USING MULTISENSOR DATA FUSION FOR ALLOCATING SYSTEMIC FAULTS OF THE HYDRAULIC CONTROL SUBSYSTEM OF A RUBBER MIXING MACHINE

Marko Orošnjak[1] [0000-0003-0929-1425], *Sandra Ramos*[2] [0000-0001-5788-1679]

[1]*University of Novi Sad, Faculty of Technical Sciences, Department of Industrial Engineering and Engineering Management, Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia*

[2] *Universidade do Porto Instituto Politécnico do Porto Instituto Superior de Engenharia do Porto, Rua Dr. António Bernardino de Almeida 431, 4249-015 Porto, Portugal*

Abstract: The rise of machine and deep learning algorithms in predictive maintenance has led to the influx of multidimensional data analysis studies. However, although studies dedicate towards increasing the accuracy of regression and classification models, many commit to resolving the issues by addressing a single fault mechanism, neglecting the latent degradation of other fault mechanisms. In this paper, we dedicate our efforts in understanding multiple and systemic faults through multidimensional data analysis. Using Knowledge Graph via Network Analysis we allocate markers of fault mechanisms that are used as features for fault classification. The features are extracted from discretised hydraulic power signal, hydraulic fluid physical and chemical data, and system response data. Using feature extraction we were able to observe latent degradation mechanisms that are used for multi label classification using machine learning algorithms. The results obtained show that neural network had highest, i.e., 85% accuracy (AUC = 0.88) among classification algorithms in allocating systemic faults within the hydraulic power system.

Keywords: predictive maintenance, multidimensional data, machine learning, network analysis, knowledge graph, systemic failures, hydraulic system

1. INTRODUCTION

Existing body-of-knowledge concerning failure mechanisms show conficting opinions regarding FPS (Fluid Power System). Many argue that particle (Jocanović et al., 2014) and air/water contamination (NORIA, n.d.) are the leading causes of stoppages, while some argue that overload and leakage are the most common causes of stoppage (Orošnjak et al., 2022). Furthermore, studies dedicated to analysing and understanding failure mechanisms show a lack of research related to multiple and systemic failures. Namely, most of the research performed dedicates to resolving the underlying mechanisms leading to a particular failure cause while neglecting the effects of latent degradation leading to a failure of other components within the system.

In such instances, many components within the system are left in an imperfect state with latent detrimental effects on the long run. For instance, in the case of a hydraulic pump, controlling and reducing the contamination by filter replacement will not reduce the risk of stoppage caused by pump-generated particle contamination that cause wear of valves and actuators down the return line of the system before the filter. Also, focusing simply on the contamination-induced wear will ommit degradation caused by poor operator manipulation of machine or system. Hence, it becomes a question whether the studies should still dedicate towards the uniform analysis of root causes or should studies turn towards resolving simultaneous and multivariate analysis of systemic failures and faults without production stoppages.

In this paper, we address the problem of systemic failures observed within the subsystem of the rubber mixing machines, specifically the hydraulic power system. As the system operates within an energy-intensive regime (Ragab et al., 2022), including around-the-clock production, there is higher theoretical probability that simultaneous and latent degradation will occur in different parts of the system. Hence, instead of monitoring the system and pinpointing a specific failure while performing diagnostics, we use multi-rate data fusion (Huang et al., 2021) for CM (Condition Monitoring) by allocating multiple deteriorating mechanisms.

The rest of the paper is structured as follows. The second section includes an in-detail description of the experimental installation, data acquisition and CM of the system. In addition, the section also presents descriptions of parameters and hyperparameters of used machine learning (ML) and Network Analysis

(NA). The third section presents the results and discusses the obtained results. Finally, the last section includes contributions to the literature, limitations, and future research.

