Internet of Things 11 (2020) 100178

Contents lists available at ScienceDirect

Internet of Things

journal homepage: www.elsevier.com/locate/iot

Research article

Reliability-as-a-Service for bearing risk assessment investigated with advanced mathematical models

Jan-M. Brandt[a,*], Márton Benedek[b,c], Jeffrey S. Guerin[a], Jörg Fliege[d]

[a] 300 Wellington Street, Suite 104 (Bathurst Entrance), London, Ontario, N6B 2L5, Canada
[b] Dept. of Differential Equations, Inst. of Mathematics, Budapest University of Technology and Economics, Egry József u. 1., Budapest, 1111, Hungary
[c] Centre for Economic and Regional Studies, Tóth Kálmán u. 4., Budapest, 1097, Hungary
[d] Mathematical Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom

ARTICLE INFO

Article history:
Received 2 September 2019
Revised 27 January 2020
Accepted 10 February 2020
Available online 18 February 2020

Keywords:
Classification
Support vector machines
Industrial analytics
Reliability
Bearing tribology

ABSTRACT

As a key player in bearing service life, the lubricant chemistry has a profound effect on bearing reliability. To increase the reliability of bearings, an Industrial Analytics solution is proposed for proactive condition monitoring and this is delivered via a *Reliability-as-a-Service* application. The performance predictions of bearings rely on customized algorithms with the main focus on digitalizing lubricant chemistry; the principles behind these processes are outlined in this study. Subsequently, independent testing is performed to confirm the ability of the presented Industrial Analytics solution for such predictions. By deciphering the chemical compounds of lubricants and characteristics of the interface, the Industrial Analytics solution delivers a precise bearing reliability assessment *a priori* to predict service life of the operation. Bearing tests have shown that the classification system of this Industrial Analytics solution is able to predict 12 out of 13 bearing failures (92%). The described approach provides a proactive bearing risk classification that allows the operator to take immediate action in reducing the failure potential during smooth operation - preventing any potential damage from occurring. For this purpose, a mathematical model is introduced that derives a set of classification rules for oil lubricants, based on linear binary classifiers (support vector machines) that are applied to the chemical compound's mixture data.

1. Introduction

Digitalization of processes play a dominant role in nearly all aspects of life. This has led to the development of a new way on how humans and technology interact, most frequently referred to as the Internet-of-Things (IoT); in industrial applications it is often referred to as Industrial IoT (IIoT) and is part of the 4th Industrial Revolution (Industry 4.0) [1]. Thus,

List of abbrevations used in the paper: SVM, support vector machine; *IIoT*, Industrial IoT (Internet-of-Things); *RaaS*, Reliability-as-a-Service; *WEC*, white-etching crack; *XFA*, X-ray fluorescence analysis; *IR*, infrared techniques; *SRR*, slip-roll-ratio; k_{4LF}, dimensionless risk factor; *SIF*, surface-induced failure; *SCC*, stress corrosion cracking; *CCI*, chemical compound interaction; *QP*, quadratic program; *ANN*, artificial neural networks.

* Corresponding author.
E-mail address: jan.m.brandt@4linesfusion.com (J.-M. Brandt).

https://doi.org/10.1016/j.iot.2020.100178

there is an ever-increasing push to digitalize applications and knowledge with the goal to create avatars/digital twins that enhance global economic growth, productivity, and competitiveness for financial benefits. In essence, the future belongs to those who can combine existing knowledge with digital tools to strategically transform data into action.

At the core of IIoT applications, algorithms are created, utilized, and interconnected in order to increase system efficiency through predictions and further providing actions; these types of digital solutions are part of the rapidly growing market, referred to as Industrial Analytics; it is the process of collecting, analyzing, and using data generated from industrial operations with the simultaneous goal of increasing cost savings and enhancing reliability. This process is highly valuable to a wide range of asset-intensive applications in order to better understand their industrial operations and ensuring their economic viability. As of recent [2], it is expected that the global Industrial Analytics market will grow from USD 11.29 billion in 2017 to USD 25.51 billion by 2022, at a CAGR of 17.7%. Industrial Analytics can include software solutions, such as operational and risk monitoring and analytics for various industrial verticals. Thus, it comes as no surprise that predictive maintenance is an integral part of Industrial Analytics to reduce operational downtime, production costs, and increase system reliability [3]; in 2018, the predictive maintenance market was valued at USD 3.3 billion and is expected to reach USD 23.50 billion by 2024, at a CAGR of 39% between 2018 and 2024, whereas much of the growth is anticipated in IIoT applications [4].

Increasing bearing reliability is an essential part of enhancing the service life of moving machine elements. The performance of bearings is largely governed by the tribological behaviour within the application; tribology is the science of friction, lubrication, and wear, and has led to large cost savings since its introduction during the 1960's [5]. Consequently, the main parameters affecting the tribology of machine elements (i.e., the tribosystem) is defined by the interactions between the mating surfaces (e.g., material, roughness, etc.), the characteristics of the lubricating medium (e.g., type of oil and additives, viscosity, oxidation, etc.), the adjacent environment (e.g., temperature, contamination, etc.), and the applied mechanical factors (e.g., load and motion) [6].

The type of bearing used in an application plays a major role on the longevity of machine elements. In particular, roller bearings are widely used to allow the operation of machine elements being used in a wide range of areas such as transportation, wind and marine applications. During the design phase of machine elements, the service life of a roller bearing is estimated by assessing the most relevant standard (DIN ISO 281) that defines the methods of calculating the basic dynamic load rating of rolling bearings within a given size range. Further, DIN ISO 281 provides guidance to the design engineer on calculating the basic rating life - which is the life associated with a 90% reliability considering many of the previously mentioned system parameters [7]. Once the estimated bearing life is exhausted, the bearings usually fail due to material fatigue, referred to as rolling-contact fatigue. Of particular interest are the bearing failures that occur irregularly and suddenly which are then associated with massive operational and maintenance costs due to repairs not being able to be scheduled effectively.

The most prominent, irregular and sudden bearing failure is the failure due to white-etching cracks (WECs). This type of failure is associated with a vast financial burden for the operator as scheduled/calendar-based maintenance principles do not apply [8–13]. Thus, it is essential to define pathways that allow condition monitoring of the most detrimental parameters that lead to WECs. Extensive research on WECs has been conducted to date but there is still a clear need for proactive condition monitoring as part of predictive maintenance initiatives to detect early and critical changes in the operation [8]. With such early-warning systems, the operation and maintenance manager would be able take immediate action, before damage occurs, to return the operation back to the specification as per design (i.e., conduct immediate actions to return the operation back to bearing life as per DIN ISO 281).

In this study, a Software-as-a-Service application is presented to provide the operator with a classification alert system when the reliability of roller bearing application deviates from DIN ISO 281 life estimate. By combining knowledge in bearing tribology, lubricant chemistry and advanced data processing, various mathematical models have been successfully applied to different classification challenges, such as clustering methods, classifying with the usage of centrality measures, or support vector machines (SVMs). Measurements on performance and systematic behaviour of oil lubricants typically yield continuous data (i.e., data points in the Euclidean space $\mathbb{R}^n$) and therefore, we rely on SVMs, as opposed to clustering and data coming from a discrete space. The advantage of using SVMs is that they allow us to derive linear (binary) classification rules, that are easy to interpret, robust against outliers and they carry sufficient distinguishing power by maximizing the gap a classifier hyperplane yields towards two sets of data points to be separated (i.e., training sets). To the best of our knowledge, this is the first occasion that SVMs are being incorporated in the classification of oil lubricants. Subsequently, the principles of the proposed bearing risk assessment approach are explained, followed by an example on the analytical steps performed, the bearing tests conducted, as well as the application of SVMs to confirm the presented risk classification system.

