Received December 18, 2020, accepted January 3, 2021, date of publication January 14, 2021, date of current version February 1, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3051583

Data-Driven Condition Monitoring of Mining Mobile Machinery in Non-Stationary Operations Using Wireless Accelerometer Sensor Modules

PABLO AQUEVEQUE[1], (Member, IEEE), LUCIANO RADRIGAN[1], (Student Member, IEEE), FRANCISCO PASTENE[1], (Member, IEEE), ANIBAL S. MORALES[2], (Member, IEEE), AND ERNESTO GUERRA[3]

[1]Electrical Engineering, Department, Universidad de Concepción, Concepción 4030000, Chile
[2]Department of Electrical Engineering, Universidad Católica de la Santísima Concepción, Concepción 4090541, Chile
[3]Department of Economics, Universidad Católica de la Santísima Concepción, Concepción 4090541, Chile

Corresponding author: Luciano Radrigan (lradrigan@udec.cl)

This work was supported in part by the Center for Industry 4.0 from the Faculty of Engineering of Universidad de Concepción, Chile, in part by ANID, Chile, and in part by CONICYT FONDEQUIP under Grant EQM150114. The work of Francisco Pastene was supported by CONICYT through the scholarship CONICYT-PFCHA/Doctorado Nacional/2020 under Grant 21202175. The work of Anibal S. Morales was supported by FIAEC N°1, Universidad Católica de la Santísima Concepción (UCSC), Chile.

ABSTRACT This paper presents the development of an easy-to-deploy and smart monitoring IoT system that utilizes vibration measurement devices to assess real-time condition of bulldozers, power shovels and backhoes, in non-stationary operations in the mining industry. According to operating experience data and the type of mining machine, total loss failure rates per machine fleet can reach up to 30%. Vibration analysis techniques are commonly used for condition monitoring and early detection of unforeseen failures to generate predictive maintenance plans for heavy machinery. However, this maintenance strategy is intensively used only for stationary machines and/or mobile machinery in stationary operations. Today, there is a lack of proper solutions to detect and prevent critical failures for non-stationary machinery. This paper shows a cost-effective solution proposal for implementing a vibration sensor network with wireless communication and machine learning data-driven capabilities for condition monitoring of non-stationary heavy machinery in mining operations. During the machine operation, 3-axis accelerations were measured using two sensors deployed across the machine. The machine accelerations (amplitudes and frequencies) are measured in two different frequency spectrums to improve each sensing location's time resolution. Multiple machine learning algorithms use this machine data to assess conditions according to manufacturer recommendations and operational benchmarks Proposed data-driven machine learning models classify the machine condition in states according to the ISO 2372 standards for vibration severity: Good, Acceptable, Unsatisfactory, or Unacceptable. After performing field tests with bulldozers and backhoes from different manufacturers, the machine learning algorithms are able to classify machine health status with an accuracy between 85% - 95%. Moreover, the system allows early detection of ''Unacceptable'' states between 120 to 170 hours prior to critical failure. These results demonstrate that the proposed system will collect relevant data to generate predictive maintenance plans and avoid unplanned downtimes.

INDEX TERMS Internet of Things, mining, non-stationary operations, data-driven, heavy machinery, condition monitoring.

I. INTRODUCTION

The safe operation of heavy machinery depends on many variables, from site design to proper work practices, and more importantly, having machine operators well informed about where other pieces of mine areas, obstacles, and personnel are about their equipment [1]. Situational Awareness (SA) is now an essential element of many mining safety programs across the industry. SA involves picking up information and cues from the environment, putting those pieces of information

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Imran.

together so operators can develop a good idea of what is going on, and then using it to predict what happens next. That is where technology can help. Many companies embrace new technologies to support and enhance the mining equipment operator's SA [2].

Typically, SA is interpreted as the mining machinery system's capability to interpret its surroundings and other agents' intentions. However, the internal system awareness is often not receiving the same R&D focus, even though any given critical mission's success is completely dependent on the condition of all the internal and external agents simultaneously. The internal system awareness in the form of vehicle health is the focus of this paper.

As the mining industry becomes increasingly automated, and vehicles become increasingly advanced, the need for condition monitoring and prognosis will continue rising [1], [3]. Condition-based Maintenance, or Predictive Maintenance, is a decision-making strategy using condition monitoring information to optimize heavy machinery availability [4]. Condition Monitoring (CM) enables the early detection of faults or failures to reduce downtime and operating costs, facilitate proactive responses, and improve the productivity, reliability, availability, maintainability, and safety (RAMS) of equipment [5] [6]. Rotating machines and main components of mobile mining machinery typically operate under conditions such as high load, high temperature, high moisture or dusty areas. Degradation in component's health is to be expected under these operating conditions. Unexpected breakdowns also can cause downtime and economic loss [7]. These situations cause accelerated degradation of critical components of heavy machinery, increasing the failure rates, and consequently, the repair costs up to 6 times and repair times up to 30 hours or more [8]. Unplanned maintenance in mining operations can reduce the availability of heavy machinery up to 37%. Correspondingly, the useful life of mobile mining machinery can be reduced by 25% [9]. In the worst case, improper maintenance programs can cause total loss of machine in less time than the manufacturer's lifespan.

Multiple machine breakdowns in mining can impact planned uptime up to critical conditions, resulting in impaired production targets downstream, with the cost ranging from \$100,000 to \$200,000 per machine per day [10]. Fig. 1 shows the relationship between maintenance cost, time to failure, and reliability in machinery. When the time to failure equals zero, the system enters a failure state. When the time to failure approaches zero, the system's failure rate increases, and the reliability and mean time to failure decreases, correspondingly [8]. Therefore, it is critical to have a well-developed CM program to improve non-stationary heavy machinery reliability.

Unforeseen failures of mobile machinery used in the mining industry significantly impact availability, overall equipment efficiency, and productivity. Due to the lack of on-line monitoring during machine operations, it cannot be verified that operators are using machines correctly or adequate maintenance is being performed according to manufacturer guidelines. These situations cause accelerated degradation of critical components of mining machines, increasing the failure rates and consequently the repair costs up to 6 times and repair times up to 30 hours or more.

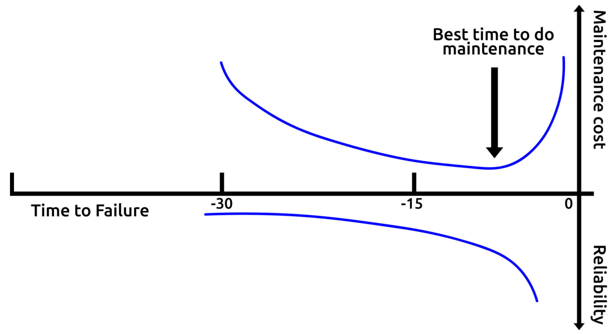

FIGURE 1. **Typical curve Maintenance cost versus Time to Failure for a heavy machinery.**

Three main research topics can be identified in the state of the art of condition monitoring solution for heavy machinery: (*i*) improvement of data processing techniques (detection, diagnosis and prognosis algorithms), (*ii*) optimization of maintenance planning and scheduling, and (*iii*) development of hardware solutions for data acquisition in harsh environment conditions. The literature review indicates that areas (*i*) and (*ii*) are the ones with the greatest R&D. Most references in (*i*) use existing data and focus on the development of new algorithms to achieve better accuracy or true-positive rate for machine health status classification. Another relevant challenge identified is the improvement of machine health status predictions using small datasets with limited historical information content. It should be pointed out that the survey and gathering of industrial machine labelled data is a complex task in mining conditions.

On the other hand, most references in (*ii*) use machine health status diagnosis derived from existing data and focus on improving maintenance strategies to extend assets lifespan and bring more productivity and economic benefits using ROI evaluations considering different maintenance plans and schedules. Commonly, these developments use data obtained through Multiphysics FEM/DEM modelling. These evaluation techniques help to provide fault diagnosis under specific boundary conditions. The use of multiphysics modelling techniques are typical for stationary machine assessment and fault diagnosis, where physical behaviour under different operating conditions are extensively reported.

CM is a broad term referring to the systematic process of data collection to evaluate asset's performance, reliability, and maintenance needs to plan repair and maintenance backlogs. Its main purpose is potential failures finding. It requires the collection of good quality asset's health data which trending is studied. The primary advantage of CM is that it incorporates health indicator monitoring activities performed while the machine is operating. Assets failures are predicted well in advance of their occurrence with the help of data-driven models. Also, machine parameter data trending allows extending assets operation as close as possible to their actual useful life. CM data provides vital information for taking important decisions affecting machine fleet operation goals. Maintenance decisions are taken based on the actual asset condition avoiding unnecessary repairs leading to startup failures. Catastrophic failures of critical assets presenting accelerated wear trends also can be avoided by using CM tasks. Sometimes, operating conditions change, causing life expectance to reduce as noted by steeper indicators trends leading to unexpected catastrophic failures. CM can detect this and earlier planned shutdowns can avoid such disasters. To achieve all these technical capabilities, it is mandatory to develop hardware, firmware and software solutions that meets industrial standards for equipment information gathering.

This paper explores data-driven methods to estimate mobile mining machinery's fault condition and health status in non-stationary operations using a novel design for low-cost wireless accelerometer sensor modules installed on-board machines for on-line condition monitoring. This work performs condition monitoring analysis by utilizing the information and signals gathered from proposed wireless 3-axis vibration sensors to make assessments of the current machine condition and tasks. The mining industry is characterized by a small series of highly specialized machines, which challenges the possibility to use traditional fault detection, diagnosis and prognosis solutions.

In addition to the proposed novel wireless vibration sensor network for condition monitoring of non-stationary heavy machinery, this paper aims to develop early fault detection algorithms using machine learning techniques. Machine health information can support routine maintenance tasks, it is an essential input to decision-making for diagnosis and prognosis systems and further operation planning, i.e., how to run the machine for minimum wear and damage while maintaining other mission targets.

This paper proposes a real-time monitoring system for non-stationary heaving machinery. The system consists of several vibration sensors (accelerometers) located at the machine's critical points, communicating with each other through a Bluetooth Low Energy (BLE) network. A central wireless hub sends the condition monitoring information using a 3G network to an external server, accessed by a web platform for O&M analysis purposes. Different machine learning data-driven algorithms for predicting the machine health status based on machine working time, and the vibrational severity gathered from real-time sensor information are proposed.

