

However, it is necessary to have enough knowledge of the degradation process. Also, most of the mentioned methods work based on a pre-established fault detection threshold that is often provided by the manufacturer; a threshold corresponding to a change from "Good Condition" (healthy stage) to "Warning" (degradation stage) and from "Warning" to "Alarm" stage (critic stage). Unfortunately, we do not know about limit values or the desired lifetime in many cases, especially when the machine is unique. In addition, this task will be more complicated when the machine works in a harsh area, where the inherent noise is a heavy tail. This may occur for several reasons. Wind inflows are one of the possible sources of heavy tail noise in wind turbine [30, 31, 32], the ore falling on devices in the mining environment is another [33, 34].

However, most of the mentioned methods are developed based on the Gaussian noise assumption, which may be less effective in actual application in the face of non-Gaussian noise. Furthermore, our previous work[35] focused on identifying and analyzing the degradation process using the real benchmark data set (available in the prognostics area), which confirms the hypothesis of non-Gaussian assumption noise even in the laboratory data sets. Therefore, in this paper, we tried to introduce a robust method that can address an online diagnostic without using a threshold in the presence of a non-Gaussian noise problem. The robust version of the switching Kalman filter (SKF) is derived to solve these issues by defining multiple dynamic systems to describe a different degradation process using Bayesian estimation. So, a threshold is no longer needed because the proposed methodology applies to dynamic behavior.

The paper is structured as follows: after the introduction, the summary of the process is presented. After that, the critical aspects of the processing methods are described in theory. Then, the proposed model is simulated, and the results will be presented. Finally, the conclusions are formed.

2. Theory

2.0.1. Kalman filter(KF) The KF is a kind of Bayesian filter used for recursive state estimation of a dynamic system by minimizing the mean squared error in the presence of process and measurement noise. The discrete state-space representation of the model is as follows:

$$\begin{aligned} x_t &= A_{t-1}x_{t-1} + q_t \\ y_t &= H_t x_t + r_t \end{aligned} \tag{1}$$

where x_t is an actual state at time t, A is the state-transition model, q_t is the process noise, y_t is observation of actual state x_t, H_t is the observation process, which maps the actual state space into the observed space, and r_t is the observation noise. The noise terms q_t and r_t are normally distributed: $q_t \sim \mathcal{N}(0, Q_t)$, $r_t \sim \mathcal{N}(0, R_t)$. In the literature, several approaches can be employed to extract KF formulation: for instance, Bayesian rule, maximum posterior approaches (MAP), orthogonal principle, weighted least square method (WLS). The quadratic objective function in WLS is defined as follows:

$$J = \tfrac{1}{2}(y_t - H\hat{x}_t)^T R_t^{-1}(y_t - H\hat{x}_t) + \tfrac{1}{2}(\hat{x}_t - A\hat{x}_{t-1})^T P_{t|t-1}^{-1}(\hat{x}_t - A\hat{x}_{t-1}) \tag{2}$$

where

$$\begin{aligned} P_{t|t-1} &= E[e_t^- e^{-T}], \qquad & e_t^- &= x_t - \hat{x}_t \\ R_t &= E[r_t r_t^T], & Q_t &= E[q_t {q_t}^T] \end{aligned} \tag{3}$$

Therefore, by solving the Equation:

$$\frac{\partial J}{\partial \hat{x}_t} = 0 \ , \tag{4}$$

KF can be derived and its formulation is described with the following equations:

$$\hat{x}_t^- = A_t x_{t-1} \tag{5}$$

$$P_{t|t-1} = A_t P_{t-1|t-1} {A_t}^T + Q_t \tag{6}$$

$$\nu_t = y_t - H_t \hat{x}_t^- \tag{7}$$

$$K_t = (H^T R_t^{-1} H + P_{t|t-1}^{-1})^{-1} H^T R_t^{-1} \tag{8}$$

$$x_t = \hat{x}_t^- + K_t \nu_t \tag{9}$$

$$P_{t|t} = (I - K_t H_t) P_{t|t-1} (I - K_t H_t)^T + K_t R_t {K_t}^T \tag{10}$$

where ν_t, K_t and P_t are innovation (or measurement pre-fit residual), Kalman gain, and covariance matrix estimation error, respectively. The KF gives the optimal solution when the process and measurement noise have a Gaussian distribution, whereas if noise has a different distribution, KF produces a sub-optimal solution. It is because only the second-order measurement information is used.