2. Methods

2.1. Reliability-as-a-Service (RaaS)

In the age of IIoT, data is usually collected through sensors, stored in the cloud, and processed by algorithms that can be based on scientific knowledge and human expertise to make processes more reliable and thus, more efficient. In this study, this type of process is provided through proprietary algorithms of a Software-as-a-Service application tailored to bearing reliability (i.e. Reliability-as-a-Service (RaaS)). This RaaS application takes a holistic approach on tribology and permits a proactive approach in assessing a tribological application during smooth operation prior to any damage occurring. In practical terms, the operation and maintenance manager can take immediate action, before damage occurs, to return the operation back to the specifications as per design (i.e., return the operation back to bearing life as per DIN ISO 281). In the following, the high-level principles and processes used in the RaaS application (SeerWorks^{TM} Reliability, 4LinesFusion, Inc., London, ON, Canada) are outlined.

One of the greatest challenges in predicting tribological events is to have fast algorithms. Over the past decade, the approaches of across-scale modeling initiatives have proposed methods that allow faster computations. Coming from the molecular perspective with a size of 10^{-9} meters it takes great computational power to interpret effects on 10^{-9} to 10^{-3} meters (e.g., magnitudes of 10^{6} in length scale). These across-scale effects are critical to understand and to model as the interactions of relevant parameters determine the reliability of contacts in relative motion (dry and lubricated). Considerable progress in multi-scale modeling have been made in recent years [14]; in order to use these multi-scale models it is necessary to create interfacial descriptors and predictors to investigate systemic effects. Basically, predicators are obtained by the properties of a molecule (e.g., coming from the chemical structure given by the nature of bonding, energy and surface of the molecules). Exploring molecules by Quantitative Structure Property Relationship [15] and the molecular properties by the use of density functional theory has become widely used [16] and some of these pathways have been considered in the present study; here, the interactions of molecules with themselves and with surfaces is part of molecular dynamics and *ab initio* methods. The tribological behaviour of any lubricated contact is strongly governed by the chemical composition and mechanical characteristics of the lubricant at the interface. In the following, the principle procedure that allows embedding molecular properties of lubricants into classical thermodynamics is introduced as an integral part of the applied tribological model for the present study. It is considered that this model is particularly relevant to the type of bearing failure addressed in this present study as it is largely driven by the lubricant composition and its interactions with electrical and electromagnetic fields at the interface. As part of this present model, the Arrhenius equation has been considered to be a crucial component; it provides insight on how fast a molecular process moves from *A to B* in a given environment. This process is affected by the presence of temperature changes and is generally expressed as:

$$k = M * \frac{e^{-Ea}}{RT} \tag{1}$$

Ea is the activation energy (J/mol) for the process of shifting the state from *A to B*; R is the gas constant (8.314 J/K*mol) and T the absolute temperature. The factor M depends on how often the molecular specie collides and at which chemical orientation it occurs. The collisions are assumed to be purely elastic; however, molecules, especially organic compounds, are not rigid in these elastic interactions due to their inherent structure and their properties. Structures of molecules are expressed by their topology (i.e., spatial arrangement of molecular bonding) and their movement is mainly expressed by thermal stress, pressure or shearing. The relaxation time of such a system is an important factor on how the collisions and stresses can be uptaken and released by the molecules. Thus, an expansion of the understanding of the pre-exponential factor M and the activation energy Ea is presented by the use of descriptors derived through molecular modeling. Transforming the Arrhenius equation leads to:

$$\ln\left(\frac{k}{M}\right) = -\left(\frac{Ea}{RT}\right) \tag{2}$$

We now compare this relationship with the chemical potential (i.e., the enthalpy of a component in solution). The chemical potential (μ_i) of a constituent i and the chemical potential (μ_j) of a component j is given by:

$$\begin{aligned} \mu_i &= \mu_i(0) + RT\ln(a_i) \\ \mu_j &= \mu_j(0) + RT\ln(a_j) \end{aligned} \tag{3}$$

R is the gas constant, T the absolute temperature and $a_{i,j}$ the activity of the components i and j given as the activity coefficient times the molar concentration.

While in equilibrium, the factor D_{ij} is zero, expressing that all chemical potentials $\mu_i, \mu_j, \mu_k, \ldots$ are equal to:

$$D_{ij} = \mu_j - \mu_i = 0 \tag{4}$$

D_{ij} values > 0 reflect the imbalance (i.e., the inherent enthalpy not in equilibrium). The temperature independent factor D_{ij} is given by the imbalance of the system and may be compared to the relationship Ea/RT, meaning that:

$$\left(\frac{k}{M}\right) = -\left(\frac{a_j}{a_i}\right) \tag{5}$$

In general terms, it is then assumed that the processes from *A to B* are governed by the activity of the components *i* and *j*. The exponential factor A transforms the Arrhenius equation and this can be expressed as:

$$Ea = -RT\ln\left(\frac{a_i}{a_j}\right) \tag{6}$$

Within a molecular system, the activity is attributed to an attraction property, such as dipole moments or polarizability, by scaling it with the molecular mass to a molecular property (i.e., the specific dipole moment or the specific polarizability). The volume in the tribological contact is almost never constant; hence, the lubricant passing through a contact is frequently exposed to high pressure as matter of the reduced volume. As the contact area approaches zero, the activation energy Ea increases as the lubricant volume decreases:

$$Ea(j) = -T * \left(\frac{(cm_{ij} + \ln(\psi(j))}{V}\right)$$
$$\frac{Ea(j)}{T} = -\left(\frac{1}{V}\right) * (cm_{ij} + \ln(\psi(j))), \tag{7}$$

where cm_{ij} is related to a specific molecular property of constituent i or j, i.e. the specific polarizability, p; ($cm_{ij} = \ln(p_i/p_j)$. The reaction rate *A to B* increases if the spatial distribution of the j-th component decreases (e.g., the j-th component gets denser in the given volume). As lubricants possess the ability to solidify in the contact zone (i.e., create a higher spatial density to decrease the activation energy), the transformation from *A to B* is facilitated and reaction rates are suggested to increase. The ability of molecules to self-order, expressed as a degeneration capability, is an important factor for the lubricated contact. As the change of the volume in a contact is transient, the lubricant components require sufficient time for self-ordering; here, dipole moments act as recognition functions between molecules and polarizability as a recognition function for electrical charge.

The probability of molecular degeneration in the contact zone will increase if the accessibility of the function is high, by the fact that the density at the surface is high and the symmetry gain is also high (i.e., the symmetry increases in the degenerated state compared to the non-degenerated state). All factors are molecular properties and shall be determined by molecular modeling. The degeneration factor is high if the molecules possess the ability for long-range coupling. Under tribological conditions, the ability of a molecular system to create self-assemblies is considered an important property; these properties are closely related to the structure of lubricant components where the dipole moments and the polarizability are key players in molecular recognition.