This work provides results for both, a rugged hardware design to acquire high-quality and reliable data in real-time from heavy machinery, and machine learning algorithms to diagnose the machine health status using measured data on the industrial environment (with accuracy over 90%). Design of hardware, firmware, and data-driven models is carried out to cover non-stationary heavy machinery operations' technical requirements. Correspondingly, features like high energy autonomy and robust wireless communication are key values of the solution. Proposed low-energy consumption hardware design for wireless sensors has IP67 protection to surpass mining harsh environmental conditions; it is a light and small size to ensure compatibility with typical O&M conditions. Wireless sensors include high magnetic field attachment system for fast and easy deployment on heavy machinery structures. This paper addresses the following challenges for condition monitoring: (i) how to acquire reliable data from heavy machinery in real time correctly, and (ii) how to use the measured data in data-driven machine learning algorithms for early fault detection and prediction of the machinery's health status.

This paper is organized in eight sections: (i) Introduction, (ii) Maintenance philosophies & strategies, (iii) Condition monitoring of mining machinery, (iv) Data-driven approaches, (v) Solution description (vi) Results, (vii) Conclusion (viii) Future scope.

II. MAINTENANCE PHILOSOPHIES AND STRATEGIES

The field of maintenance can be divided into three main areas. The most fundamental is reactive maintenance, where errors are fixed as they arise, known as breakdown maintenance. This can be a good option for failures, not causing considerable production loss or damage. However, it is a low option for most critical components given its unpredictability, causing a sudden loss of functionality and unplanned detentions [11]

On the other hand, preventive maintenance is when parts are exchanged based on some parameter different from failure, such as the machine's running hours. The replacement of parts is performed at set intervals. The benefits include a reduced number of unplanned stops, i.e., reduced risk of failure and secondary damage, and reduced degradation of critical components. The main drawback is a high, potentially unnecessary cost for parts and labour since different individual machines do not deteriorate at the same rate [12].

Condition-based and/or predictive maintenance is maintenance based on some measurable parameters on the individual machine. Ideally, operators can measure some degradation parameters and then change parts necessary at the most convenient time right before failure. Thus, minimum maintenance resources can be used without compromising reliability and availability. A good review of such methods in predictive maintenance is closely related to prognostics [13].

Fig. 2 shows a summary of the maintenance strategies used for preventive maintenance. Comprehensive preventive maintenance in modern mining operations include activities of corrective maintenance (run-to-failure), routine maintenance (scheduled approach), random maintenance (opportunity-driven), and predictive maintenance based on condition monitoring and failure prediction [14].

The field of prognostics is defined by the standard ISO 13381-1:2015: E as "*analysis of the symptoms of faults to predict the future condition and residual life within design parameters*." [15]. The estimated time to failure (*ETTF*) is

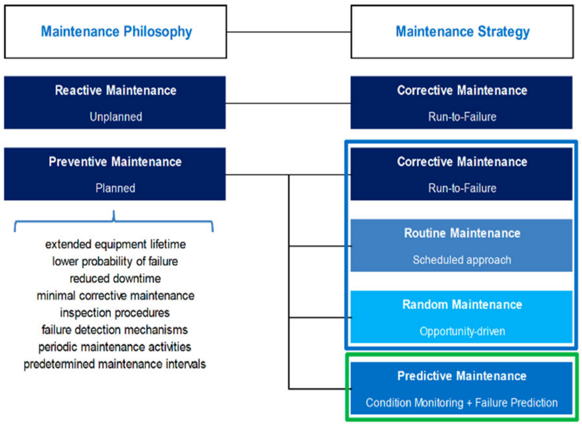

FIGURE 2. Summary of the maintenance strategies used for preventive maintenance.

the time from the point of prediction (t_{pred}) until the estimated failure time. The remaining useful time (*RUL*) is the time between t_{pred} and the end of life. The prognosis relies on some measured signal from the asset, together with a model relating this signal to a deterioration process. There are many ways to develop such models, but in principle, they can be divided into three main groups: experience-based models [8] [16], physics-based models [17], and data-driven models [18].

Data-driven models relieve the need for physical understanding by using measurement data to find the relationships between sensor signals and damage. A good review of data-driven statistical approaches is given by [19]. Benefits and drawbacks are as follows: (i) there is no need for physical models of degradation and dynamics, (ii) the approach can capture unknown failure modes, (iii) the available techniques are not specific to a certain domain, i.e., the methods can be transferred to different applications, (iv) data from an operational machine is required, model development can only be done after machine is built and used, (v) a lot of data is typically required, and also run-to-failure data, which is often expensive and in some cases dangerous to obtain; (vi) the approach suffers from the fact that error is rare events, leading to too unbalanced data sets, which are hard to learn from, (vii) the learned relations rarely take into account of causality, i.e., the direct relation between problems and symptoms, (viii) anomaly detection on real-time through machine learning techniques [1].

Using their definitions, this paper's main work can be categorized as a data-driven model to infer a measurement model and use of a data-driven model to predict the time to failure and prognosis.

III. CONDITION MONITORING OF MINING MACHINERY

A condition monitoring system can increase the RUL of the asset while keeping uptime and high production quality. Currently, most industries are working in the term called "Time on Tools," which means improving the maintenance process based on statistical analysis and real-time asset measurements. This way, "Time on Tools" allows increasing the RUL of the asset, also increasing overall labour effectiveness. Nowadays, technological advancements make it feasible to collect data from many industrial processes and apply the "Time on Tools" concept [20].

State of the art on condition monitoring of heavy machinery can be divided into two groups: the Original Equipment Manufacturer (OEM) and the Original Technology Manufacturer (OTM) solutions. The OEM solutions refer to solutions for some product which are developed by the own product developer. On the other hand, OTM solutions refer to some products developed for other companies.

Within the OEM solution for heavy machinery are the Engine Control Modules (ECMs) [21]. These systems use actuators like high-pressure pumps, cylinder injectors, among others. Typically, the ECM controls four subsystems in heavy machinery: air-fuel rate, idle speed, valves synchronization, and turn on synchronization. The ECMs receive abnormal signals from the OEM sensors, and they deliver a failure alert, tagged with a code. So, these devices work under reactive maintenance logic, which is not enough to avoid losses associated with catastrophic failures.

Additionally, OEMs have developed some improvements to measure in real-time when a component of the engine fails [22]. However, these solutions rely on the same reactive and non-predictive fault detection strategy. ECM-based solutions exhibit a TRL 8 or 9.

Other OTM solutions with a TRL 4 or less are outlined below. Oppenheimer *et al.* proposed a physical model to predict shafts cracked in a machine. The proposed system uses a combination of machine-fault models and measured machine signatures to identify and classify the machine's state. Some issues were found related to the identification of the machine's loads in measured data, which impaired the machine state classification [23].

A logistic regression model was used for calculating the probability of failure of a machine using given condition variables. Also, the remaining useful life of the asset was estimated using the proposed approach [24]. The system achieves promising results in health predictions for an elevator system. However, tests were not reported for heavy equipment case. Then, not enough data is available to validate its accuracy.

Gebraeel *et al.* [25] proposed a neural network-based degradation model to calculate the residual life distribution of partially degraded components. In this solution, predicted failure times are estimated using a dynamic Wavelet neural network using real-time sensory signals (vibrations). These estimations are used to derive prior failure time distribution for the component being monitored. This system was tested for stationary machinery with good results (error = 7.56%).

Other solutions have been developed for machinery monitoring. In [26] a new wireless sensor network for indoor industrial monitoring to optimize the data packet and energy consumption with high reliability was developed. This solution focused on optimizing network protocol and topology to support the condition monitoring of industrial machinery.

TABLE 1. Commercial solutions for machinery monitoring.

Item \ Solution's Company	System 1	System 2	System 3	PROPOSED SYSTEM
Effective work time measurement	Yes.	Yes.	Yes.	Yes.
Vibration severity assessment	No.	No.	No.	Yes.
Fuel consumption	Yes.	Yes.	Yes.	Yes, optional feature.
Asset Geolocation	Yes.	Yes.	Yes.	Yes, optional feature
Operation cost optimization through machine learning and unstructured data analysis	Yes. Calculation method not disclosed	Yes, method based on reliability data.	Yes, the method based on reliability data.	Yes
Preventive maintenance plans through machine learning and unstructured data analysis	Yes. Calculation method not disclosed.	Yes, method based on reliability data.	Yes, method based on reliability data.	Yes.
Predictive maintenance plans through machine learning and unstructured data analysis	Yes. Calculation method not disclosed.	No.	Yes, method based on control curve delimitation.	Yes.
Supplies and spare parts purchase plans through machine learning and unstructured data analysis	No.	Yes, method based on reliability data.	No.	No, under R&D.
Suitable for mining heavy equipment monitoring	Yes.	Yes.	Yes.	Yes

This work makes more evident the relevance of the communication network and protocol for condition monitoring purposes.

Several works addressed the instrumentation of critical mining machinery with external condition monitoring systems. These solutions improved the OEE of machinery operating under stationary conditions such as conveyor belts, crushers, ball mills and vibratory screens. The solutions proposed in these works aims to improve the OEE while reduces measurement noise in electronic instrumentation [27]–[30].

Another work developed a system with an accelerometer and a microphone to assess a diesel engine's health status. In this work, the vibration magnitude (RMS value) and the vibration frequency spectrum obtained from 3-axis accelerometer signals were used for condition monitoring and fault analysis [31]. This system was developed only for diesel engine monitoring. The early fault detection in other critical components using the proposed solution is not reported.

Table 1 shows a feature comparison between the existing commercial OTM solution (TRL 7-9) to monitor heavy machinery. These systems use a plurality of sensors, typically temperature, pressure, and Hobb meter sensors for performing the condition monitoring. None of the solutions in Table 1 provides health status predictions based on the heavy machinery's reliability. All solutions in Table 1 establish empirical control curves based on physics-models and experience-based models. These control curves are used to trigger failure alarms using fixed thresholds according to the machinery's condition variables. Thus, Table 1's monitoring systems are not based on the machinery's current state for maintenance diagnostics. These systems provide a diagnosis based on hours of work and do not quantify the intensity of work.

In summary, still, there is an opportunity for R&D in condition monitoring systems for heavy machinery. A number of technology gaps were identified as mandatory to reach the TRL level demanded by the application: (*i*) smaller size of sensors, (*ii*) device's weight reduction, (*iii*) increase of energy autonomy, (*iv*) avoid wiring for instrumentation and/or power, (*v*) smart wireless communications, (*vi*) improve IP protection to withstand harsh environmental conditions, (*vii*) proper attachment system for easy and fast deployment, (*viii*) avoid drilling or welding for device attachment to machines, (*ix*) short-term training of data-driven models, and (x) use of small datasets with limited historical information content for the training of predictive condition monitoring models, among others.