2. METHODOLOGY

2.1 Experimental Installation and Data Acquisition

The experiment is performed on a hydraulic control system of a rubber mixing machine that performs movement for opening and closing the chamber. The hydraulic control system performs a hydraulic cycle for unloading the bulk material in three regimes: Opening Saddle, Idle Saddle and Closing Saddle. Data acquisition is performed through non-destructive measurements (Horvatic et al., 2016), including (1) inline lubricant condition monitoring (automatic particle counter, aqua sensor and turbine flow meter); (2) online system monitoring (SCADA) for measuring hydraulic cycles, idle time, and actuators' movement; (3) offline fluid condition monitoring (ICP-OES – Inductively Coupled Plasma Optical Emission Spectrometry and laboratory analysis) for elemental and physio-chemical analysis of contaminants present in the oil. Offline oil analysis and online lubricant condition monitoring include measurements of physio-chemical characteristics of oil: oil density, viscosity (at 40°C and 100°C), viscosity index, flame point, flow point, total acid number, water (ppm), water saturation (%), ISO4406 contamination level, and elemental analysis using ICP-OES. The hydraulic power signal is discretised at 0.1 seconds (f = 10 Hz) from the HYDROTECHNIK MultiHandy 2045 with measurements from 15.10.2021-12.12.2021.

2.2 Data processing and feature extraction

From discretised hydraulic power signal, the signal is split into opening saddle, idle saddle and closing saddle regime. From the discretised signal, each part is used for feature generation (Table 1), after which features are used for machine learning classification based on quasi-faults generated by the deviations in the signal. Although the system (excluding total failures) was performing in order, the anomalies noticed in the signal are classified as „quasi-faults". Quasi-faults are labelled via the concept of functional-productiveness (see (Orošnjak et al., 2023)). The signal sample contains 980 performed cycles (15.10.2021-12.12.2021). Hence, the sample contains n = 980 cycles and designated 52 features used for machine learning classification. The interpolation of offline measurements is performed with polynomial regression for at least R^2 > 0.95 to eliminate missing data the sample used for machine learning classification.

Table 1: Feature generation from discretised hydraulic signal

Time-domain feature	Feature notation[1]	Formula for feature generation
Mean (average) value	N_Mean_*X*S	$MEAN = \frac{1}{n}\sum_{i=1}^{n} N_i$
Standard deviation	N_StDev_ *X*S	$StDev = \sqrt{\frac{(x_i-\bar{x})^2}{n-1}}$
Root Mean Square	N_RMS_ *X*S	$RMS = \sqrt{\frac{1}{n}\sum_i x_i^2}$
Quartile ranges (Q1,Q3)	N_nQ_ *X*S	$nQ_{1;3} = x_{(k)} + a\left(x_{(k+1)} - x_{(k)}\right)$
Interquartile Range	N_IQR_ *X*S	$\mathrm{IQR} = \mathrm{N\_3Q\_}XS - \mathrm{N\_1Q\_}XS$
Peak-to-peak	N_P-P_ *X*S	$P\_P = \lvert \mathrm{N\_MIN\_}XS - \mathrm{N\_MAX\_}XS\rvert$
Skewness	N_Skew_ *X*S	$Skew = \tilde{\mu}_3 = \frac{\sum_i^n (x_i-\bar{x})^3}{(n-1)\cdot\sigma^3}$
Kurtosis	N_Kurt_ *X*S	$Kurt = \tilde{\mu}_4 = \frac{\sum_i^n (x_i-\bar{x})^4}{(n-1)\cdot\sigma^4}$

[1]*X*S can represent: **OS** for Opening Saddle regime; **IS** for Idle Saddle Regime; **CS** for Closing Saddle Regime.

2.3 Machine learning algorithms and parameters settings

The algorithms used for classification are Random Forest (RF), Decision Tree (DT), Gradient Boosting Classification (GBC), Support Vector Machine (SVM), Linear Discriminant Analysis (LDA) and Multilayer Perceptron (MLP) Neural Network. All algorithms are trained with 60% of data, using the holdout data of 20% tested, and for validation, 20% of data is used, splitting the data into 60/20/20. The following gives settings for the parameters and hyperparameters for all algorithms. The RF parameters are set: training data is used per tree at 60%, features are scaled, and splits are optimised with a maximum of 100 trees. The DT algorithm parameters are set: minimum observations per split are set at 20, minimum observations in the terminal are 7, maximum interaction depth is 30 and complexity penalty is set at 0.01 with scaled features. The GBC boosting parameters are set as follows: Shrinkage = 0.1, Interaction depth = 1, Minimum observations in node = 10, training data used per tree 50% with scaled features and optimised with a maximum of 100 trees. For the SVM, the settings are set as: weights = linear, cost of constraints violation = 1.0, tolerance of termination 0.001, epsilon parameter $\xi = 0.01$ with scaled features. The LDA parameters include scaled features, and for the relationships estimation MLE (Maximum Likelihood Estimation) method is used. For the MLP-NN, the activation function is a logistic sigmoid function with a backpropagation algorithm with a learning rate set at 0.05. The stopping criteria for loss function is set to 1 with 100000 maximum training repetitions. The network optimisation used is a Genetic Algorithm (GA) with a population size of 20 and 10 generations. Parent selection is a roulette wheel with reset mutations (10% probability), a uniform crossover method, and a fitness-based survival method with elitism set at 10%. Given parameters and obtained results can be replicated via open-source JASP.