Coming to the reduction of energy density, it is also obvious that these properties are closely related to the degeneration ability of such systems and may act as suitable descriptors for lubricants under technical conditions. This is related to the electron population in the highest occupied molecular state and the distance to the lowest unoccupied state. Once excited, the molecules completely relax over time; thus, the relaxation time is a measure of a molecular system to accommodate shear stress frequency; if the relaxation time is quick, the molecules are capable to uptake a higher shear frequency, whereas if the relaxation time is slower, the molecular systems will tend to accumulate more energy.

Reactivity in a tribosystem is accompanied by high rates of charge transfer (i.e., shearing molecules against each other and the creation/decay of electrical and electromagnetic fields during a tribological process. The response of a molecular system is dedicated to the dipole moment and the induced dipole moments (i.e., polarizability) as the main properties of a system. The driving force of a tribological reaction is explained by the gain of energy during the process (i.e., the difference of the free energy of the educts and the products). In a tribosystem, the lubricant involved may be written as a vector of all components and molecular digits may address the properties of the individual components of the lubricant. The individual properties may be given by the relaxation time, the ability to degenerate, the dipole moment as a molecular recognition digit, and the polarizability as a molecular reactivity digit that refers to the ability of the molecule to react via transient electrical fields within the contact. The dipole moment and polarizability are key players in molecular recognition and these parameters are closely related to the degeneration ability of such systems and may act as suitable descriptors for lubricants under tribological conditions.

Observing bearing failure through a classical view, the reliability is given by the load capacity of the component (e.g., the bearing) and the loading (i.e., the forces present in the tribosystem) [7]. In the classical view, lubricants act as a supporting element, that are mainly responsible for separating mating surfaces to reduce or eliminate friction and wear. Lubricant interactions and its degradation can have a profound effect on machine elements, altering the load capacity (e.g., corrosion, stress-corrosion cracking, sludge build-up, and a more recent phenomenon - WECs). Looking more specifically at WECs, it may not be the lubricant itself acting on changes in load capacity but its sub-components (i.e., base oil and additives), foreign elements (e.g., water, detergents, etc.), and also the interactions of all of these constituents that pass through the interfaces. Chemical structures and the complexity of their interactions render assessing WEC risk difficult for an end-user of a product. End-users rely on sending oil samples to a lab for an element and chemical analysis. Unfortunately, lab analytics do not go far enough in assessing the complexities of the chemical interactions. This leaves the end-user lost in assessing complex issues, like WECs, and thus, they have to manage with the uncertainty of failures.

The aformentioned principles and processes of SeerWorks^{TM} Reliability are applied to determine WEC risk classes. At the core, this RaaS application deciphers the main tribological parameters that define the bearing articulation [17]. More specifically, the physical, chemical and dynamic parameters of the interface are described in digital form; these parameters include lubricant constituents l_i ($l_{i-1}, \ldots, l_{i-n}$), the anisotropic, physical properties of the mating surfaces (s_i, s_j), their inhered dynamics (d_{ij}), the inherent structure of the materials surface (m_{si}, m_{sj}), the bulk properties ($b_i, \ldots, b_n$), under considerations of intrinsic or imposed electrical fields ($e_i, \ldots, e_n$). These parameters comprise the tribosystem and serve as the foundation of the algorithms. Their second derivative describes the magnitude of the curvature of the hyperplane in specific locations that indicate the likelihood of bearing failure. The associated unreliability of the tribological application is then

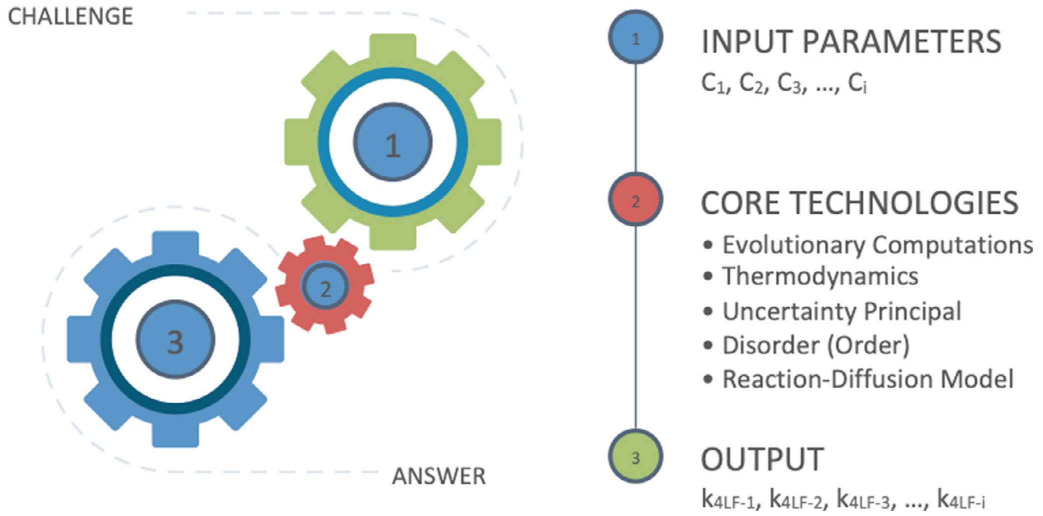

Fig. 1. Schematic showing the process flow to generate an application/situation-specific dimensionless risk coefficient, k_{4LF}, that is then subjected to a classification system.

directly linked to the extent of the discontinuity in the hyperplane. Dimensionless parameters describe, for every given incident in the hyperplane, the conditions that have shown to lead to bearing failure due to WEC. Lubricants are mixtures of functional components that are mainly based on organic chemistry, and thus are extremely difficult to describe in respect to their conjoint interactions in a tribosystem (e.g., surfaces, viscosity, temperature, and others). As an integral part of the SeerWorks™ Reliability, lubricants are defined as a component vector; within SeerWorks™ Reliability each lubricant component is modeled in respect to their transient activities, covering their reactivity in a defined tribosystem. These calculations provide valuable information of lubricant compounds, and the combinations thereof, that contribute to bearing failures due to WEC, among other bearing failures.

SeerWorks™ Reliability alleviates the uncertainty of failures due to the lubricant and gives the end-user an understandable risk assessment and concrete recommendations for reducing the risk of complex issues, like WEC. This is accomplished through an analysis of the media (e.g., lubricants, their sub-components, foreign elements, the system components and the entirety of the interactions between all these constituents). Starting with information about the current state of the lubricant using X-ray fluorescence analysis (XFA) and Infrared (IR) techniques, SeerWorks™ Reliability reduces the lubricants to their possible best-fit chemical structures by using several molecular modeling systems. Once the possible chemical structures are known, SeerWorks™ Reliability virtually simulates the structures property relationships combined with the known technical data. More specifically the simulations look at the molecules and their exposure to mechanical energy (e.g., Slip-Roll-Ratio, SRR), electrical energy and/or a combination of the two. Furthermore, SeerWorks™ Reliability also considers the mechanical system's activity of a given surface (from low roughness to high roughness) and phase deviations of the rotating parts (i.e., in-phase movement is represented with lower values and out-of-phase movement with higher values).

In principle, the known parameters that define the system serve as the input into SeerWorks™ Reliability and are subjected to algorithms that are based on various principles (Fig. 1). The output is the dimensionless risk factor, k_{4LF}; k_{4LF} describes attributes in relation to the input parameters/components in an applied system. With respect to WECs, it is known that both chemical and mechanical systems play a role in the reliability of a system, thus SeerWorks™ Reliability creates a dimensionless k_{4LF} risk coefficient that incorporates key parameters.