IV. DATA-DRIVEN APPROACHES

Data-driven approaches addressed in this paper contain both statistical and machine learning techniques. These techniques include Artificial Neural Network, Particle Filtering, Neuro-Fuzzy Inference system, Hidden Markov model, Gaussian Process Regression, Match Matrix, Fuzzy Logic, Extended Kalman Filtering, and Support Vector Machine (SVM).

Fig. 3 shows the different requirements for each technique to achieve adequate operation of prediction models [32]. Gaussian process regression and Match matrix demand heavy computational time while Particle Filtering, the hidden Markov model, Artificial Neural Network, and Neuro-Fuzzy systems require considerable amounts of historical data to perform prognosis. These requirements limit the applicability of these techniques to real-world implementation for real-time prediction.

Heavy historical data requirement is not desirable for the application since historical data is not always available or well labelled. The use of unreliable and/or low-quality data

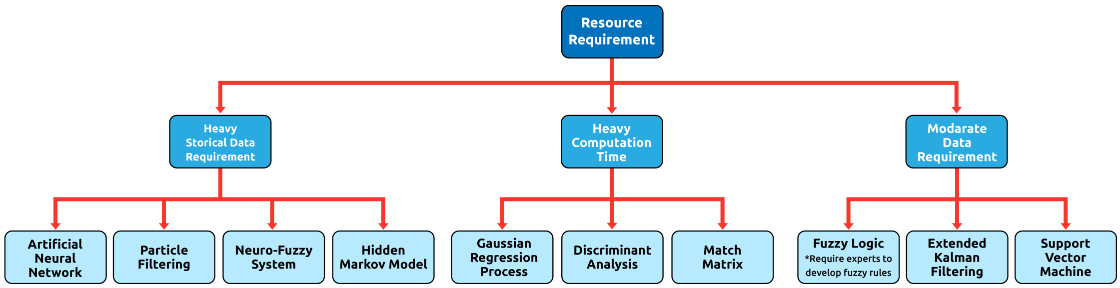

FIGURE 3. **Machine learning models classification according to their specific requirements.**

results in unreliable, inaccurate forecasts. It causes false alarms and unnecessary machine downtime [32], [33]. On the other hand, heavy computation time can be avoided using modern powerfull multi-core computers/servers and/or cloud computing. Despite the processing hardware available on premise or in the cloud, the most suitable algorithms for the application will be the ones with moderate data required to ensure real-time operation when scale-up the number of machines being monitored.

Considering the design boundary conditions of data requirement and computation time, the best machine learning algorithms for the application should not require a significant computational capacity, must be able to provide good predictions with small datasets, and be easy to implement as industrial monitoring systems [34]. The algorithms that meets these characteristics are [35], [36]:

(i) Support Vector Machine (SVM): this technique has the goal of finding a hyperplane in N-dimensional space (with n data classes) that separates the data. In most situations, this can have multiple solutions. However, this algorithm finds the one with the most significant margin to the data classes. Thus, higher reliability in the classification is achieved. To maximize the margin, support vectors are used, which correspond to the data closest to the hyperplane. These support vectors will condition the position and orientation of the hyperplane. A cost function is calculated, from which its partial derivatives are obtained to optimize using a gradient-based algorithm finally. Table 2 shows the advantages and disadvantages of the SVM algorithm [37] [38] [39].

(ii) Naive Bayes: it is a non-parametric and Bayesian approach process that provides uncertainty measurements to predictions. Unlike other classification approaches, the Bayesian infers the probability distribution of the data, starting from an initial probability and recalculating it with the data's evidence using the Bayes rule. Thus, it weights each possible prediction with its last calculated probability distribution from the training to predict unknown data. Using Gaussian regression calculates the probability distribution over all potential curves that fit the data. Equal starts from a probability distribution and recalculates it with the data. Normally, one begins by assuming a Gaussian distribution of the data. However, knowing the input data, specific parameters can be preset to improve the regression process, such as the mean, covariance, and the data.

TABLE 2. **Main features of SVM.**

Advantages	Disadvantages
• Flexible, probabilistic non-parametric technique that offers prediction of uncertainty through the variance around the mean prediction of the bayes theorem • High adaptability to handle high-dimensional data and can achieve accurate prediction even when the sample size is small	• High computational cost •There are numerical stability problems in restricted quadratic programming. • The parameters need to be tailored specifically to the problem at hand and this can be difficult.

Table 3 shows the advantages and disadvantages of the Naive Bayes algorithm [40]–[42].

(iii) Discriminant Analysis: Pattern recognition and machine learning algorithm that finds a linear combination used to build up predictive models and forecast the group to which an observation belongs based on specific characteristics, that is, to identify the characteristics that differentiate and define its profile.

Table 4 shows the advantages and disadvantages ofDiscriminant analysis [43].

V. SOLUTION DESCRIPTION

The proposed system is a comprehensive solution that comprises hardware and firmware for data collection, and software for O&M data analysis of heavy machinery in non-stationary operations.

Each heavy machinery has its own vibratory patterns [44], which can help to detect machinery's health status in real-time. The proposed solution uses several wireless accelerometer sensor modules located at critical points in the machinery. The solution requires sensing the heavy machinery in a non-invasive way to obtain acceleration signals during normal operation. The sensor modules (named Bluetooth accelerometer sensors) use a dedicated BLE wireless network for data transmission [45], [46]. The distributed sensor network considers one hub module (named gateway publisher) for backhaul of vibrations readings from Bluetooth accelerometer sensors using a star network topology [45]. The hub module concentrates the measurements from all the sensor modules of the network, and send this information through a 3G/4G/LTE network to the system server connected

TABLE 3. Main Features Naive Bayes.

Advantages	Disadvantages
• Can calculate degradation prediction for dynamic classification problems.	• High computational load, especially when training data sets are large
• Flexible, probabilistic non-parametric technique that offers prediction of uncertainty through the variance around the mean prediction of the Gaussian process model.	• Assume that all points are normally distributed and that the error between each point is correlated
• High adaptability to handle high-dimensional data and can achieve accurate prediction even when the sample size is small	• It isn't easy to find optimal values of the scale parameters
• Allows non-parametric learning of a regression function from noisy data, avoiding simple parametric assumptions.	• Assume that the noise in the training data is constant throughout the input domain.

TABLE 4. Main features of Discriminant analysis.

Advantages	Disadvantages
• Robust and accurate results with large input data.	• There is no standard method for choosing the kernel function.
• High precision with maximized decision limit.	• There are numerical stability problems in restricted quadratic programming.
• Efficient for small or large data sets and real-time analysis.	• It is difficult to construct a univariate time series for the remaining life and the sampling time.
• It is successfully applied to the diagnosis of machinery failure	• The parameters need to be tailored specifically to the problem at hand and this can be difficult.
• Good generalization performance on a limited number of learning patterns	

to network backbone on the mine site [46], [47]. The data is stored in a SQL compatible database server, which is accessed by a web platform for O&M monitoring and analysis purposes. Table 1 shows a comparison between three commercial condition monitoring solutions and proposed system. The developed solution exceeds the features of commercial system studied.

A. GATEWAY Publisher–Hub MODULE DESIGN

The design of the Hub module of the sensor network or Gateway Publisher (GP) meets different technical and functional requirements to ensure embedded capabilities: (*i*) network's management, (*ii*) data packets routing and traffic control, (*iii*) power supply unit, (*iv*) power management unit, (*v*) data processing unit, and (*vi*) sensing unit. Fig. 4 shows the GP device functional diagram. It should be pointed out that the hub module includes the same embedded sensing capabilities that sensor modules. Each functional block embedded in the GP device performs the following tasks:

1) PROCESSING SYSTEM

This functional block is responsible for coordinating the communication with Bluetooth Accelerometer Sensor (BAS) modules using BLE links. The BAS devices' vibration

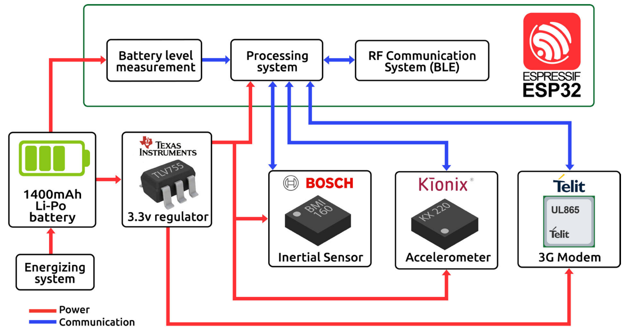

FIGURE 4. GP device functional diagram.

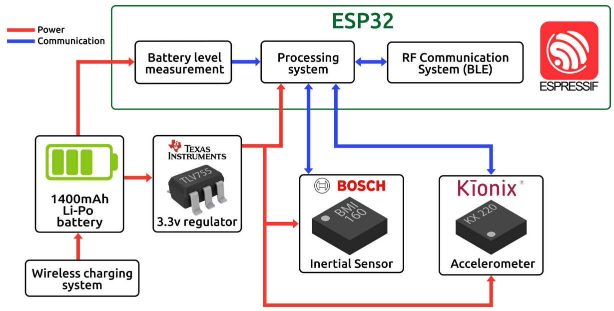

FIGURE 5. BAS device functional diagram.

information is sent to the WEB server through the 3G/4G/LTE module.

2) COMMUNICATION SYSTEM

This functional block comprises two elements, the BLE controller, which communicates using a star topology to the BAS units. The second element is the 3G/4G/LTE controller, which establishes communication with the WEB server.

3) POWER MANAGEMENT SYSTEM

The primary power source of the Hub modules is the on-board power system of the heavy machines. A voltage regulation stage is then used to supply a LiPo battery to provide autonomy to the system if the machine power fails.

B. BLUETOOTH ACCELEROMETER SENSOR DESIGN

The design of the BAS measuring devices of the sensor network considers the following embedded capabilities: data storage and processing functionalities, wireless communication, 3-axis inertial measurement, accelerometer, power supply, and power management. Fig. 5 shows the BAS device functional diagram. Each available block embedded in BAS devices performs the following tasks:

1) PROCESSING SYSTEM

Its function is processing and managing the vibration signals coming from the IMU sensor (Bosch BMI160) and the accelerometer (Rohm Semiconductor Kx220). The BAS device uses a combination of IMU and accelerometer to extend the frequency range without lowering its resolution.