To address the non-Gaussian problem, Izanloo et al.[36] presented the correntropy filter (C-Filter) for state estimation and indicated that the C-Filter has better performance than KF in the presence of non-Gaussian noise, thanks to using correntropy criterion, which utilizes higher-order measurement signal information. In the following subsection, the maximum correntropy criterion Kalman filter(MCKF) is derived to handle non-Gaussian noise based on the following references: [37, 36, 38, 39]

2.0.2. Maximum correntropy Kalman filter(MCKF) In the following, another form of a cost function based on the maximum correntropy criterion is introduced, see Eq.11, that is more robust in the presence of non-Gaussian noise.

$$J_m = G_\sigma(\|y_t - H\hat{x}_t\|_{R_t^{-1}}) + G_\sigma(\|\hat{x}_t - A\hat{x}_{t-1}\|_{R_{t|t-1}^{-1}}) \tag{11}$$

where $G_\sigma(x_i - y_i) = \exp(-\frac{\|x_i - y_i\|^2}{2\sigma^2})$ and σ is kernel size(bandwidth). The MCKF can be driven by minimizing the cost function Eq. 12 .

$$\frac{\partial J_m}{\partial \hat{x}_t} = 0 \tag{12}$$

The MCKF equations are given as follows:

$$\hat{x}_t^- = A_t x_{t-1} \tag{13}$$

$$P_{t|t-1} = A_t P_{t-1|t-1} {A_t}^T + Q_t \tag{14}$$

$$L_t = \frac{G_\sigma(\|y_t - H\hat{x}_t\|_{R_t^{-1}})}{G_\sigma(\|\hat{x}_t - A\hat{x}_{t-1}\|_{R_{t|t-1}^{-1}})} \tag{15}$$

$$\nu_t = y_t - H_t \hat{x}_t^- \tag{16}$$

$$K_t = (L_t H^T R_t^{-1} H + P_{t|t-1}^{-1})^{-1} L_t H^T R_t^{-1} \tag{17}$$

$$x_t = \hat{x}_t^- + K_t \nu_t \tag{18}$$

$$P_{t|t} = (I - K_t H_t) P_{t|t-1} (I - K_t H_t)^T + K_t R_t {K_t}^T. \tag{19}$$

When a significant outlier appears, the innovation term ν_t diverges, but k_t controls the divergence of the estimator $\hat{x}_t$. Please see these references [37, 36, 38, 39] for more details about the procedure of driving equations and stability.

The most significant limitation of KF and MCKF used for diagnostics and prognostics is that the degradation process must be time-invariant. However, in practical application, the

degradation process is composed of several components that change over time. Therefore, using a single model to describe the whole process would make incorrect state estimations and cause divergence or fluctuation. In lieu of this, the SKF, also known as the dynamic linear model, is proposed to handle the mentioned issue when the dynamic behavior of a process changes during the time and, if not possible, to use a unique model. However, general SKF has assumed the noise was Gaussian. In the next subsection, we introduce a new version of SKF based on MCKF to address these challenges.

2.0.3. Switching maximum correntropy Kalman filter (SMCKF) In this subsection, we introduce the switching maximum correntropy criterion Kalman filter (SMCKF), which is a robust version of SKF. The SMKF can be described as a dynamic Bayesian network. In each time step, S_t as the switching model variable and X_t as the state variable are hidden, which must be figured out from the observation Y_t. This may cause numerical problems, especially when the number of the regimes is increased, as argued in [40].