To generalize the risk assessment, SeerWorks™ Reliability takes the analysis and reduces it down to a simple set of generalized classes that can be used to more easily assess the risk of the system. In the present study, three (3) risk classes have been generalized:

- **Class I** indicates that the lubricant, its' components and mixtures create a low risk with respect to mechanical and electrical exposure. If a failure occurred, then it is unlikely that the failure was due to the lubricant as a prime cause.
- **Class II** indicates that the lubricant shows a pronounced risk due to the sensitivity of the components toward mechanical (SRR) and electrical impact. If the impact is high, then this would be considered to be critical. Class II could lead to further issues if the end-user does not reduce the risk levels (e.g., stress corrosion cracking (SCC), surface-induced failure (SIF) or white-etching cracks (WECs).
- **Class III** indicates that the lubricant, its' components, and/or foreign elements contribute to a high risk of a system failure. It is most likely that the risk being observed is of a WEC nature but the risk could also lead to other failure modes (e.g., SCC, SIF, or WECs).

In a real-life environment, sparingly available data can be gathered and analyzed with respect to the structures and their properties. The reliability assessment in this circumstance can only be as accurate as its inputs and data as provided. Even

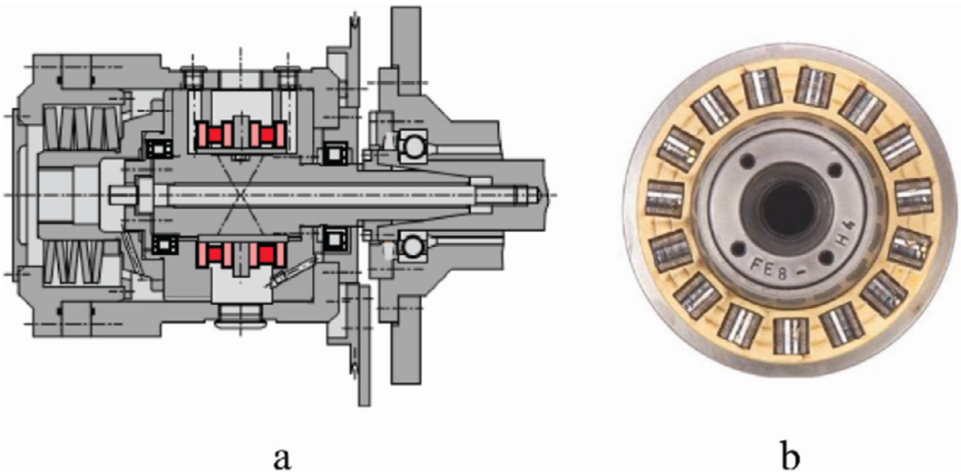

Fig. 2. (a) A general FE-8 test rig schematic and (b) cylindrical roller bearing (adopted from [19])

Table 1
Showing three possible variant structures of a determined lubricant consisting of base oil and additives ($C_1, \ldots, C_8$).

Chemical Compound	Variant Structure 1 (%)	Variant Structure 2 (%)	Variant Structure 3 (%)
$\mathbf{C_1}$	0.37	0.00	0.00
$\mathbf{C_2}$	0.00	0.67	0.67
$\mathbf{C_3}$	0.35	0.00	0.00
$\mathbf{C_4}$	0.00	0.64	0.64
$\mathbf{C_5}$	0.08	0.08	0.08
$\mathbf{C_6}$	0.28	0.00	0.00
$\mathbf{C_7}$	0.00	0.31	0.00
$\mathbf{C_8}$	0.00	0.00	0.44
Base Oil	98.92	98.29	98.16

with the most advanced systems the end-user is faced with a lack of precision due to the lack of information available for the assessment. SeerWorks^{TM} Reliability begins to reduce this uncertainty and for the first time the end-user has a real-time assessment tool.

2.2. Simplified Analytical Steps

A brief description of the analytical steps is provided in the following. It is assumed that the lubricant being used is unknown (Note: when information is unknown SeerWorks^{TM} Reliability performs an analysis using a best-fit approach which can reduce the precision of the risk analysis). As an example, the XFA breakdown has been supplied through an oil sample analysis:

- Calcium = 388 ppm
- Magnesium = 229 ppm
- Phosphorus = 3 ppm
- Sodium = 53 ppm
- Sulfur = 854 ppm
- Zinc = 365 ppm

SeerWorks^{TM} Reliability uses the XFA data to render it into three (3) possible lubricant variants in terms of chemical structures of each component (chemical structure components' names are obfuscated; see Fig. 1, input parameters: $C_1, C_2, \ldots, C_n$), likely one of which is near to the true composition lubricant (e.g., from the above XFA data we get Table 1). Additional precision can be added when the lubricant being used is provided. For the purpose of this study, SeerWorks^{TM} Reliability is not aware of the lubricant in use. SeerWorks^{TM} Reliability then performs analytics, through various statistical tools and proprietary algorithms, on each of the Chemical Compound Interactions (CCI) and the known surface material (Note: if the surface material is not known, then a non-stochiometric oxide, Fe-Fe_3O_4, is assumed as the default material). SeerWorks^{TM} Reliability then creates a k_{4LF} risk coefficient and a corresponding SRR (indication of where the threat is likely to occur) for each of the CCIs for phases 0.1 through 0.9, reflecting the phase distortion between the moving parts and activity 0.1 through 0.9. For example, using phase 0.1 and activity 0.1, the output for just the 1st variant from the above

Table 2
Showing various combinations of a variant structures and their corresponding k_{4LF} and SRR.

Variant Structure 1			
CCI		**k_{4LF}**	**SRR**
CC A	**CC B**		
C_1	C_5	0.96	3.09
C_1	C_6	.016	11.17
C_1	C_7	0.55	1.56
C_1	C_8	0.63	1.73
C_5	C_6	0.49	−1.35
C_5	C_7	0.31	0.73
C_5	C_8	0.84	−2.63
C_6	C_7	0.33	0.89
C_6	C_8	0.89	−2.84
C_7	C_8	1.26	5.19

Table 3
Showing Class categorization of each CCI for each of the three possible lubricant variants.

Variant Structure 1			Variant Structure 2			Variant Structure 3		
CCI		**Class**	**CCI**		**Class**	**CCI**		**Class**
CC A	**CC B**		**CC A**	**CC B**		**CC A**	**CC B**	
C_1	C_5	I	C_2	C_4	I	C_2	C_3	I
C_1	C_6	I	C_2	C_6	I	C_2	C_6	I
C_1	C_7	I	C_2	C_7	I	C_2	C_7	I
C_1	C_8	I	C_2	C_8	I	C_2	C_8	I
C_5	C_6	I	C_4	C_6	I	C_3	C_6	I
C_5	C_7	I	C_4	C_7	I	C_3	C_7	I
C_5	C_8	I	C_4	C_8	I	C_3	C_8	I
C_6	C_7	I	C_6	C_7	II	C_6	C_7	I
C_6	C_8	I	C_6	C_8	I	C_6	C_8	I
C_7	C_8	II	C_7	C_8	I	C_7	C_8	I

lubricant variants is as in Table 2. The information supplied in Table 2 can be useful to the end-user but we perform several simplifications to assist the end-user in interpreting their overall risk. To simplify the communication of the risk, a general classification system is used. A risk class is assigned to each CCI in each of the three variants (it is important to understand that the risk classes are soft rules and not hard lines of classifying the risk. Each CCI risk should be considered and reviewed carefully). For example, based on the k_{4LF} and SRR for each CCI above is shown in Table 3. To further simplify the risk level, the predominant class for all CCIs, within a variant, dictate the risk class for that variant. However, if there is a CCI risk class III identified then the variant is automatically assigned a risk class III. For example, the predominate CCI classes from above are as follows:

– Variant 1 = Class I
– Variant 2 = Class I
– Variant 3 = Class I

Each variant will then have its own risk class as the variants are independent of each other. However, we can reduce the complexity even further. That is, if all three variants have the same risk class then it is safe to assume that the lubricant, no matter of the variant, is the risk class that they are all assigned. For example, all variants above are risk class I; therefore, this lubricant is a risk class I. However, if the variants differ in classes then one must consider them carefully. A conservative approach is to assume that the highest risk level across all three variants is the risk class for that lubricant but it does not necessarily indicate the true risk. Although, if the client knows the true variant being utilized then the proper risk can be assessed using the risk class for that variant (SeerWorks^{TM} Reliability can assist in determining the true variant if details on the additives and base oils used are provided). As part of SeerWorks^{TM} Reliability, the following assumptions are made:

- Risk is defined as the activity at the surface as a combination of the interfacial medium (e.g., lubricant, all components and their permutations) and the surface (e.g., bearing material) attributed to the lubricant and its constituents.
- Risk does not comment on SRR but rather uses a virtual SRR as a threshold for determining where, within the construction, the risk coming from the lubricant could play a significant role.
- The Risk Classes are soft rules for helping to determine the magnitude of risk levels with respect to either the lubricant or the construction. However, the detailed k_{4LF} for all the compounds and variance comparisons along with the corresponding SRR values should be considered carefully.

- It is important to understand that the k_{4LF} is calculated to give a simplified general risk based on the lubricant. SeerWorks™ Reliability uses small amounts of data to perform these calculations. However, with more accurate data from the end-user SeerWorks™ Reliability can be more precise with assessing the risk level.
- Typically, data for phase 0.1 and activity 0.1 are used for the general classification unless the construction model and surface roughness are known; surfaces that slide in tune with each other are considered of being in-phase (in-phase = 0 and out of phase unequal zero). The greater the phase the more the activity is out of phase.
- Each chemical compound interaction (CCI) is rated using the general classes.
- Count the number for each class across the CCIs and the predominate class then defines the variant class. However, if a class III is observed in any of the CCIs then the variant is automatically considered a class III (this must be performed on each of the three possible lubricant variants).
- If all three Variants have the same risk class then that defines the risk class for the lubricant. However, if the variants differ in classes then conservatively one must choose the highest risk class to define the risk for the lubricant.

2.3. Bearing Life Testing

As previously mentioned, sudden, irregular bearing failures that do not follow classical calculations defined in DIN ISO 281 [7], WECs in particular, are of great interest to manufacturers of bearings and the end-users. In order to evaluate the performance of bearings prior to their use in the field, bearing life tests are frequently performed. One of the most frequent tests is the FE-8 bearing test following DIN 51819-3 [18]. The FE-8 test is a standardized bearing test that is widely used in the industry to evaluate bearing life in the laboratory setting under predefined conditions (Fig. 2).

The bearings tested were cylindrical roller bearings (type 81212, brass cage at $\kappa < 1$ (explanation: $\kappa >> 1$ then full fluid film lubrication; $\kappa << 1$ mixed/boundary lubrication) at a speed of 750 rpm, an axial force 30 – 60 kN (1700 – 2000 MPa), and temperature set constant to 90°C at the oil circuit to simulate a typical WEC failure mode. The lubricants used consisted of the following constituents and various combinations: poly-α-olefine, ISO VG 46, ISO VG 68, ISO VG 100; zinc-/molybdenum- alkylphosphate as antiwear additives; corrosion preservatives on sulfonate base, amine phosphates; boronic acid esters; VI improvers on PMMA base; magnesiumsulfonate; antioxidants on the base on alkylated diphenylamines. This present test setup and lubricant compositions were selected to specifically create test environments that are known to facilitate lubricant-mediated WECs.

As part of the FE-8 bearing test post analysis, the bearing materials were assessed using common metallurgical techniques to investigate whether the indications for WECs were present. After cutting the samples perpendicular to the raceway, the surface was etched with nital solution (2% nitric acid in ethanol) to reveal the white areas to confirm the presence of WECs [20], [8], [10]. The raceway was also assessed for wear using scanning electron microscopy to determine early signs of surface-induced failure (SIF) and crack networks below the raceway; SIF together with surface-near crack networks are known to be a precursor of WECs. In particular, SIFs occur in bearings that have been tested for extended periods ($>>$ 50h).

2.4. Support Vector Machine (SVM) for Binary Linear Classification

Let us start by noting that there are numerous classification methods one can use, such as logistic regression, decision trees, random forests, neural networks and SVMs. The main driving force of one being more suitable than the other lies in the nature of the considered application and the available data. We considered (binary and multiclass) decision trees [21] as a good starting point, however the data available (cf. Sections 3.2 and 3.3) proved insufficient for these methods, as these provided overly simplistic classification rules. Accordingly, random forest methods seem excessive for the application considered, while regarding neural networks one has to worry about overfitting. Our attention naturally focused on SVMs and, to some degree, logistic regression, therefore the remainder of this section is devoted to very briefly introduce the methodology that the main conclusions rely on, followed by the presentation of the results.

Among the above mentioned and in general, one of the most widespread mathematical tools for binary linear classification is the support vector machine (SVM) ([22], [23], [24]). At a given data set of n data points $x_1, x_2, \ldots, x_n \in \mathcal{X}$, all of them belong to a m-dimensional vector space $\mathcal{X} \subseteq \mathbb{R}^m$. This is the case of n measurements on a system described by m characteristics were performed. Suppose that for each of the n measurements x_i, $i = 1, \ldots n$, there is a corresponding binary indicator $y_i \in \{-1, 1\}$ (e.g. indicating *success* (1) or *failure* (-1)).

If the convex hulls of $\{x_i : y_i = 1\}$ and $\{x_i : y_i = -1\}$ are disjointed, then there exists a linear binary classifier defined by $w \in \mathbb{R}^m$ and $b \in \mathbb{R}$: $w^\top x_i + b > 0 \iff y_i = 1$. In order to find the separating hyperplane that carries the largest discriminatory power we define b such that $d := \min_{y_i=1} w^\top x_i + b = \max_{y_i=-1} w^\top x_i + b$.

With the choice of d we have $w^\top x_i + b \geq d$ for all $y_i = 1$ and $w^\top x_i + b \leq -d$ for all $y_i = -1$, determining two hyperplanes, called *the hard margin*, around the classifier, the hyperplane lying in between the sets equidistantly from their convex hulls. These hyperplanes are supported by data points that are located exactly on them, called *support vectors* (hence the name SVM), as shown in Fig. 3.