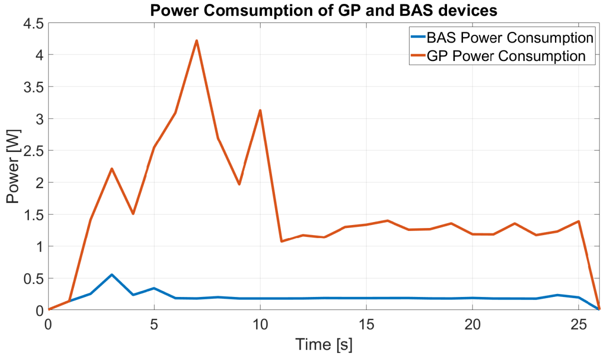

FIGURE 6. Power consumption of BAS and GP devices during the communication tests.

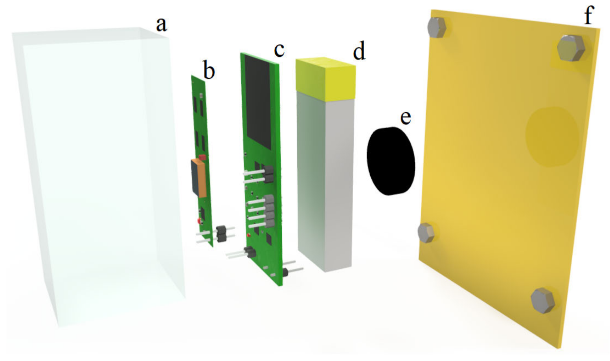

FIGURE 7. 3D model for GP and BAS devices. a) Epoxy Resine. b) Wireless charging system (only in BAS device), c) Hardware (GP and BAS devices), d) LiPo Battery, e) high-field magnet f) piece of heavy equipment body and/or chassis.

Thus, the system achieves a total frequency range of 1-5000 [Hz] for inertial signals measurement. Some parameters are calculated in the edge BAS device from the acceleration signals, which are described in section VII.D.

2) RF COMMUNICATION SYSTEM

This functional block is responsible for the data transmission of estimated parameters of inertial signals measured by the BAS device to the GP device or hub module of the sensor network using BLE.

3) POWER MANAGEMENT SYSTEM

An embedded 3.7V 1400mAh LiPo battery is used to power the systems of each BAS device. The proposed design considers battery wireless recharging.

C. DEVICES IMPLEMENTATION

1) ELECTRONICS ENGINEERING

Prototypes of GP and BAS devices were implemented for functional performance testing and evaluation. During lab testing, measurements of power consumption were accomplished using Fluke 287 multimeter (Fluke Corporation, Everett, Washington, USA) and the TDS1002C-EDU Tektronix Oscilloscope (Tektronix, Beaverton, Oregon, USA). The test protocol was as follows: (*i*) both devices start to measure inertial signals simultaneously, (*ii*) then establishes a BLE connection between each other, (*iii*) after BAS device complete a measurement cycle, it sent the vibration measurement parameters to the GP device, (*iv*) when the GP receives the BAS devices' information, it starts the 3G connection to send the information to the application web server, (*v*) finally, the BLE connection ends when the last data sent triggers the 3G connection's ending in the GP device.

Fig. 6 shows the power consumption of the GP and the BAS modules. The average power consumption of the GP is 468 mA and 58 mA for the BAS. Thus, an energy autonomy of 3 hrs and 24 hrs is achieved in the devices, respectively.

2) MECHANICAL ENGINEERING

To ensure an IP67 protection against harsh environment conditions in mining, an Epoxy resin's block of $90 \times 40 \times 29$[mm] size for the GP device and $70 \times 22 \times 30$[mm] for BAS device were manufactured. Fig. 7 shows a 3D assembly model for GP and BAS devices.

D. DATA PROCESSING

After acceleration signal acquisition, the RMS value of accelerations is calculated, and the frequency spectrum analysis of measured signals is performed in each sensor module (GP and/or BAS). A Fast Fourier Transform (FFT) is applied to the vibration signals for frequency spectrum analysis. The frequency spectrum analysis is implemented on the microcontroller embedded on each sensor module. Thus, the FFT uses a one-minute window for each frequency spectrum analysis. The data processing algorithm stores the five highest predominant frequencies with their respective amplitudes in 3-axis.

The BMI160 and KX220 accelerometers of the GP and BAS devices measure at a frequency of 500Hz and 5000Hz, respectively. This data is stored in a buffer per one minute to be processed. Every 15 minutes, the calculated indicators (1 per minute, 15 indicators per sampling cycle) are averaged and sent to the GP for further processing and analysis at the application web server. This strategy is chosen to decrease data traffic between sensor network devices.

E. COMMUNICATION PROTOCOL

The main challenge to design the sensor devices is LiPo batteries' limited energy capacity in BAS devices.

To handle the energy restriction, the BAS device first measures the sensors' variables and then pre-process the information to assemble a package with less data than the originals. This way, fewer transactions using BLE connection are needed, which results in less power demand. Once the pre-processing is complete, the BLE is enabled, and the data is sent; finally, the BLE is disabled.

The data packet transferred from a BAS to a GP is 78 bytes length. This data packet comprises 4 bytes for the header and 74 for the accelerometer's pre-processed data. Table 5 shows the data packet distribution. It should be noted that a BLE transaction (data packet of Table 5) is made every 15 minutes.

TABLE 5. BAS data packet specification.

Item	Content	Bytes	Data type
ID header	ID Slave Device	2	UINT8
	Long msg	2	UINT8
BMI160	RMS per axis (x, y, z).	12	UINT8
	Five frequencies with major amplitudes per axis (x, y, z).	10	UINT8
	Five major amplitude per axis (x, y, z).	10	UINT8
KX220	RMS per axis (x, y, z).	12	UINT8
	Five frequencies with major amplitudes per axis (x, y, z).	20	UINT8
	Five major amplitude per axis (x, y, z).	10	UINT8
Total			**78**

The GP or Hub module of the sensor network act as a gateway for the BAS modules, so the GP's BLE is always enabled. GP device is always connected to a 3G network to send the BAS data to a remote server for further processing and analysis. Due to these two conditions, the GP device is connected to the machine power system with a backup battery (see section V). The data packet that GP module sent to the server over the 3G network has a minimum length of 84 and a maximum length of 552 bytes. The data packet length depends on the amount of BAS modules connected to the GP, which depends on the heavy machinery under evaluation. Table 6 specifies the data packet distribution sent by the GP device to the remote server. The data packet comprises 4 bytes for header and 80 bytes (up to 548 bytes) for pre-processed data. This pre-processed data is constituted by the machine effective working time and the vibration data from the accelerometers.

F. HEALTH STATUS DETECTION ALGORITHM

This research proposes a cost-effective solution for implementing a vibration sensor network with wireless communication and machine learning data-driven capabilities for condition monitoring of non-stationary heavy machinery in mining operations without interference with existing WiFi 2.4/5GHz networks. During machine operation, 3-axis accelerations are measured using a plurality of sensors deployed across the machine. The machine accelerations (amplitudes and frequencies) are measured in two different frequency spectrums to improve each sensing location's time resolution. In this paper, multiple machine learning algorithms use these machine data to assess condition according to vendor recommendations and operational benchmark. Proposed data-driven machine learning models classify the machine condition in a state according to the ISO 2372 standards for vibration severity: Good, Acceptable, Unsatisfactory, or Unacceptable. The parameters described in section VII are used as input for machine health status classification.

The Good state indicates that the machine is working in conditions that do not impair asset lifespan.

TABLE 6. GP data packet specification.

Item	Content	Bytes	Type	Character
ID header	ID Master Device	2	UINT8	Required
	Long msg	2	UINT8	
Machine working hours		2	UINT8	
Vibration sensor data	Master measurement data packet	78	UINT8	
	Slave 1 measurement data packet	78	--	Optional
	Slave 2 measurement data packet	78	--	
	Slave 3 measurement data packet	78	--	
	Slave 4 measurement data packet	78	--	
	Slave 5 measurement data packet	78	--	
	Slave 6 measurement data packet	78	--	
Minimum total		**84**		
Maximum total		**552**		

The Acceptable state indicates work conditions in which the machine's useful life is not significantly reduced. The Unsatisfactory state indicates that the machine can operate for a limited period in the current operating condition until the need for corrective action arises. The Unacceptable state indicates that the machine will suffer a critical failure in short, maintenance is mandatory immediately.

The machine learning algorithms chosen for machine health status classification during non-stationary operations are the following: Naïve Bayes, Support Vector Machine (SVM), and Discriminant Analysis.

The three algorithms were tested considering a prior optimization process of method's parameters, kernel, support, and type with different combinations to achieve a better possible classification. The protocol to train the machine learning algorithm was the following:

1) Test an algorithm with default parameters.
2) Change the kernel and choose the best.
3) Change support.

The training protocol was performed using the Matlab software classification learner toolbox (MATLAB R2020a, Mathworks, Inc., Natick, MA, USA).

VI. RESULTS

For testing the proposed system, two backhoes machines were used. These backhoes were model 3CX ECO 15FT from the JCB manufacturer (JCB, Rocester, UK). The measurements in both industrial machines were performed during 60 continuous working days. Both backhoes integrated the same machine fleet. Both machines were used as earthmovers in road maintenance activities in a mining site located in Chile. The measurements elapsed by 60 working days until one of the backhoes suffer a diesel engine critical failure. The measurement survey end for the two machines for data labelling and standardizing after this unplanned downtime. Collected data was analyzed and classified by an expert in heavy machinery maintenance and condition monitoring (data manual labelling).

For practical purposes, only the vibration data from a BAS module installed in the machine for diesel engine monitoring is discussed in this section. However, it is important to note that the vibration data from any BAS device installed to monitoring the machine has the same structure.

During field test and measurements, the backhoe that suffers the critical breakdown (machine #1) was labelled by the condition monitoring expert into four states of vibratory severity during the 60 days trial. On the other hand, the backhoe without critical faults during the 60 days trial (machine #2) was labelled by the expert only into the states Good and Acceptable.

A. GENERAL FRAMEWORK

The proposed approach is simple but efficient. It uses the classical steps of data preparation, feature extraction, feature selection and classifier training [47].

The approach's input is the raw data generated by sensor modules installed in each heavy machinery every 15 minutes (5 highest predominant frequencies in measured vibrations with their respective RMS amplitudes in 3-axis).

In other words, the input from each sensor is a multivariate time series with ten variables $\{T_1, T_2, \ldots, T_{10}\}$, where T_2 is a series of floating point numbers $\{x_1, x_2, \ldots, x_{10}\}$ made sequentially through time.

After receiving the input data, datasets are cleaned and prepared for statistical analysis, feature extraction and classifiers training. Several techniques were used along the data analysis process to get the best performance in speed and accuracy. The output of the approach predicts each monitored main component's health status and heavy machinery in real-time.