Kevin et al.in [41], developed an approximation approach, namely the generalized pseudo-Bayesian (GPB) method, to solve this issue. In every time step, the covariance and state estimation from all MCKFs in the last time step is mixed using weights based on the switch variable $S_t^{i|j}$ and the transition probabilities Z_{ij}, as expressed with the equations 20, 21, respectively.

$$S_t^{i|j} = \frac{Z_{ij} {S^i}_{t-1}}{\sum_{i=1}^{n} Z_{ij} {S^i}_{t-1}} \tag{20}$$

Weighted state and covariance estimates are derived as follows:

$$\begin{aligned} \tilde{x}_{t-1}^{j} &= \sum_{i=1}^{n} \tilde{S}^{i|j} {x^i}_{t-1} \\ \tilde{P}_{t-i}^{j} &= \sum_{i=1}^{n} \tilde{S}_t^{i|j} \{\bar{P}_{t-1}^{i} + [x_{t-1}^{i} - x_{t-1}^{j}][x_{t-1}^{i} - x_{t-1}^{j}]^T \end{aligned} \tag{21}$$

Also, by using measurement residuals, the likelihood of MCKF is calculated as Eq.22.

$$L_t^i = \mathcal{N}(v_t^i; 0, \bar{C}_t) \tag{22}$$

where $\bar{C}_t = (L_t H^T R_t^{-1} H + P_{t|t-1}^{-1})$. In the end, the probability of i-th MCKF model is equal to Eq.23.

$$S_t^i = \frac{L_t^i (\sum_{i=1}^{n} Z_{ij} S_{t-1}^i)}{\sum_{i=1}^{n} (L_t^i (\sum_{i=1}^{n} Z_{ij} S_{t-1}^i))} \tag{23}$$

The weighted state and covariance estimates are given by MCKF from Eq.13 to Eq.19 for each filter, which was used before to estimate the predicted state $\tilde{x}_{t-1}^{j}$ and covariance $\tilde{P}_{t-i}^{j}$, see Eq. 21. More details about the procedure of the switching structure and stability of the model are available in [41]. The SMCKF, which is composed of multiple MCKF with different state-space models, is applied to track changes in the degradation process. Then SMCKF is switched between the selected models according to their likelihood calculated from health index(HI). It should be noted that for applying this approach, the degradation model should be selected based on the case's physics.

Based on our primary assumption in this analysis, the degradation process comprises three stages. Zero-, first-, and second-order MCKF are used to simulate healthy, degradation, and critic stage, respectively. According to [42, 43], the actual state, error covariance, and observation matrices representing the polynomial MCKFs are illustrated with the subscripts 1, 2, and 3

below, showing the zero-, first-, and second-order MCKFs, respectively. Eq. 24 represent state matrices.

$$x_1(t) = \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}, x_2(t) = \begin{bmatrix} x \\ \dot{x} \\ 0 \end{bmatrix}, x_3(t) = \begin{bmatrix} x \\ \dot{x} \\ \ddot{x} \end{bmatrix} \tag{24}$$

$$A_1(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_2(t) = \begin{bmatrix} 1 & t_s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_3(t) = \begin{bmatrix} 1 & t_s & \frac{{t_s}^2}{2} \\ 0 & 1 & t_s \\ 0 & 0 & 1 \end{bmatrix} \tag{25}$$

where A_1, A_2, and A_3 are state transition matrices and t_s is discretisation step size. Also, the covariance matrices of the process are defined by the following:

$$Q_1(t) = q \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, Q_2(t) = q \begin{bmatrix} \frac{{t_s}^3}{3} & \frac{{t_s}^2}{2} & 0 \\ \frac{{t_s}^2}{2} & t_s & 0 \\ 0 & 0 & 0 \end{bmatrix}, Q_3(t) = q \begin{bmatrix} \frac{{t_s}^5}{20} & \frac{{t_s}^4}{8} & \frac{{t_s}^3}{6} \\ \frac{{t_s}^4}{8} & \frac{{t_s}^3}{3} & \frac{{t_s}^2}{2} \\ \frac{{t_s}^3}{6} & \frac{{t_s}^2}{2} & t_s \end{bmatrix} \tag{26}$$

where q is a scalar hyper-parameter (related to the noise) that can describe the uncertainty of the filter in actual application, which can be used for tuning of the SMCKF for a different machine.