The method is indicative of how it is computed as well: maximizing the distance, defined by $2d/||w||$, subject to keeping the points on the correct side with at least a given distance. Since for a given instance d is fixed, we maximize the distance by minimizing $||w||$. Correct classification is forced by linear constraints

$$y_i\left(w^\top x_i + b\right) \geq d \quad \forall x_i \in N, \tag{8}$$

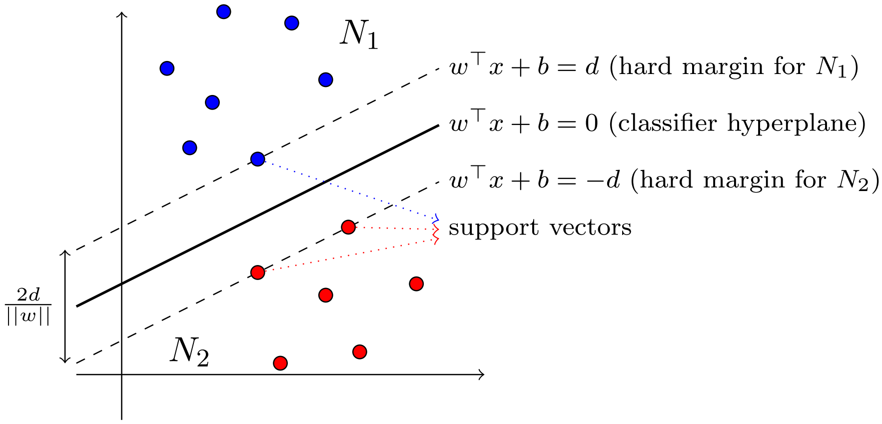

Fig. 3. SVM hard margin

while minimizing $||w||$ over (8) is a convex quadratic program (QP) that can be solved efficiently.

However, if the convex hulls are not linearly separable, the hard margin does not exist and the possibilities are embedding the data into a larger dimensional space (at the expense of less tractable computations), or consider the *soft margin* balancing, through parameter μ, the maximum distance $||w||$ and misclassification errors $z_i := \max(0, d - y_i(w^\top x_i + b))$, the smallest non-negative number satisfying $y_i\big(w^\top x_i + b\big) \geq d - z_i$. Consequently, the soft margin classifier can be computed using the convex QP

$$\begin{array}{llll} \min & \sum_{i=1}^{n} z_i + \mu||w|| & & \\ \text{s.t.} & y_i\big(w^\top x_i + b\big) + & z_i & \geq d \quad \forall x_i \in N \\ & & z & \geq 0 \end{array} \tag{9}$$

3. Results

3.1. Bearing Tests

The XFA and IR data of 24 different oil lubricants were analyzed independently using SeerWorks™ Reliability to be able to identify which lubricants would lead to bearings with WECs or SIF. Out of the 24 oils, 13 oils were tested on the FE-8 bearing test rig and each test was suspended if the maximum test period of 1000 h was reached - regardless of bearing survival status. The SeerWorks™ Reliability risk classifications performed prior to the FE-8 tests were compared to the actual results from the 13 FE-8 tests; 12 out of the 13 tests matched to their results (Table 4).

3.2. k_{4LF} and SRR

An experimental data set covering 24 lubricants collected in a sample size of more than 43,000 was available for this study. The data sample was focused on two parameters that were virtually expressed by the SeerWorks™ Reliability: the WEC coefficient k_{4LF} and the slip-roll-ratio SRR. We also used the risk class assessment by SeerWorks™ Reliability as an *ex ante* classification for 13 lubricants out of 24 as in Table 4, the remaining 11 being unclassified, using the general rules of risk classification as above. In terms of the parameter values, this means the following: if $k_{4LF} > 2$ and $0 < SRR < 30$, then the corresponding sample record likely leads to WEC; if $1 < k_{4LF} < 2$ and $0 < SRR < 30$, then SIF; while $k_{4LF} < 1$ and $SRR > 30$ typically leads to no breakdown. Table 5 and Fig. 4 show the distribution of the sample data, with the empirical classification rules depicted on the latter as rectangular areas in the k_{4LF}/SRR space.

The first things that can be noted from Table 5 and on Fig. 4 is the sizeable classification error between the ex ante classification and the empirical rules, as well as the significant noise in the *SRR* values. It is apparent that the sample data in terms of the latter is not suitable to deduce classification rules from, and as for the former, classifying the sample in terms of the k_{4LF} values rather simply just by grouping them based on shape of the data set, we can already reduce the error by 1.5% compared to the classification provided. In any case, there is room for improvement, therefore we focus on the components of the tested lubricants in order to derive classification rules.

3.3. Lubricant Component Analysis

In the present study, the calculations were applied to the composition of the 24 oil lubricants: altogether the tested lubricants consisted of 18 different components. That means for each oil i there is an $x_i \in \mathcal{X} \subseteq \mathbb{R}^{18}$, describing the components of the oil such that $\mathcal{X} = [0, 1]^{18}$ and $\sum_{j=1}^{18} (x_i)_j = 1$, i.e. the j-th coordinate of x_i, $(x_i)_j$ shows what is the ratio of component

Table 4
Risk class assessment using SeerWorks™ Reliability in relation to the outcome from physical laboratory tests.

Oil	Lubricant Variant[b]			Overall Lubricant Risk Class	FE-8 Results (SIF, WEC; hours)[a]
2	I	I	I	**I**	No Failure (> 1000)
3	II	I	III	**III**	Failure (SIF; 1000)[c]
4	III	III	III	**III**	Failure (SIF; 100)
7	III	I	I	**III**	Failure (WEC; 100)[c]
9	III	III	III	**III**	Failure (SIF; 100)
10	III	I	I	**III**	Failure (WEC)[c]
11	I	I	I	**I**	No Failure
13	I	I	I	**I**	Failure (SIF; 1000)[d]
15	III	I	I	**III**	Failure (SIF/WEC; 100)
16	I	I	I	**I**	No Failure (1000)
17	III	III	III	**III**	Failure (WEC; 100)
22	I	I	I	**I**	No Failure (1000)
24	III	III	III	**III**	Failure (WEC; 37.5)

[a] SIF=surface-induced failure; WEC=white-etching cracks; not match for oil number 13
[b] If class III only present once in the list of variants, worst case of Class III is assumed for Overall Risk Class
[c] Risk level dependent on Lubricant variant (further study required)
[d] Test suspended at max hours and bearing failed late (Oil 13 was predicted a “class I” and failed late but due to SIF).

Table 5
Experimental data summary of k_{4LF} and SRR parameters.

	Sample size	Min	Max	Mean	Var
k_{4LF}	43484	0.0018	16.5607	1.2339	2.6848
SRR	43389	−99.9998	97.2387	−1.2627	370.65

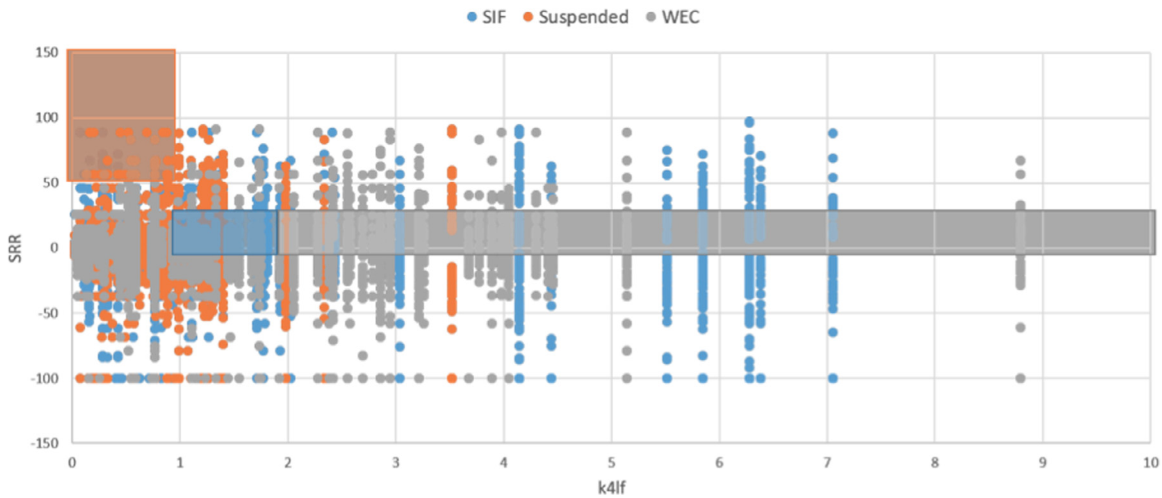

Fig. 4. Sample data of oil lubricants (k_{4LF}/SRR dimension)

j in oil i. Furthermore, there are y^{WEC}, y^{SIF}, y^{Susp} and y^{Unc} in $\{-1, 1\}$, indicating with value 1 whether each oil is classified as *WEC*, *SIF*, *Suspended* or otherwise *Unclassified* in the ex ante classification (training set).