B. DATA PREPARATION

The raw sensor data are not directly ready to build up the classification models. In several cases, these data contain outliers and missing values that will influence the features' calculation accuracy.

Data cleansing is a mandatory phase that should precede all other machine learning phases. In data cleansing, two tasks were executed: (*i*) identification and handling of missing values, and (*ii*) identification and handling of outliers.

The outlier is a numeric value with an unusually high deviation from the mean or median value. Although there are numerous sophisticated algorithms for outlier detection, a simple statistical method is used in this work. This method is based on the interquartile range, which measures the variability of data [47]. After the identification, the outliers, as well as the missing values, were substituted by the mean value of the dataset in the neighbourhood. The data cleansing procedure is simple but very efficient. It can remove most erroneous values from the data.

C. DATA NORMALIZATION

Data were normalized to reduce unwanted variation between datasets as well as to allow data from different scales to be compared by converting them into a common unified range [48].

Since the vibration's excursion range of heavy machinery differs from one machine to another, for comparison all vibration data measurements were normalized by dividing by the maximum vibratory severity reported by vendors and operational data benchmark per type of heavy machine.

D. STATISTICAL FEATURE EXTRACTION

The second step in the proposed approach is feature extraction, which is transforming patterns in time series into features that are considered a compressed representation of datasets.

Feature extraction is related to dimensionality reduction of datasets. In machine learning, feature extraction starts from an initial set of measured data and builds derived values or features intended to be informative and non-redundant, assisting the subsequent learning and generalization steps for better process interpretations and predictions. The selected features are expected to contain the relevant information from the input data, so feature extraction involves reducing the number of resources required to describe large datasets. Another objective of feature extraction methods is to avoid overfitting the data to make data analysis and model prediction possible.

Heavy machinery time series have very high dimensionality. Therefore, mining such data is a challenge because a huge number of features can be extracted from the raw data [47]. A high-level representation is built to reduce the data dimensionality, where a set of significant features are calculated. These features provide an approximation of the original time series datasets.

For each time series variable $T_{\mathrm{i}} = \{x_1, x_2, \ldots, x_{10}\}$, $i = 1 \ldots 10$, a number of statistical features were calculated to measure different properties of each variable. In this paper, the statistical feature analysis considers: (*i*) measures of central tendency, (*ii*) measures of variability, (*iii*) measures of shape, (*iv*) measures of position, and (*v*) measures of impurity. In specific, the statistical analysis included the following evaluations: (*i*) arithmetic mean value, (*ii*) standard deviation, (*iii*) root-mean-square, (*iv*) coefficient of variation or relative standard deviation, and (*v*) standardized moment, to measure tendency and dispersion of frequency distribution of measured data; (*vi*) kurtosis, (*vii*) skewness, (*viii*) Fisher-Pearson asymmetry coefficient of skewness, to measure shape attributes and other information about curve behaviour; and, (*ix*) percentiles (p10, p25, p50, p80) and (*x*) entropy to measure position and impurity. In addition to the above-mentioned measures, basic statistical functions were calculated like min, max, sum, first, last and range.

The sensor modules designed are capable of acquiring the first 10 predominant frequency components of the vibratory signal generated by the machine's piece element under

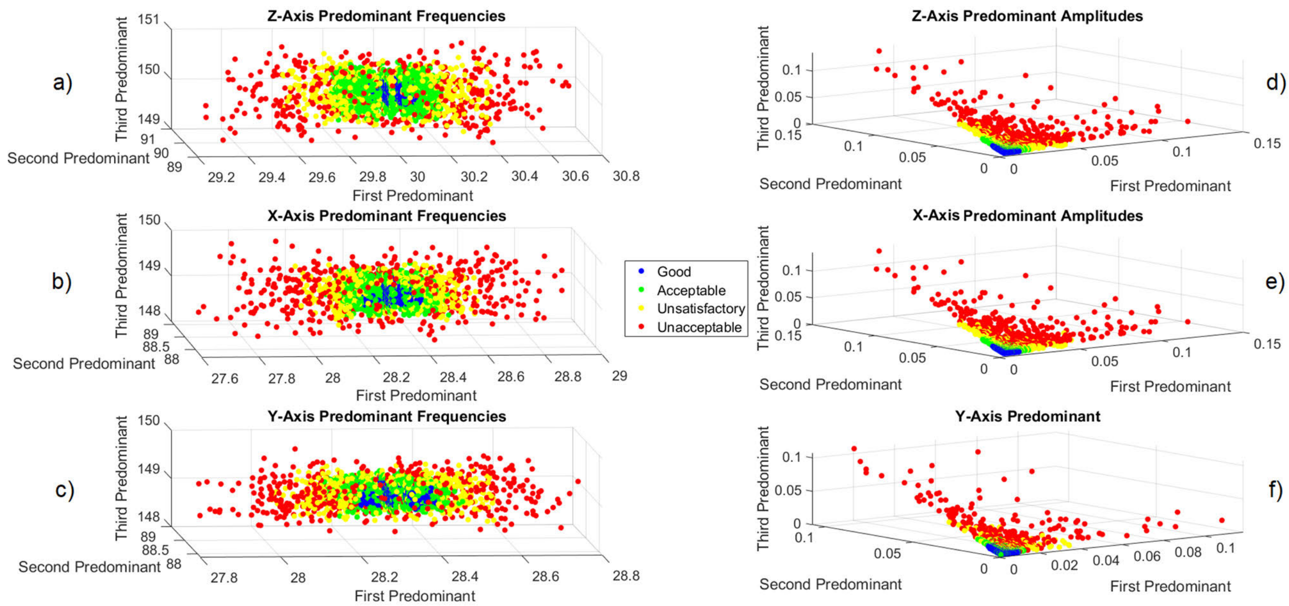

FIGURE 8. Scatter plots of predominant frequency components (amplitudes and frequencies) per axis. a) b) c) d) e) f).

evaluation (sampling frequency = 440 Hz, bandwidth up to 220 Hz). Operational data shown that the cumulative sum of magnitudes of the 5 predominant frequency components of the signal accounts for 85% of the signal's information, approximately [47], [48].

The statistical feature analysis of the actual frequency values for the 5 predominant components of vibratory signals per axis shown:

- The machine #1 presents a coefficient of variation of the average values of predominant frequencies of 0.2% versus 0.11% for machine #2 ($\sim$100% difference in dispersion of frequency values between a machine in critical failure condition and a machine in good operational conditions). The machine #1 (that reached the unacceptable health status) showed greater variability in the frequency data under the same operating conditions during the 60 trial days. This feature analysis of field data shown a direct relationship between dispersion of predominant frequencies in the vibratory signals and the machine health status. The higher the coefficient of variation and dispersion of predominant frequency values, the worse the machine's health status.
- The Fisher-Pearson asymmetry coefficients of skewness for the average values of predominant frequencies in x-, y- and z-axis for machine #1 (heavy machinery with information in the 4 vibratory severity states according to ISO 2372) are -0.0013, -0.0043 and 0.011, respectively. The dominant frequency in 3-axis presents a leptokurtic distribution, i.e. signals with a positive excess kurtosis, while the other predominant frequencies present a platykurtic distribution, i.e. signals with a negative excess kurtosis. This information was key for kernel estimation in the application of Naïve Bayes and SVM algorithms for real-time predictions based on frequency information. In this work, the field data collected from the two backhoes during the 60 days trial was used to estimate the initial kernel parameters for both machine learning techniques. Then, sliding window trend analysis was applied to the time series to update the kernel parameters in real-time using sampled and resampled time data series to get an accuracy above 95% in vibratory severity predictions and improved forecasting horizon based on models fitted to present and past observations.
- In Fig. 8.a-c shows the dispersion of the frequency values of the 3 predominant components of the signal measured with their respective vibration severity diagnostic labels according to the ISO 2372 standards. The predominant accelerations were concentrated around 30 Hz (29.3 to 30.4 Hz), 90 Hz (89.2 to 90.7 Hz) and 150 Hz (149 to 151 Hz) for the x-, y- and z-axis, respectively. The data on z-axis presents a higher dispersion than the measurements on the other axes. The z-axis corresponds to the direction with highest degree of freedom of mobile heavy machinery. Finally, it should be noted that the actual vibrations with predominant frequencies away apart from the average values reported (30 Hz, 90 Hz and 150 Hz) indicated the machinery condition evolution to critical health status in terms of vibratory severity.

The statistical feature analysis of the actual amplitude values for the 5 predominant components of vibratory signals per axis shown:

- The machine #1 presents a coefficient of variation of the average amplitudes of predominant frequency

TABLE 7. Classification results with different classfiers algorithms.

Machine Learning Algorithm	Accuracy [%]	Recall Good class [%]	Recall Acceptable class [%]	Recall Unsatisfactory class [%]	Recall Unacceptable class [%]
Naive Bayes[1]	92.66	97.52	85.82	77.74	93.76
Naive Bayes[1] with smoothing density estimate model.	96.02	99.97	89.09	86.61	98.12
Naive Bayes[1] with smoothing density estimate model and positive support	96.14	99.97	88.66	88.38	98.96
SVM[2] Linear Kernel	95.02	98.20	89.40	86.77	97.71
SVM[2] polynomial order 2 kernel	92.00	96.56	83.91	82.09	93.34
SVM[2] gaussian kernel with a scale of sqrt(datalength)*2	95.60	99.50	90.01	83.70	96.67
SVM[2] gaussian kernel with a scale of sqrt(datalength)/2	89.01	98.99	89.71	85.64	86.44
SVM[2] gaussian kernel with a scale of sqrt(datalength)*0.56	95.02	99.16	89.89	78.70	98.33
Linear Discriminant	84.22	87.88	77.88	73.87	87.94
Quadratic Discriminant	93.25	97.59	87.18	79.03	95.21

[1] The best results with Naïve Bayes was always the technique with gaussian kernel.
[2] SVM was tested only with unbonded support because it requires normalized data as input.

components of 127% versus 53.7% for machine #2 (a machine in critical failure condition reached 2.4 times higher dispersion of average amplitudes of predominant frequency components than a machine in good operational conditions). The machine #1 (that reached the unacceptable health status) showed greater variability in vibration frequency components' amplitudes under the same operating conditions during the 60 trial days. This feature analysis of field data shown a direct relationship between amplitude dispersion of predominant frequency components in the vibratory signals and the machine health status. The higher the coefficient of variation and dispersion of amplitude values, the worst the machine health status.