2.4 Network Analysis for multivariate data visualisation

For the NA the FR (Fruchterman-Reingold) is the common algorithm used in network analysis (Leme et al., 2020), however, the *huge* network we used relies on Meinshausen-Buhlmann graph or the graphical lasso since they are building the *huge* network (Zhao et al., 2012). The network analysis allows a user to research data relationship and data complexity between features of interest. In our case, we extend the ML classification results with insights about features and data used for classification. As it is most often the case, especially with Random Forest, Decision Trees and Ensemble aglorithms such as the case with Gradient Boosting, some feature can be used for reaching >99% accuracy, even though the feature would not lead to meaningful conclusions about the case. Therefore, we use NA for gaining inisights about the system behavior and degradations that can lead up to systemic failures. It is also useful to emphasise that NA can sometimes be referred to as KG (Knowledge Graph) in literature (Xia et al., 2022), if it relies on multi-source data. The *huge* package in R (Zhao et al., 2012) is used for NA while the estimation criterion the EBIC (Extended Bayes Information Criterion) (Foygel & Drton, 2010) with tunning parameter of 0.5.

3. RESULTS & DISCUSSION

Using quasi-events for binary classification as „None“ and „Quasi-fault“, i.e., [0, 1], the results show high accuracy. However, even though classification performed well, suggesting the possibility of upcoming fault or stoppage, there is a lack of information regarding the part of the system that has degraded. However, from the observed discriminative learning algorithms (GCB, DT, RF), the feature importance suggests that only a few features are used to classify 100% (e.g., Flame point, N_1Q_OS, Viscosity). At the same time, the rest of them are excluded. In particular, the GCB algorithm suggests that only Flame point, N_Stdev_CS, N_1Q_OS and T3 are features used for classification, which is a misconception and can be misleading. The underlying reasons is that different features suggest different changes as percentage of degradation that require maintenance action. For instance, degradation of 10% of viscosity and density of oil would require preventive actions and suggest replacement of the oil. In other instances, rise of *Fe* or *Cr* particles with elemental analysis would sometimes require 10-fold changes (i.e., from 5 to 50 ppm) that would indicate changes in surface texture (Alar et al., 2010) leading to a wear. For that instance, it further requires pre-set maintenance plan that would indicate changes needing maintenance actions. For that instance, we use NA that would suggest changes in nodes and edges and indicate potential degradation of the system.

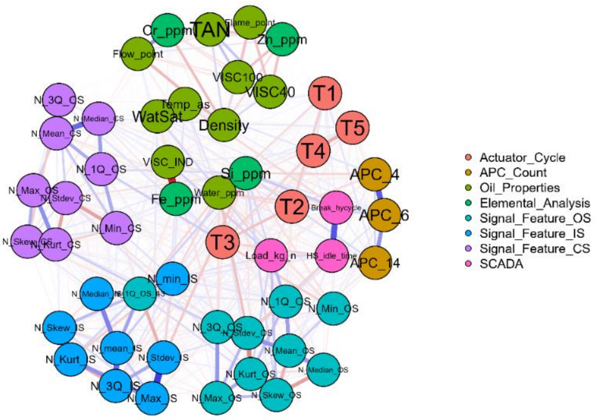

Figure 1. Network analysis of the system in the initial state with labels

Performing network analysis, we observe changes in the weight between the initial state (Figure 1) and in quasi-fault state (Figure 2), including clustering and labelling properties of the network.