Observation matrices H_1, H_2, and H_3 corresponding to models $1, 2, 3$ are defined as following:

$$H_1 = H_2 = H_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{27}$$

The transition matrix is defined by Eq. 28. Note that there is no possibility of transition to the previous state.

$$Z = \begin{bmatrix} 0.998 & 0.001 & 0.001 \\ \sim 0 & 0.999 & 0.001 \\ \sim 0 & \sim 0 & 1 \end{bmatrix} \tag{28}$$

The initial model probabilities are set with Eq. 29

$$\tilde{S}_0 = \begin{bmatrix} 0.9 & 0.05 & 0.05 \end{bmatrix} \tag{29}$$

State and covariance estimates have the following form, respectively:

$$\vec{x}_0 = \begin{bmatrix} y_0 \\ 0 \\ 0 \end{bmatrix}, \tag{30}$$

$$\bar{P}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{31}$$

3. Simulation

In this section, At first, the long-term data is generated based on the lifetime curve shape following the model known in the literature [35]. this degradation model is developed to explain real degradation of real machine that work in harsh environment such as mining and wind turbine which can be seen the noise with non-Gaussian caractristic. This degradation model consists of 3 regimes: healthy condition, slow degradation, and rapid degradation. The data could be modeled as a mixture of trend and random components. More details about simulated model can find, for example, in [35] . Then the proposed methodology is used to segment data into three regimes in the presence of Gaussian noise and non-Gaussian noise.

3.1. Gaussian Noise

In this subsection, the introduced methodology is applied to simulated data set in the presence of non-Gaussian noise with by considering Student's t distribution with a degree of freedom $V = 3$. For this case, it assumed Time=1000 and Time =1600 are changing points between the healthy stage and degradation stage (CP1) and the degradation stage and healthy stage (CP2), respectively.

Fig. 2 ilistrated the result of the applied SKF and MCSKF. The estimated probability of each state using SKF and MCSKF during the degradation process is presented in panels (b) and (c), respectively. On the panel (b), we can see the probability of the healthy stage (the green line) is starting to reduce from Time=300; after a few fluctuations with the degradation stage (the yellow line), somewhere around Time=880, the degradation state will be the most probable stage which remains on this situation until Time=1446. Behind this time, the critic stage has more probability than the rest stages, but as we can see on panel (b) in Fig.2, the SKF has diverged after Time=1500. Panel (c) presented the result of SMCKF. Also, panels (d) and (e), the most probable stage during the degradation process and changing points extracted by employing SKF and SMCKF, are demonstrated. By comparing these two last panels, as expected, the MCSKF result is closer to the actual changing points while SKF has been significantly affected by non-Gaussian noise and, in the end, is diverged.

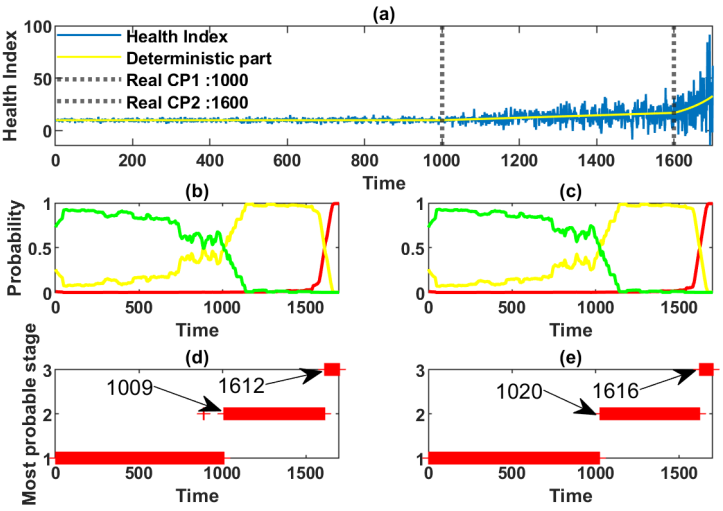

Figure 1: Detection of the stages in the presence of Gaussian noise, (a) health index (HI), (b) probability of stages performed by SKF, (c) probability of stages performed by SMCKF, (d) most probable stages based on the implementation of SKF, (e) most probable stages based on the implementation of SMCKF.

3.2. Non-Gaussian Noise

In this part, the proposed methodology is applied to data generated by the offered model, considering that the noise term has Student's t distribution with a degree of freedom $V = 3$. According to the simulation procedure, CP1 and CP2 is equal to Time=1000 and Time=1600, respectively.