- The Fisher-Pearson asymmetry coefficients of skewness for the average amplitudes of predominant frequency components in x-, y- and z-axis for machine #1 (heavy machinery with information in the 4 vibratory severity states according to ISO 2372) are 6.4, 7.5 and 0.011, respectively. The predominant frequencies in 3-axis presents a leptokurtic distribution, i.e. signals with a positive excess kurtosis. This information was also key for kernel estimation in applying Naïve Bayes and SVM algorithms for real-time predictions based on amplitude information.
- In Fig. 8.d-f shows the dispersion of the average amplitude values of the 3 predominant components of the signal measured with their respective vibration severity diagnostic labels according to the ISO 2372 standards. The predominant accelerations' amplitude was concentrated below 0.01, 0.01 and 0.02 for Good and Acceptable states in the x-, y- and z-axis, respectively. Any excursion beyond those limits indicated the evolution of machine condition to Unsatisfactory and Unacceptable states in terms of vibratory severity.

E. FEATURE SELECTION

High dimensional datasets, which has hundreds of possible features, can contain a high degree of irrelevant and redundant information which might greatly reduce the performance of machine learning algorithms [49]. Therefore, feature selection becomes necessary.

Designers must choose a subset of relevant features with high predictive value for creating robust machine learning models [45], [46]. In this work, feature selection was implemented at hardware/firmware level to improve selected machine learning models' performance while reducing the data traffic between sensor network devices.

All the machine learning techniques shown in Fig.3 can be trained using time series in raw data format and/or processed values like the RMS value of vibratory signal that represents the mechanical energy release in our study. The decision of what type of data should be used to train the data-driven models in each case is given by technique requirements such as precision, computational capacity and amount of data available, among others.

Data clustering on the measured time series was used in this work for feature selection using key indicators for diagnosis of machinery health status like RMS value of vibratory signals and the FFT frequency analysis of the signals over

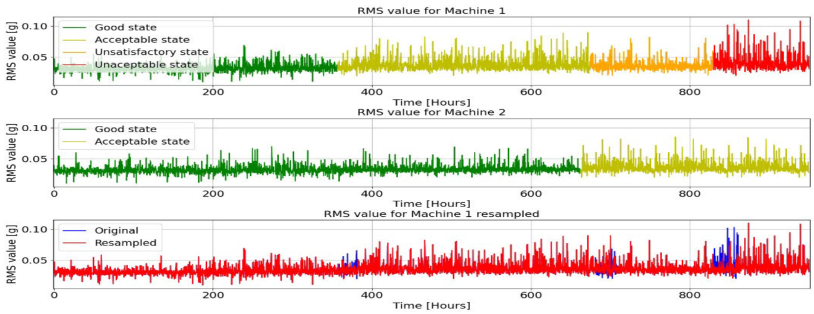

FIGURE 9. Original time series of RMS vibration data for machines #1 and #2.

sliding windows. The amplitude spectrum analysis was used to determine the 5 predominant frequency components, the dispersion of the actual frequency and amplitude values of the predominant frequencies, and the statistical frequency distribution of the vibratory signal in each time window. These operating variables are considered relevant for the diagnosis health state of heavy machinery [49].

F. CLASSIFICATION RESULTS

With all the data from the measurements, the machine learning algorithms described in section V-F was trained and evaluated. These algorithms were trained only with the data from a BAS module located in chassis close to the diesel engine.

The classification results attained with the machine learning algorithms selected are shown in Table 7. The best results in both, prediction accuracy and true positive rate or recall for all states of vibratory severity according to ISO 2372 as a whole, were achieved with the Naïve Bayes algorithm with a smoothing density estimate model and positive support. This algorithm achieves a prediction accuracy exceeding 96% with an average true positive rate or recall for health state prediction of 94%. Moreover, the same machine learning algorithm allows to predict a critical failure of the machinery with a forecasting horizon up to 170 hours, being the best-in-class among techniques used in this work. Thus, data-driven models help improving early fault detection up to 50 hours before manual diagnosis. This indicates that proposed design satisfies the condition monitoring requirements and will help to generate predictive maintenance plans. For more details see Tables 7 and 10.

Fig. 9 shows the RMS values of the vibrations of machine #1 and machine #2. A condition monitoring expert labelled these signals based on their vibratory severity for health state classification according to ISO 2372. The vibrations collected during the 60 days trial in both backhoes are denominated Original signals in this work. Note that machine #1 exhibited the four possible health state conditions (Good, Acceptable, Unsatisfactory and Unacceptable) while machine #2 only reached the first 2 states. Table 8 shows each health state's specific start time along measured time series dataset for machine #1. This table shows the health status predictions provided by machine learning models versus manual

TABLE 8. Health state transition time for machine #1 using original signal.

	Start time [h]			
	Good state	Acceptable state	Unsatisfactory state	Unacceptable state
Expert Label	0	357	675	830
Naïve. Bayes1 with smoothing density estimate model and positive support.	0	358	676	831
Naive Bayes.	0	360	678	832
SVM^2 Linear. Kernel	0	359	677	832
Linear Discriminant.	0	365	688	840

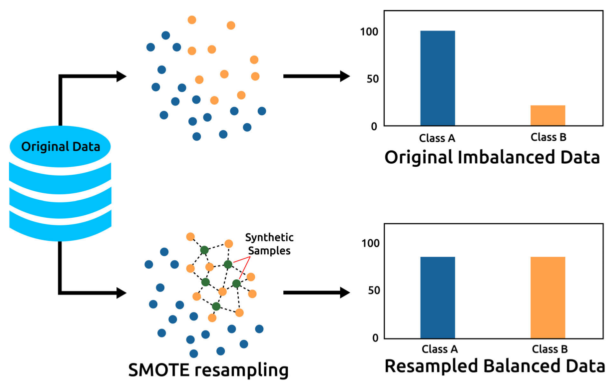

FIGURE 10. Schematic diagram of synthetic minority random oversampling technique (SMOTE resampling) used in this work.

diagnosis results. It should be pointed out that none of the proposed machine learning models were able to provide better forecasting horizon than manual diagnosis based on original signals datasets. This result is mainly due to the fact that changes in machine health state occur very quick and abruptly. Therefore, the system cannot capture enough information on the transitions between health states to improve forecasting horizon while ensuring good prediction accuracy. Most common machine learning algorithms usually work

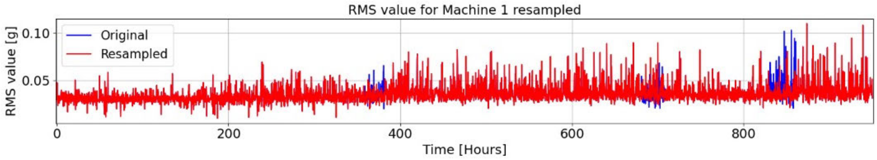

FIGURE 11. Resampled RMS vibration data for machine #1 using SMOTE resampling technique.

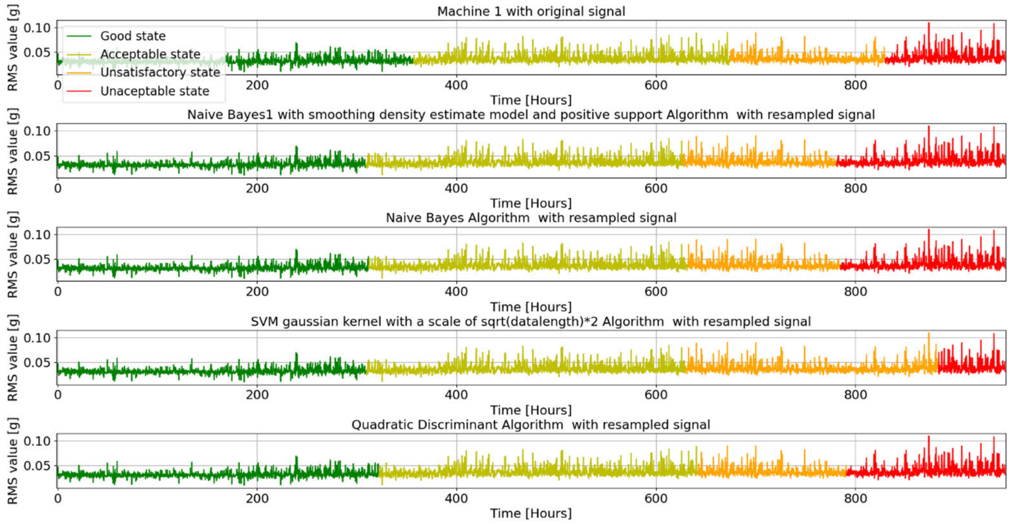

FIGURE 12. Machine learning health state classification results based on RMS vibration data for machine #1 according to ISO 2372.

well on balanced training sets, that is, datasets in which all classes are approximately represented equally [50]. Because these algorithms treat all misclassifications equally, they bias classes with many instances, resulting in false accuracy estimates. Therefore, the accuracy estimates may be unreliable and classes with only a few values are often misclassified or neglected [50]. This issue is known as a class imbalance problem in machine learning. Datasets that do not meet this criterion are referred to as imbalanced data [50]. Most condition monitoring datasets are imbalanced data in some extent. Time series of vibration excursions collected from heavy machinery in non-stationary operations under typical operating conditions in mine sites usually are not equal in data density for each class. It can additionally be influenced by the sampling strategy [50].

Several approaches have been developed in the machine learning community to handle imbalanced data. One is the design of new models that can directly handle imbalanced datasuch aspplying cost functions that penalize wrong classification [51]. Another approach is to apply different evaluation metrics instead of the overall accuracy, such precision and recall [52], [53]. A third approach is to resample the data [54].

Several resampling approaches have been proposed which can be separated in two groups: (a) data-driven and (b) algorithm driven methods [52], [53] [54]. Most researchers have employed data-driven methods [54] which use resampling techniques to adjust the ratio between the classes in the training set [51]. In their simplest forms, random oversampling (ROS) increase the minority class data by random replication of their occurrence, and random undersampling (RUS) decreases the number of majority class data by randomly removing data from the original dataset. This consequently allows machine learning algorithms to be trained from the balanced data without bias [52] [53].

With the aim to improve forecasting horizon in health state predictions, synthetic minority random oversampling technique (SMOTE resampling) was used in this work over the RMS vibration measured signals focused in state transition data. (see Fig. 10). Random oversampling is the simplest oversampling technique to balance the imbalanced nature of the dataset. It balances the data by replicating the minority class samples. This does not cause any information loss, but the dataset is prone to overfitting as the same information is replicated. To avoid this issue, SMOTE is preferred. It creates new synthetic samples to balance the dataset. SMOTE works by utilizing a k-nearest neighbor algorithm to create synthetic data. Steps samples are: (*i*) identify the feature vector and its nearest neighbor, (*ii*) compute the distance between the two sample points, (*iii*) multiply the distance with a random number between 0 and 1, (*iv*) identify a new point on the line segment at the computed distance, and (*v*) repeat the process for identified feature vectors taking into account average value, standard deviation, precision, recall and accuracy.