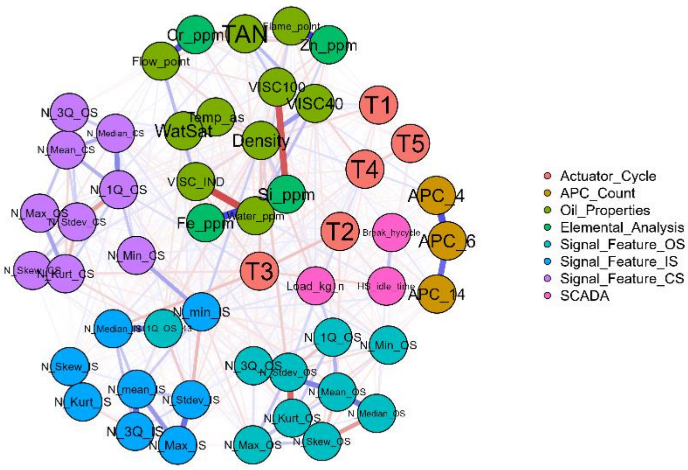

Figure 2. Network analysis of the system in the quasi-fault state with labels

Some nodes show high association with higher-order clusters such as oil and elemental analysis that cluster together. This can suggest that changes in oil physio-chemical properties cluster with elements rise within the fluid, which can indicate changes in not only fluid condition but the presence of wear, since aside from *Zn* and *Si*, rise in *Fe* and *Cr* particles can suggest starting wear of hydraulic pump. Following the statement, it is reasonable to assume that differentation between the states does not necessarily lead to wear. This is because different component degrade differently, and changes in percentage rise suggest different outcomes and not necessarily failure of a component. However, instead of using ML algorithms, such as discriminative ones, where for instance GCB or RF show perfect classification properties just by using 3% of the features extracted and neglect latent changes in other features, we use NA to gain insight about the changes in feature properties (Figure 3).