Fig. 2 shows the result of the proposed methods when noise is non-Gaussian. The estimated probability of each state using SKF and SMCKF during the degradation process is presented in panels (b) and (c), respectively. On the panel (b), we can see the probability of the healthy stage (the green line) is starting to reduce from Time=300; after a few fluctuations with the

degradation stage (the yellow line), somewhere around Time=880, the degradation state will be the most probable stage which remains on this situation until Time=1446. Behind this time, the critic stage has more probability than the rest stages, but as we can see on panel (b) in Fig.2, the SKF has diverged after Time=1500. Panel (c) presented the result of SMCKF. Also, panels (d) and (e), the most probable stage during the degradation process and changing points extracted by employing SKF and SMCKF, are shown. By comparing these two last panels, as expected, the SMCKF result is closer to the actual changing points while SKF has been significantly affected by non-Gaussian noise and, in the end, is diverged.

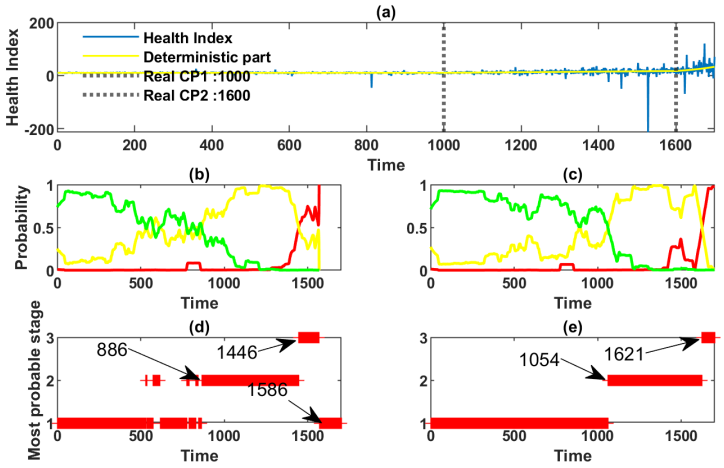

Figure 2: Detection of the stages in the presence of non-Gaussian noise, (a) health index, (b) probability of stages by SKF, (c) probability of stages by SMCKF, (d) most probable stages based on the implementation of SKF, (e) most probable stages based on the implementation of SMCKF.

4. Conclusions

This paper introduces a robust version of SKF to address the non-Gaussian noise issue. At first, we extracted an MCKF as a robust version of SKF. We are assuming that the degradation process is composed of three regimes corresponding to the machine life phases: good condition, degradation, and critic stage. Each of these regimes is modeled by MCKF. The SMCKF is introduced to infer from the current observations of the underlying process by calculating which of the three filters has the most significant probability in each time step. This method can provide more information to the decision-maker by providing probabilistic information about each degradation stage over time. Also, the proposed methodology was applied to the simulated degradation data set, and the result was presented for two cases in the presence of Gaussian and non-Gaussian noise. The results confirmed the performance of SMCKF in detecting different stages in both Gaussian and non-Gaussian noise in the comparison of SKF. Also, this method works according to the dynamic trend of the degradation process, and the threshold is no longer needed.

Funding

This work was supported by European Commission via the Marie Sklodowska Curie program through the ETN MOIRA project (GA 955681) - Hamid Shiri

Acknowledgments
The authors (Hamid Shiri) gratefully acknowledge the European Commission for its support of the Marie Sklodowska Curie program through the ETN MOIRA project (GA 955681).

Contributions
Hamid Shiri: methodology, software, writing, preparation, Paweł Zimroz : methodology, validation, formal analysis, Jacek Wodecki: supervision . All authors have read and agreed to the published version of the manuscript.