TABLE 9. Health state transition time for machine #1 using original signal.

	Start time [h]			
	Good state	Acceptable state	Unsatisfactory state	Unacceptable state
Expert Label	0	357	675	830
Naive Bayes1 with smoothing density estimate model and positive support	0	309	627	781
Naive Bayes	0	312	632	785
SVM^2 Linear Kernel	0	310	630	783
Linear Discriminant	0	322	645	791

TABLE 10. Summary of forecasting horizon for critical failure (unacceptable health state) for machine #1 reached for each machine learning method using resampled data.

	Original	Resampling
Expert label	120	
Naive Bayes1 with smoothing density estimate model and positive support	119	169
Naive Bayes	118	165
SVM^2 Linear Kernel	119	167
Linear Discriminant	119	159

Fig.11 shows resampled RMS value of vibration measurements for machine #1 considering up to 100 extra samples in health state transitions to get less abrupt changes and enough synthetic information to improve forecasting horizon.

Table 9 and Fig.12 show each health state's start times for machine #1 predicted by machine learning methods with resampled data versus manual diagnosis results.

Finally,

Table 10 shows a summary of the forecasting horizon results for the different machine learning algorithms, using the original and resampled data for both cases.

VII. CONCLUSION

A rugged, easy-to-deploy and smart vibration monitoring system based on IoT devices and data-driven models to assess heavy machinery's real-time condition in non-stationary operations for mining industry have been proposed. Proposed solutions for hardware, firmware and machine learning algorithms for software were cost-effective for implementing a wireless vibration sensor network for condition monitoring using Bluetooth and 3G/4G/LTE data transmission. Designed sensor modules can be attached to the chassis of heavy machinery in different locations using magnetic coupling to measure 3-axis accelerations in real-time. Once deployed on-board, the sensor modules interconnected with a Hub module forming a star network topology using Bluetooth links and externally through mobile cellular networks to a remote monitoring server on-site avoiding interference with existing wireless networks. Machinery vibration measured data was sent every 15 minutes to a remote server for health state evaluation using machine learning according to ISO 2372 vibratory severity standards. The design of sensor modules meets different technical and functional requirements to ensure embedded and distributed capabilities: (i) network management, (ii) data routing and traffic control, (iii) power supply unit, (iv) power management unit, (v) data processing unit, and (vi) sensing unit. The proposed wireless sensor module are low-energy consumption, it has IP67 protection to surpass mining harsh environmental conditions; it is a light and small size to ensure compatibility with typical O&M conditions, and it includes wireless battery recharging to avoid cable connections for powering and communications.

The best classification results in both, prediction accuracy and true positive rate for all states of vibratory severity according to ISO 2372 as a whole, were achieved in this work with the Naïve Bayes algorithm with a smoothing density estimate model and positive support. This algorithm achieves a prediction accuracy exceeding 96% with an average recall for health state prediction of 94%. Moreover, this machine learning algorithm allows to predict a critical failure of the machinery with a forecasting horizon up to 170 hours, being the best-in-class among techniques used in this work. Thus, data-driven models help improving early fault detection up to 50 hours before manual diagnosis. This indicates that proposed design satisfies the condition monitoring requirements and will help to generate predictive maintenance plans. With the aim to improve forecasting horizon in health state predictions, synthetic minority random oversampling technique (SMOTE resampling) was used in this work over the RMS vibration measured signals focused in state transition data. Random oversampling is the simplest oversampling technique to balance the imbalanced nature of the dataset. It balances the data by replicating the minority class samples. This does not cause any information loss, but the dataset is prone to overfitting as the same information is replicated. To avoid this issue, SMOTE was preferred in this work to balance the dataset utilizing a k-nearest neighbour algorithm to create synthetic data.

VIII. FUTURE SCOPE

The future scope has three significant challenges:

1- Monitor a highernumber and more diverse group of backhoes to find the minimum number of sensing points in order to predict the different critical components' failures.

2- This paper presents a family of machine learning data-driven algorithms that are trained under a robust set of assumptions. However, there is a wide and specialized literature of alternative algorithms and alternative settings with potential applications in this problem. Furthermore, the question of what algorithm and what predictor variables to use, to achieve the best machine operation performance, could have a state-dependent answer; based on current real-time data, data frequency, and computational power. Indeed, in the context of non-stationary operations, these are interesting topics and challenging questions that leave room for future research.

REFERENCES

[1] E. Jakobsson, *Data-Driven Condition Monitoring in Mining Vehicles*. Linköping, Sweden: Linköping Univ., Department of Electrical Engineering, 2019.
[2] M. S. Value, ''New tech new deal, technology impacs review,'' Winnipeg, MB, Canada, Tech. Rep., 2019.
[3] B. S. Dhillon and O. C. Anude, ''Mining equipment reliability: A review,'' *Microelectron. Rel.*, vol. 32, no. 8, pp. 1137–1156, Aug. 1992.
[4] R. Ahmad and S. Kamaruddin, ''An overview of time-based and condition-based maintenance in industrial application,'' *Comput. Ind. Eng.*, vol. 63, no. 1, pp. 135–149, Aug. 2012.
[5] S. Chakraborty, *Handbook of Offshore Engineering*. Amsterdam, The Netherlands: Elsevier, 2005.
[6] C. Okoh, R. Roy, J. Mehnen, and L. Redding, ''Overview of remaining useful life prediction techniques in through-life engineering services,'' in *Proc. 6th CIRP Conf. Ind. Product Services Syst. Value Creation*, 2014, pp. 158–163.
[7] O. A. Makinde, ''Establishment of the best maintenance practices for optimal reconfigurable vibrating screen management using decision techniques,'' *Int. J. Qual. Rel. Manage.*, Sep. 2016.
[8] Y. Peng, M. Dong, and M. J. Zuo, ''Current status of machine prognostics in condition-based maintenance: A review,'' *Int. J. Adv. Manuf. Technol.*, vol. 50, nos. 1–4, pp. 297–313, Sep. 2010.
[9] F. W. Boelter, J. W. Spencer, and C. E. Simmons, ''Heavy equipment maintenance exposure assessment: Using a time-activity model to estimate surrogate values for replacement of missing data,'' *J. Occupational Environ. Hygiene*, vol. 4, no. 7, pp. 525–537, May 2007.
[10] K. Komonen, P. Oedewald, M. Peltonen, T. Reiman, and V. Rouhiainen, ''Technical and economic assessment, developed and optimising techniques of industrial maintenance,'' Saf. Rel. Technol. Rev., Tech. Rep. BTUO43-031209.
[11] A. K. S. Jardine, D. Lin, and D. Banjevic, ''A review on machinery diagnostics and prognostics implementing condition-based maintenance,'' *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, Oct. 2006.
[12] N. Sakib and T. Wuest, ''Challenges and opportunities of condition-based predictive maintenance: A review,'' in *Proc. 6th CIRP Global Web Conf. Envisaging Future Manuf., Design, Technol., Syst. Innov. Era*, Morgantown, WV, USA, 2018, pp. 267–272.
[13] R. S. Beebe, *Predictive Maintenance of Pumps Using Condition Monitoring*. Melbourne, VIC, Australia: Elsevier, 2001.
[14] X. Liu, D. He, G. Lodewijks, Y. Pang, and J. Mei, ''Integrated decision making for predictive maintenance of belt conveyor systems,'' *Rel. Eng. Syst. Saf.*, vol. 188, pp. 347–351, Aug. 2019.
[15] *ISO/TC 108/SC 5 Condition Monitoring and Diagnostics of Machine Systems*, Standard ISO 13381-1:2015, 2015.
[16] D. An, N. H. Kim, and J. Choi, ''A review of data-driven and physics-based prognostics,'' in *Proc. Annu. Conf. Prognostics Health Manage. Soc.*, 2013.
[17] M. Chiachío, J. Chiachío, G. Rus, and J. L. Beck, ''Predicting fatigue damage in composites: A Bayesian framework,'' *Struct. Saf.*, vol. 51, pp. 57–68, Nov. 2014.
[18] F. J. Montns, F. Chinestab, R. Gómez-Bombarelli, and J. N. Kutzd, ''Data-driven modeling and learning in science and engineering,'' *Comp. Rendus Mécanique*, vol. 347, no. 11, pp. 845–855, 2019.
[19] P. Baraldi, F. Mangili, and E. Zio, ''Investigation of uncertainty treatment capability of model-based and data-driven prognostic methods using simulated data,'' *Rel. Eng. Syst. Saf.*, vol. 112, pp. 94–108, Apr. 2013.
[20] B. Samanta, B. Sarkar, and S. K. Mukherjee, ''Reliability centred maintenance (RCM) for heavy earth-moving machinery in an open cast coal mine,'' *CIM Bull.*, vol. 94, no. 1056, pp. 104–108, 2001.
[21] M. Haddar, W. Bartelmus, F. Chaari, and R. Zimroz, *Advances in Condition Monitoring of Machinery in Non-Stationary Operations*. Lyon, France: Springer, 2017.
[22] L. D. Hasfjord, ''Maintenance alert system for heavy-duty trucks,'' U.S. Patent 6 172 602 B1, Sep. 6, 2001.
[23] C. H. Oppenheimer and K. A. Loparo, ''Physically based diagnosis and prognosis of cracked rotor shafts,'' in *Proc. 2nd Compon. Syst. Diag., Prognostics, Health Manage.*, Orlando, FL, USA, vol. 4733, Jul. 2002, pp. 122–132.
[24] J. Yan, M. Koç, and J. Lee, ''A prognostic algorithm for machine performance assessment and its application,'' *Prod. Planning Control*, vol. 15, no. 8, pp. 796–801, Dec. 2004.
[25] N. Z. Gebraeel and M. A. Lawley, ''A neural network degradation model for computing and updating residual life distributions,'' *IEEE Trans. Autom. Sci. Eng.*, vol. 5, no. 1, pp. 154–163, Jan. 2008.
[26] Z. Iqbal, K. Kim, and H.-N. Lee, ''A cooperative wireless sensor network for indoor industrial monitoring,'' *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 482–491, Apr. 2017.
[27] J. Brodny and M. Tutak, ''Analysing the utilisation effectiveness of mining machines using independent data acquisition systems: A case study,'' *Energies*, vol. 12, no. 13, p. 2505, Jun. 2019.
[28] J. Barabady and U. Kumar, ''Reliability analysis of mining equipment: A case study of a crushing plant at Jajarm Bauxite Mine in Iran,'' *Rel. Eng. Syst. Saf.*, vol. 93, no. 4, pp. 647–653, Apr. 2008.
[29] J. Brodny, M. Tutak, and M. Michalak, ''The use of the TGŚP module as a database to identify breaks in the work of mining machinery,'' in *Proc. Int. Conf., Beyond Databases, Archit. Struct.*, 2017.
[30] J. Brodny, M. Tutak, and M. Michalak, ''A data warehouse as an indispensable tool to determine the effectiveness of the use of the longwall shearer,'' in *Proc. Int. Conf., Beyond Databases, Archit. Struct.*, 2017.
[31] H. Yıldırım, A. N. Özsezen, and A. Çınar, ''Vibration and noise depending upon engine speed in a diesel engine fueled with biodiesel,'' in *Proc. 6th Eur. Conf. Ren. Energy Syst.*, Istanbul, Turkey, 2018, pp. 25–27.
[32] M. S. Kan, A. C. C. Tan, and J. Mathew, ''A review on prognostic techniques for non-stationary and non-linear rotating systems,'' *Mech. Syst. Signal Process.*, vols. 62–63, pp. 1–20, Oct. 2015.
[33] B. Sun, S. Zeng, R. Kang, and M. G. Pecht, ''Benefits and challenges of system prognostics,'' *IEEE Trans. Rel.*, vol. 61, no. 2, pp. 323–335, Jun. 2012.
[34] J. Veldman, W. Klingenberg, and H. Wortmann, ''Managing condition-based maintenance technology: A multiple case study in the process industry,'' *J. Qual. Maintenance Eng.*, vol. 17, no. 1, pp. 40–62, Mar. 2011.
[35] A. Tripathy, A. Agrawal, and S. K. Rath, ''Classification of sentimental reviews using machine learning techniques,'' *Procedia Computer Science*, vol. 57, pp. 821–829, Jan. 2015.
[36] K. Sabanci, M. Ünlerşen, and K. Polat, ''Classification of different forest types with machine learning algorithms,'' Res. Rural Develop., Bolu, Turkey, Tech. Rep., 2016.
[37] C. J. C. Burges, ''A tutorial on support vector machines for pattern recognition,'' *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
[38] C. Campbell and Y. Ying, *Learning With Support Vector Machines*, vol. 25, no. 1. San Rafael, CA, USA: Morgan & Claypool, 2011, p. 95.
[39] X. Tan, W. Bi, X. Hou, and W. Wang, ''Reliability analysis using radial basis function networks and support vector machines,'' *Comput. Geotech*, vol. 38, no. 2, pp. 78–186, 2011.
[40] C. E. Rasmussen, ''Gaussian processes in machine learning,'' *Adv. Lect. Mach. Learn*, vol. 14, no. 2, pp. 63–71, 2004.
[41] S. Hong and Z. Zhou, ''Application of Gaussian process regression for bearing degradation assessment,'' in *Proc. Inf. Sci. Service Sci. Data Mining*, 2012, pp. 644–648.
[42] S. Saha, B. Saha, A. Saxena, and K. Goebel, ''Distributed prognostic health management with Gaussian process regression,'' in *Proc. IEEE Aerosp. Conf.*, Mar. 2010, pp. 1–8.
[43] H. Abdi, ''Discriminant correspondence,'' in *Encyclopedia of Measurement and Statistics*. Richardson, TX, USA: The Univ. of Texas at Dallas, 2007.
[44] J. Örn, ''Vibration guideline for large diesel engines,'' Vaasa, Finland, Tech. Rep., 2014.
[45] J. C. Haartsen, ''The Bluetooth radio system,'' *IEEE Pers. Commun.*, vol. 7, no. 1, pp. 28–36, Feb. 2000.
[46] J. Nieminen, C. Gomez, M. Isomaki, T. Savolainen, B. Patil, Z. Shelby, M. Xi, and J. Oller, ''Networking solutions for connecting Bluetooth low energy enabled machines to the Internet of Things,'' *IEEE Netw.*, vol. 28, no. 6, pp. 83–90, Nov. 2014.