Using the component data as an input for the SVMs, classification rules are derived from much less noisy samples compared to the experiments above, but the 18 components means that the dimension of $\mathcal{X}$ is still relatively low enough that allowed the analysis and interpretation of the results.

As previously described, the SVM method was applied to repeat the same process with the same $x_i, i = 1, \ldots, 24$ data points finding 4 binary classifiers for each y identifier. Starting with classifiers for y^{WEC} and y^{SIF}, it was found in both cases that the WEC (SIF) lubricants can be clearly separated by a (classifier) hyperplane from the non-WEC (non-SIF) lubricants, with corresponding distances of $d^{WEC} = 0.000175$ and $d^{SIF} = 3.5178 * 10^{-6}$. The separating hyperplanes were given by $b^{WEC} =$

Table 6
Binary classifier hyperplane coefficients for WEC and SIF.

Comp.	129	130	131	132	133	134	135	136	137
w^{WEC}	−0.001	0.0019	0.0032	0.0034	0.0008	-10^{-9}	0.0004	0.0013	-10^{-5}
w^{SIF}	0.0003	-10^{-4}	-10^{-4}	-10^{-4}	0.0006	-10^{-6}	0.0009	0.0004	0.0005
Comp.	138	141	142	143	144	145	146	154	155
w^{WEC}	0.0008	0.0002	0.0009	0.001	−0.0069	−0.0007	10^{-4}	−0.001	−0.003
w^{SIF}	−0.0003	-10^{-6}	-10^{-7}	−0.0002	0.0002	-10^{-6}	0.0003	-10^{-5}	−0.0004

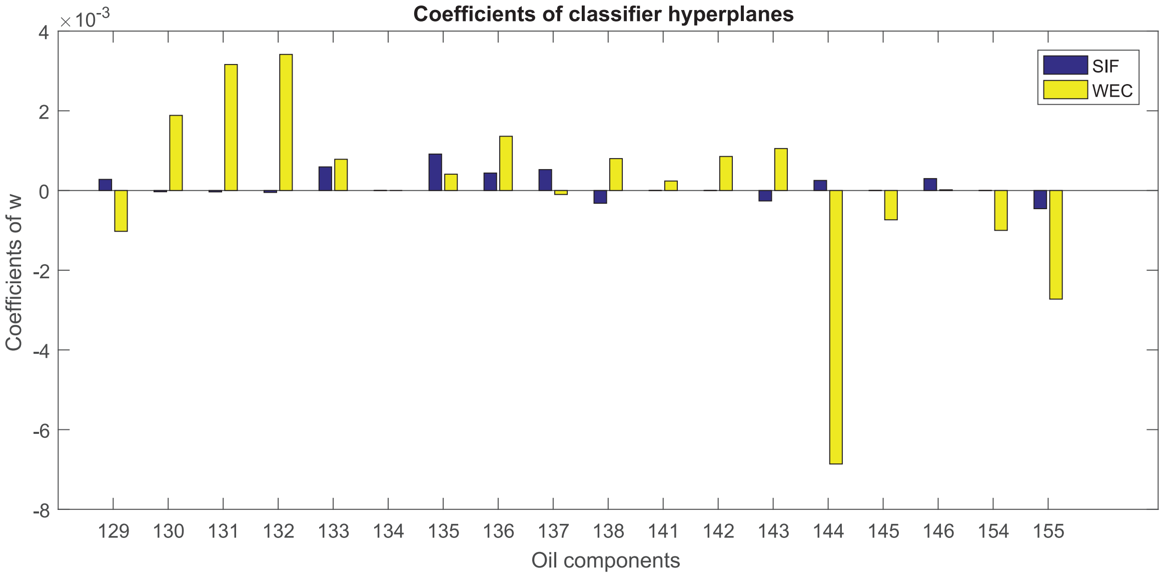

Fig. 5. Binary classifier hyperplane coefficients for WEC and SIF

Table 7
Binary classification consistency for each lubricant.

Lubricant	Ex ante	WEC	SIF	Suspended
2	Susp.			?
3	SIF		SIF	
4	SIF			
7	WEC	WEC		
9	SIF		SIF	
10	WEC	WEC		
11	Susp.			
13	SIF		SIF	
15	SIF		SIF	
16	Susp.			
17	WEC	WEC		
22	Susp.			?
24	WEC	WEC		

0.000175, $b^{SIF} = 3.5178 * 10^{-6}$ and the coefficients of w^{WEC} and w^{SIF} for each of the components are shown in Table 6 and Fig. 5.

The same tests were conducted for the remaining y^{Susp} and y^{Unc} with the only difference that in these cases the data points were not linearly separable, hence it was relied on the soft margin formulation. Once the separate binary classifier hyperplanes were determined, their consistency needed to be evaluated. Therefore, it was investigated whether the sets of lubricants classified as WEC and SIF were disjointed, while all of them being classified as non-Suspended at the same time (Table 7).

As shown in Table 7, the blank cells are classified as non-WEC (-SIF, -Suspended) for the WEC (SIF, Suspended) classifier, "?" denotes lubricants that violated (8) for the non-separable y^{Susp} case, while the remaining 11 blank rows of lubricants that were not classified ex ante (as in Table 7) were omitted. It can be seen that the binary classifiers that were generated for WEC and SIF were consistent with each other and the computed WEC and SIF classifying hyperplanes can be used as stand-alone lubricant classifiers, as an input for learning algorithms or for further analysis. For the latter purposes, it is of

Table 8
Most influential components for WEC and SIF classifiers.