[1] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, J. Lin, Machinery health prognostics: A systematic review from data acquisition to rul prediction, Mechanical systems and signal processing 104 (2018) 799–834.
[2] X.-S. Si, W. Wang, C.-H. Hu, D.-H. Zhou, Remaining useful life estimation–a review on the statistical data driven approaches, European journal of operational research 213 (1) (2011) 1–14.
[3] Z.-S. Ye, M. Xie, Stochastic modelling and analysis of degradation for highly reliable products, Applied Stochastic Models in Business and Industry 31 (1) (2015) 16–32.
[4] A. Grzesiek, K. Gasior, A. Wyłomańska, R. Zimroz, Divergence-based segmentation algorithm for heavy-tailed acoustic signals with time-varying characteristics, Sensors 21 (24) (2021) 8487.
[5] A. Wyłomańska, R. Zimroz, Signal segmentation for operational regimes detection of heavy duty mining mobile machines-a statistical approach, Diagnostyka 15 (2014).
[6] R. Zimroz, M. Madziarz, G. Żak, A. Wyłomańska, J. Obuchowski, Seismic signal segmentation procedure using time-frequency decomposition and statistical modelling, Journal of Vibroengineering 17 (6) (2015) 3111–3121.
[7] D. Kucharczyk, A. Wyłomańska, J. Obuchowski, R. Zimroz, M. Madziarz, Stochastic modelling as a tool for seismic signals segmentation, Shock and Vibration 2016 (2016).
[8] K. Gasior, H. Urbańska, A. Grzesiek, R. Zimroz, A. Wyłomańska, Identification, decomposition and segmentation of impulsive vibration signals with deterministic components—a sieving screen case study, Sensors 20 (19) (2020) 5648.
[9] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, D. I. Fotiadis, Machine learning applications in cancer prognosis and prediction, Computational and structural biotechnology journal 13 (2015) 8–17.
[10] K. Thomsen, L. Iversen, T. L. Titlestad, O. Winther, Systematic review of machine learning for diagnosis and prognosis in dermatology, Journal of Dermatological Treatment 31 (5) (2020) 496–510.
[11] A. Diez-Olivan, J. Del Ser, D. Galar, B. Sierra, Data fusion and machine learning for industrial prognosis: Trends and perspectives towards industry 4.0, Information Fusion 50 (2019) 92–111.
[12] F. Moosavi, H. Shiri, J. Wodecki, A. Wyłomańska, R. Zimroz, Application of machine learning tools for long-term diagnostic feature data segmentation, Applied Sciences 12 (13) (2022) 6766.
[13] A. Heng, S. Zhang, A. C. Tan, J. Mathew, Rotating machinery prognostics: State of the art, challenges and opportunities, Mechanical systems and signal processing 23 (3) (2009) 724–739.
[14] M. S. Kan, A. C. Tan, J. Mathew, A review on prognostic techniques for non-stationary and non-linear rotating systems, Mechanical Systems and Signal Processing 62 (2015) 1–20.
[15] L. Liao, F. Köttig, Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction, IEEE Transactions on Reliability 63 (1) (2014) 191–207.
[16] Z. Zhao, J. Wu, T. Li, C. Sun, R. Yan, X. Chen, Challenges and opportunities of ai-enabled monitoring, diagnosis & prognosis: A review, Chinese Journal of Mechanical Engineering 34 (1) (2021) 1–29.
[17] P. Tamilselvan, Y. Wang, P. Wang, Deep belief network based state classification for structural health diagnosis, in: 2012 IEEE Aerospace Conference, IEEE, 2012, pp. 1–11.
[18] J. Liu, F. Lei, C. Pan, D. Hu, H. Zuo, Prediction of remaining useful life of multi-stage aero-engine based on clustering and lstm fusion, Reliability Engineering & System Safety 214 (2021) 107807.
[19] J. Singh, A. Darpe, S. P. Singh, Bearing remaining useful life estimation using an adaptive data-driven model based on health state change point identification and k-means clustering, Measurement Science and Technology 31 (8) (2020) 085601.
[20] W. Mao, J. He, B. Sun, L. Wang, Prediction of bearings remaining useful life across working conditions based on transfer learning and time series clustering, IEEE Access 9 (2021) 135285–135303.
[21] S. Sharanya, R. Venkataraman, G. Murali, Estimation of remaining useful life of bearings using reduced affinity propagated clustering, Journal of Engineering Science and Technology 16 (5) (2021) 3737–3756.
[22] K. Javed, R. Gouriveau, N. Zerhouni, A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering, IEEE transactions on cybernetics 45 (12) (2015) 2626–2639.