[47] B. Esmael, A. Arnaout, R. Fruhwirth, and G. Thonhauser, "A statistical feature-based approach for operations recognition in drilling time series," in *Proc. ESMAEL ASF*, 2012, pp. 454–461.
[48] A. Choudhary, D. Goyal, S. L. Shimi, and A. Akula, "Condition monitoring and fault diagnosis of induction motors: A review," *Arch. Comput. Methods Eng.*, vol. 26, no. 4, pp. 1221–1238, Sep. 2019.
[49] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in *Proc. Sci. Inf. Conf.*, London, U.K., 2014, pp. 372–378.
[50] R. Taghizadeh-Mehrjardi, K. Schmidt, K. Eftekhari, T. Behrens, M. Jamshidi, N. Davatgar, N. Toomanian, and T. Scholten, "Synthetic resampling strategies and machine learning for digital soil mapping in Iran," *Eur. J. Soil Sci.*, vol. 71, no. 3, pp. 352–368, May 2020.
[51] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, no. 16, pp. 321–357, Jun. 2002.
[52] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Hong Kong, Jun. 2008, pp. 1322–1328.
[53] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.
[54] S. Piri, D. Delen, and T. Liu, "A synthetic informative minority oversampling (SIMO) algorithm leveraging support vector machine to enhance learning from imbalanced datasets," *Decis. Support Syst.*, vol. 106, pp. 15–29, Feb. 2018.
[55] L. F. Johnson, "Electronic control system for mobile heavy equipment machinery," U.S. Patent 9 116 787 B1, Aug. 25, 2015.
[56] H. Nugra, A. Abad, W. Fuentes, F. Galárraga, H. Aules, C. Villacís, and T. Theofilos, "A low-cost IoT application for the urban traffic of vehicles, based on wireless sensors using GSM technology," in *Proc. IEEE/ACM 20th Int. Symp. Distrib. Simulation Real Time Appl.*, 2016, pp. 161–169.
[57] M. Pagani and C. H. Fine, "Value network dynamics in 3G–4G wireless communications: A systems thinking approach to strategic value assessment," *J. Bus. Res.*, vol. 61, no. 11, pp. 1102–1112, Nov. 2008.
[58] A. Laukka, J. Saari, J. Ruuska, E. Juuso, and S. Lahdelma, "Condition-based monitoring for underground mobile machines," *Int. J. Ind. Syst. Eng.*, vol. 23, no. 1, pp. 74–89, 2016.
[59] K. L. Tsui, N. Chen, Q. Zhou, Y. Ha, and W. Wang, "Prognostics and health management: A review on data driven approaches," *Math. Problems Eng.*, vol. 2015, May 2015, Art. no. 793161.
[60] A. Jablonski and T. Barszcz, "Validation of vibration measurements for heavy duty machinery diagnostics," *Mech. Syst. Signal Process.*, vol. 38, pp. 248–263, Jul. 2013.
[61] P. Lewis, H. Fitriani, and I. Arocho, "Engine variable impact analysis of fuel use and emissions for heavy-duty diesel maintenance equipment," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2482, no. 1, pp. 8–15, Jan. 2015.
[62] I. Amihai, R. Gitzel, A. M. Kotriwala, D. Pareschi, S. Subbiah, and G. Sosale, "An industrial case study using vibration data and machine learning to predict asset health," in *Proc. IEEE 20th Conf. Bus. Informat. (CBI)*, Vienna, Austria, Jul. 2018, pp. 178–185.
[63] T. Schneider, N. Helwig, and A. Schütze, "Industrial condition monitoring with smart sensors using automated feature extraction and selection," *Meas. Sci. Technol.*, vol. 29, no. 9, Sep. 2018, Art. no. 094002.
[64] W. Caesarendra and T. Tjahjowidodo, "A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing," *Machines*, vol. 5, no. 4, p. 21, Sep. 2017.

PABLO AQUEVEQUE (Member, IEEE) was born in Santiago, Chile, in 1976. He received the B.S. degree in electronics engineering and the Ph.D. degree from the University of Concepción, Concepción, Chile, in 2000 and 2008, respectively. He is currently a Professor with the Department of Electrical Engineering, University of Concepción. He is also the Director of the Center for Industry 4.0 (C4I), Faculty of Engineering, Universidad de Concepción. His research interests include bioelectronics, embedded electronic, electrochemical process, power electronics, and wireless power transfer.

LUCIANO RADRIGAN (Student Member, IEEE) was born in Chillan, Biobio, Chile, in 1991. He received the B.S. degree in electronic engineering and the M.S. degree in electrical engineering from the University of Concepción, in 2019 and 2020, respectively, where he is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, Faculty of Engineering.

He currently works at the Center for Industry 4.0, University of Concepción, where the research is carried out for industrial processes. His research interests include sensors for condition monitoring, the IoT, industrial sensorization, machine learning, energy-harvesting systems, embedded systems, and wireless energy transfer.

FRANCISCO PASTENE (Member, IEEE) was born in Concepción, Biobio, Chile in 1995. He received the B.S. degree in biomedical engineering and the M.S. degree in electrical engineering from the Universidad de Concepción, in 2018 and 2020, respectively, where he is currently pursuing the Ph.D. degree with the Electrical Engineering Department, Faculty of Engineering.

He is the author of three conference papers and two articles. His research interests include a sensor for physiological signal acquisition, signal processing, medical devices, the IoT, and wireless power transfer.

Mr. Pastene received the Faculty of Engineering Prize and the Universidad de Concepción Prize from the Universidad de Concepción, in 2018 and 2019, respectively. He was a recipient of the CONICYT Fellowship for pursuing master's and Ph.D. degrees.

ANIBAL S. MORALES (Member, IEEE) was born in Concepción, Chile, in 1983. He received the B.S. degree, the Electronics Engineering degree (Hons.), and the Ph.D. degree in electrical engineering from the University of Concepción, Concepción, in 2004, 2007, and 2012, respectively. He is currently an Assistant Professor with the Department of Electrical Engineering, Universidad Católica de la Santísima Concepción, Chile. His current research interests include power converters, high-current rectifiers, multiphysics FEM modeling, energy efficiency, electrical safety, energy harvesting and technology for mining industry, copper electrowinning, and electrorefining.

ERNESTO GUERRA was born in Chile, in 1985. He received the B.S. degree in economics from the Universidad de Concepción and the M.S. and Ph.D. degrees in agricultural and resource economics from the University of California, Berkeley, in 2018. He is currently with the Economics Department, Universidad Católica de la Santísima Concepción, teaching econometrics and microeconomics. His research interests include econometrics estimations and numeric solutions of dynamic stochastic models.