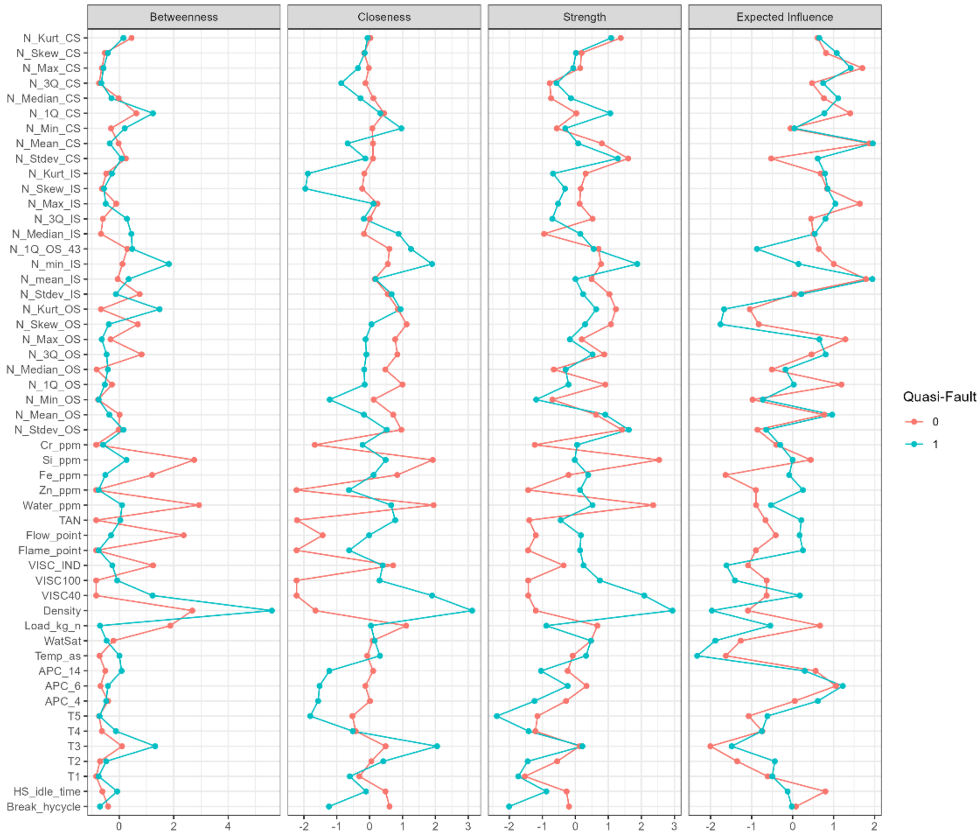

Figure 3. Centrality plot of conditions without and with quasi-faults

The strength and expected influence of the NA is important since it provides information about the sum of the weights of all edges connected to the particular node (strength), while expected influence provides information about local and global influence. For instance, we can see that density significantly changes between quasi-events suggesting high strength, i.e., weighted sum of the edges to a particular node. In the case of expected influence, the N_Mean_CS, N_Mean_IS, N_Max_CS have the highest influence, followed by N_Max_IS, N_1Q_CS, N_Max_OS, and N_1Q_OS, and APC_6. From a practical standpoint this was confirmed since significant variations in the hydraulic power signal is noticed that would indicate degradation and potential wear. However, since other parameters, like Fe, Cr, oil density, that would indicate presence of wear were not noticed (e.g., filtering, filter replacement, oil refilling) and left unoticed, the changes in signal helped improve diagnostic and decision-making.
Instead of using features for binary classification, the expert opinion for multiclass labels provided helped determine the degradation of components. Hence, using binary classification as quasi-events, we labelled quasi-faults as follows: y_i = C, where C = {None, Closing saddle, Saddle position, Sensor response, Total failure, Wear component}. The recorded total and sensor failures are used for classification from the experiment; the rest are added as labels. For the classification, we use a neural network with the following parameters. The rprop+ (Riedmiller & Braun, 1993) with logistic sigmoid activation functions are used. Stopping criteria loss function 1 with 100000 max training repetitions. Features are scaled and seed is set at 1234. Topology of the network uses GA (Hajnayeb et al., 2011) with 20 population size and 10 generations with roulette wheel parent selection and uniform crossover method. Reset mutations with probability of 10% and Survival method with elitism of 10% are used. The obtained results show good prediction properties with overall prediction of potential failures of 84% and 87.9% AUC (Area Under Curve) metrics (table 2).

Table 2. Prediction accuracy of neural network for systemic failures

Evaluation metrics	None	Closing saddle	Saddle position	Sensor response	Total failure	Wear component	Average
Accuracy	0.561	0.561	0.985	0.959	0.995	0.980	0.841
AUC	0.955	0.977	0.997	0.857	0.500	0.988	0.879

Based on the obtained classification we see that combinations of different changes in nodes and edges, can lead to different types of failure mechanisms. For instance, changes in system response and idle position of the saddle, can suggest slow response of the sensor. However, the biggest issues encountered were degradations of mechanisms that are not classified accurately. Namely, anomalies and spikes in the signal (presumably due to sensor disturbances) were classified as “None”. The underlying reasons include not statistically significant degradation of used features, which can lead to innacurate classifications. Hence, since there are numerous failure mechanisms that are left undetected or can be associated with different proposed failure mechanisms, the poor accuracy suggests that more background knowledge and associated underlying mechanisms need to be provided in order to increase the accuracy of the model.

4. CONCLUSIONS

The paper investigates the use of NA for allocating latent degradational mechanisms that can be used for labelling and condition monitoring. Namely, since ML models can indeed predict the outcomes with “perfect” classification results, this often leads to misscclassification and misinterpretation of the results. This is often the case with discriminative learning (e.g., RF, SVM, DT). In such instances, latent degradation of different components go by unnoticed since they are not captured as important features used for classifications. Therefore, we turn towards NA for observing changes in network structures, edge weights and centrality indices in understanding and observing potential systemic failure mechanisms that can go by unnoticed. The results obtained are used for gaining insights about the system behavior and from the observed distances metrics we generate multiclass features that would indicate existance of systemic failure mechanisms within practical hydraulic control system of a rubber mixing machine. The obtained results show classification accuracy and AUC metrics >80%, suggesting good results.
There are several limitations of the study. Namely, the NA is still in the infancy and interpretation centrality indices is not fully understood, especially considering the weighted undirected NA, such as this case. Also, the interpretation of expected influence and strength centrality measures have different mathematical estimation since in some cases the expected influence can be calculated using both betweenness and strength centrality measures suggesting overlap in the interpretation. Next, since huge package is used

with multidimensional data it is usually applied in high-dimensional statistics with sparse data, where there is higher features than sample sizes. Also, the partial correlation metric is used for estimating the edge weights to provide information how each pair of variables are related when the influence of all other variables are removed, which can be misleading in this case. Finally, there is only a handful of labels used, which is far from the actual potential failure mechanisms that potentially exist in the system.
The future studies will include expanding features and dataset with maintenance activity logs as qualitative data that can be used in generating high-dimensional knowledge graphs and gain deeper understanding about insights and system behavior that will initially increase classification accuracy and maintenance decision-making. The author will also in the future engage in graphical neural networks based on the expanded sample of the same system.