[23] Z. Chen, M. Wu, R. Zhao, F. Guretno, R. Yan, X. Li, Machine remaining useful life prediction via an attention-based deep learning approach, IEEE Transactions on Industrial Electronics 68 (3) (2020) 2521–2531.
[24] A. Giantomassi, F. Ferracuti, A. Benini, G. Ippoliti, S. Longhi, A. Petrucci, Hidden markov model for health estimation and prognosis of turbofan engines, in: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol. 54808, 2011, pp. 681–689.
[25] E. Ramasso, T. Denoeux, Making use of partial knowledge about hidden states in hmms: an approach based on belief functions, IEEE Transactions on Fuzzy Systems 22 (2) (2013) 395–405.
[26] F. Sloukia, M. El Aroussi, H. Medromi, M. Wahbi, Bearings prognostic using mixture of gaussians hidden markov model and support vector machine, in: 2013 ACS International Conference on Computer Systems and Applications (AICCSA), IEEE, 2013, pp. 1–4.
[27] A. Soualhi, H. Razik, G. Clerc, D. D. Doan, Prognosis of bearing failures using hidden markov models and the adaptive neuro-fuzzy inference system, IEEE Transactions on Industrial Electronics 61 (6) (2013) 2864–2874.
[28] R. B. Chinnam, P. Baruah, Autonomous diagnostics and prognostics in machining processes through competitive learning-driven hmm-based clustering, International Journal of Production Research 47 (23) (2009) 6739–6758.
[29] Q. Liu, M. Dong, W. Lv, X. Geng, Y. Li, A novel method using adaptive hidden semi-markov model for multi-sensor monitoring equipment health prognosis, Mechanical Systems and Signal Processing 64 (2015) 217–232.
[30] K. Gong, X. Chen, Influence of non-gaussian wind characteristics on wind turbine extreme response, Engineering structures 59 (2014) 727–744.
[31] K. R. Gurley, M. A. Tognarelli, A. Kareem, Analysis and simulation tools for wind engineering, Probabilistic Engineering Mechanics 12 (1) (1997) 9–31.
[32] A. Kareem, J. Zhao, Analysis of non-gaussian surge response of tension leg platforms under wind loads (1994).
[33] J. Hebda-Sobkowicz, R. Zimroz, A. Wyłomańska, J. Antoni, Infogram performance analysis and its enhancement for bearings diagnostics in presence of non-gaussian noise, Mechanical Systems and Signal Processing 170 (2022) 108764.
[34] J. Nowicki, J. Hebda-Sobkowicz, R. Zimroz, A. Wyłomańska, Dependency measures for the diagnosis of local faults in application to the heavy-tailed vibration signal, Applied Acoustics 178 (2021) 107974.
[35] Framework for stochastic modelling of long-term non-homogeneous data with non-gaussian characteristics for machine condition prognosis, Mechanical Systems and Signal Processing 184 (2023) 109677. doi:https://doi.org/10.1016/j.ymssp.2022.109677.
[36] R. Izanloo, S. A. Fakoorian, H. S. Yazdi, D. Simon, Kalman filtering based on the maximum correntropy criterion in the presence of non-gaussian noise, in: 2016 Annual Conference on Information Science and Systems (CISS), IEEE, 2016, pp. 500–505.
[37] B. Chen, X. Liu, H. Zhao, J. C. Principe, Maximum correntropy kalman filter, Automatica 76 (2017) 70–77.
[38] G. Wang, R. Xue, J. Wang, A distributed maximum correntropy kalman filter, Signal Processing 160 (2019) 247–251.
[39] G. Wang, Y. Zhang, X. Wang, Iterated maximum correntropy unscented kalman filters for non-gaussian systems, Signal Processing 163 (2019) 87–94.
[40] L. Reuben, D. Mba, Diagnostics and prognostics using switching kalman filters, Structural Health Monitoring 13 (3) (2014) 296 – 306.
[41] M. Kevin, Learning switching kalman filter models (1998).
[42] N. A. Mustafa, et al., Design of smart wearable system for sudden infant death syndrome monitoring, Ph.D. thesis, Sudan University of Science and Technology (2021).
[43] L. C. K. Reuben, D. Mba, Diagnostics and prognostics using switching kalman filters, Structural Health Monitoring 13 (3) (2014) 296–306.