REFERENCES

Alar, V., Runje, B., & Baršić, G. (2010). Influence of surface texture on electrochemical potential. Der Einfluss vom Zustand der Oberflächenbeschaffenheit auf das elektrochemische Potenzial. *Materialwissenschaft Und Werkstofftechnik*, *41*(10), 875–878. https://doi.org/10.1002/mawe.201000632

Foygel, R., & Drton, M. (2010). Extended Bayesian Information Criteria for Gaussian Graphical Models. In J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems* (Vol. 23). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2010/file/072b030ba126b2f4b2374f342be9ed44-Paper.pdf

Hajnayeb, A., Ghasemloonia, A., Khadem, S. E., & Moradi, M. H. (2011). Application and comparison of an ANN-based feature selection method and the genetic algorithm in gearbox fault diagnosis. *Expert Systems with Applications*, *38*(8), 10205–10209. https://doi.org/10.1016/j.eswa.2011.02.065

Horvatic, A., Runje, B., & Butkovic, D. (2016). *Influence of Geometrical Magnification on Computed Tomography Dimensional Measurements* (pp. 0615–0622). https://doi.org/10.2507/27th.daaam.proceedings.090

Huang, K., Wu, S., Li, Y., Yang, C., & Gui, W. (2021). A multi-rate sampling data fusion method for fault diagnosis and its industrial applications. *Journal of Process Control*, *104*, 54–61. https://doi.org/10.1016/j.jprocont.2021.06.003

Jocanović, M. T., Karanović, V. V., Ivanišević, A. V., & Knežević, D. M. (2014). Hydraulic hammer excavator failure due to solid particle contamination. *Military Technical Courier*, *17*(1), 112–139. https://doi.org/10.5937/vojtehg62-4676

Leme, D. E. da C., Alves, E. V. da C., Lemos, V. do C. O., & Fattori, A. (2020). Network Analysis: a Multivariate Statistical Approach for Health Science Research. *Geriatrics, Gerontology and Aging*, *14*(1), 43–51. https://doi.org/10.5327/z2447-212320201900073

NORIA. (n.d.). *Air Contamination in Hydraulic Systems*. Machinery Lubrication. Retrieved June 6, 2021, from https://www.machinerylubrication.com/Read/28461/air-contamination-hydraulic-systems

Orošnjak, M., Brkljač, N., Šević, D., Čavić, M., Oros, D., & Penčić, M. (2023). From predictive to energy-based maintenance paradigm: Achieving cleaner production through functional-productiveness. *Journal of Cleaner Production*, *408*(April), 137177. https://doi.org/10.1016/j.jclepro.2023.137177

Orošnjak, M., Delić, M., & Ramos, S. (2022). Influence of Maintenance Practice on MTBF of Industrial and Mobile Hydraulic Failures: A West Balkan Study. In M. Rackov, R. Mitrović, & M. Čavić (Eds.), *International Conference on Machine and Industrial Design in Mechanical Engineering* (pp. 617–625). SpringerLink. https://doi.org/10.1007/978-3-030-88465-9_62

Ragab, A., Ghezzaz, H., & Amazouz, M. (2022). Decision fusion for reliable fault classification in energy-intensive process industries. *Computers in Industry*, *138*, 103640. https://doi.org/10.1016/j.compind.2022.103640

Riedmiller, M., & Braun, H. (1993). A direct adaptive method for faster backpropagation learning: the RPROP algorithm. *IEEE International Conference on Neural Networks*, *16*(3), 586–591. https://doi.org/10.1109/ICNN.1993.298623

Xia, L., Zheng, P., Li, X., Gao, Robert. X., & Wang, L. (2022). Toward cognitive predictive maintenance: A survey of graph-based approaches. *Journal of Manufacturing Systems*, *64*, 107–120. https://doi.org/10.1016/j.jmsy.2022.06.002

Zhao, T., Liu, H., Roeder, K., Lafferty, J., & Wasserman, L. (2012). The huge Package for High-dimensional Undirected Graph Estimation in R. In *Journal of Machine Learning Research* (Vol. 13).