Component	w^{WEC}	w^{SIF}
130	0.0019	
131	0.0032	
132	0.0034	
144	−0.0069	
155	−0.003	
133		0.0006
135		0.0009

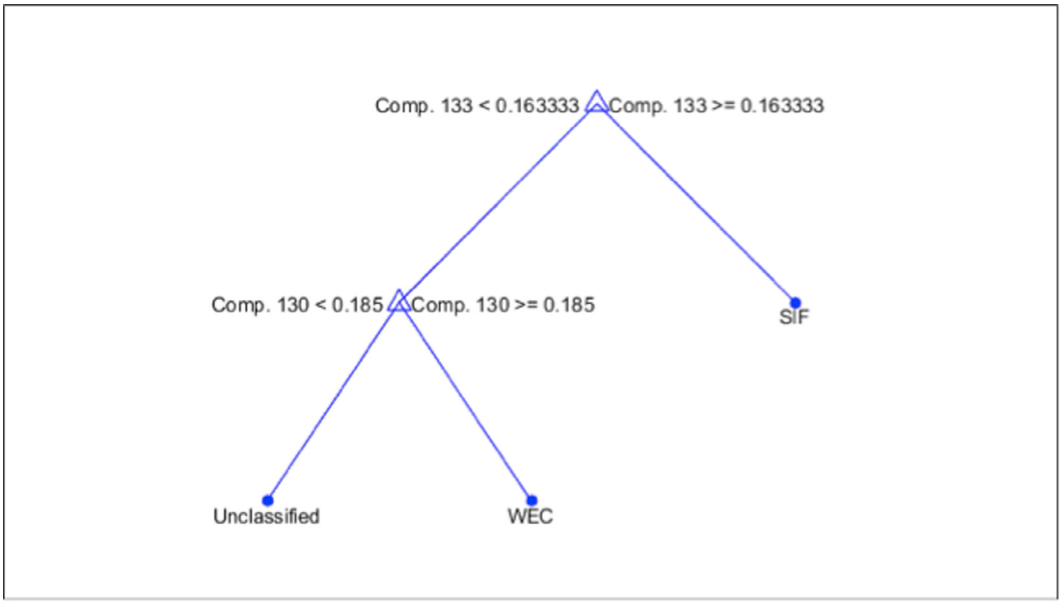

Fig. 6. Multiclass decision tree for classification

particular interest to highlight the (relatively) most influential components for each classifier in Table 8. For each classifying hyperplane, these are the defining coefficients with the largest absolute value, relative to the other defining coefficients of the hyperplane. Therefore, the coefficients have no stand-alone meaning, rather their magnitude indicates the influence each lubricant component has in the classification process. Thus, ideally, for the WEC and SIF binary classifiers, we would like to find different components that are most influential, the components mainly determining whether the lubricant will lead to WEC or not, or whether it will lead to SIF or not. Indeed, these formed two disjointed sets of components, indicating that the key lubricant components that are of increased risk for WEC and SIF have been identified.

3.3.1. Logistic Regression and Decision Trees

To provide a more complete understanding of the quality of the method used, the SVM methodology was compared with two other classification approaches, i.e. decision trees and logistic regression. Due to the lack of sufficient data, only overly simplistic classification rules were achieved for decision trees (Fig. 6)

In the case of logistic regression, parameters were estimated to achieve the highest probability of correct classification by maximizing the logarithm of the odds of correct classification. Hence, we consider the task of a maximum log-likelihood parameter estimation that requires solving a non-linear, but unconstrained optimization problem. Since in these problems most gradient search algorithms do not converge, Newton's method is used to obtain the estimated parameters, again, focusing on obtaining binary classifiers for each lubricant outcome category (e.g., WEC, SIF, etc.). It is of no surprise that the classification using logistic regression is completely in line with what was found using SVMs (Table 7). However, SVMs are not only less prone to outliers than logistic regression, interpreting the estimated parameters in the latter is far less straightforward compared to the case of SVMs with the soft and hard margin (Fig. 3) hyperplanes (coefficients in Tables 6 and 8), because these estimated parameters (Table 9) have no geometrical meaning[1].

[1] Moreover, Newton's method did not even converge to obtain a binary classifier for the 'Unclassified' class.

Table 9
Logistic regression estimated parameters for WEC and SIF binary classifier.

Comp.	-	129	130	131	132	133	134	135	136	137
WEC	8875	−89	5754	30176	−25228	−84	−97	900	−91	−73
SIF	−517800	5183	12872	803	3015	5280	5118	13443	5143	5080
Comp.	138	141	142	143	144	145	146	154	155	
WEC	−128	10817	−3041	−1152	−136	−1222	−89	−79	−89	
SIF	5694	41256	−4551	−3773	5342	−3098	5178	5235	5400	

4. Discussion

It has become increasingly important to preserve energy and material resources around the globe to boost sustainability. As part of industrial analytics, the rapidly growing predictive maintenance market has already shown to provide vast savings. These concepts are being used in a wide range of industrial applications by reducing failures and downtime while increasing longevity of machine elements. SeerWorks^{TM} Reliability, as a RaaS application, has proven to provide the user with stress points of an application that relates to events when the present lubricated contact is at risk for various types of failures and WECs in particular. In today's digital world, monitoring sensors are becoming increasingly connected as part of IIoT and the collected data can then be stored and processed on-premise or in the cloud. SeerWorks^{TM} Reliability processes data in a hardened private cloud (with third-party penetration testing grade of "exceeding best practices") that has been collected by sensors, or related equipment, either in real-time or offline. SeerWorks^{TM} Reliability simulates the lubricant interactions with the type of bearing surface to provide systemic risk assessment of a lubricated contact; therefore, the risk assessment is governed by the lubricant constituents (i.e., base oil and additives; lubricant changes due to e.g., water, detergents, sump residue and other degradation compounds), type of bearing surface materials, roughness, adjacent electricity and the interactions thereof. It needs to be noted that SeerWorks^{TM} Reliability provides a risk assessment for very early WEC failure that is mediated by the lubricant; it does not assess risk factors that can originate from material inclusions or other sources [8].

SeerWorks^{TM} Reliability, as a proactive condition monitoring application as part of predictive maintenance permits relationships to evolve – as in the case of real-life applications that are characterized by cascading events (e.g., tribology). While others teach machines based on data from past events, SeerWorks^{TM} Reliability uses deterministic approaches for these evolving/cascading events [25–28]. Thus, SeerWorks^{TM} Reliability can be applied to any type of system where cascading events play a role as in most real-world situations. SeerWorks^{TM} Reliability can show significant abnormalities early on, i.e. a risk class I lubricant changing to a risk class II/III lubricant as machine operating time progresses. As part of this process, lubricant samples are drawn, analyzed following industry-standard protocols on an ongoing basis where then SeerWorks^{TM} Reliability uses that standard data to report about significant abnormalities. This RaaS application deciphers the chemical compounds of the lubricant and the characteristics of the interface to provide proactive recommendations to guarantee machine reliability long before any major damage/root-cause initiates.

Bearing life can be separated into two phases: the time frame when the application is in smooth operation and the time frame when damage has been initiated, progresses and eventually leads to catastrophic failure; the latter is associated with the highest operational and maintenance cost as failure is inevitable (e.g., bearing replacements in an offshore wind turbine can cost millions of USD and reduce its availability to generate power). SeerWorks^{TM} Reliability manages the threats in smooth operation by reducing the uncertainty through recommendations that will eliminate damage initiation (Fig. 7).

There are plenty of monitoring sensors available that are able to detect damage and estimate the remaining useful life [29,30], e.g. including industry-standard lubricant analyses, acoustic, vibration, and particle contamination sensors. The risk classes that are created within SeerWork^{TM} Reliability have shown highly effective in providing proactive WEC alerts. Therefore, this permits the operator to perform evidence-based rather than calendar-based maintenance and implement counter measures that will extend the service life of the operation. The greater challenge is to predict how the application is performing during smooth operation, where no damage has yet developed, and there lies the greatest amount of uncertainty. The best time for reducing threats of damage exists in the smooth operation period; this is where proactive condition monitoring has its greatest value. Based on the holistic tribological approach, SeerWorks^{TM} Reliability provides a risk assessment and recommendations to reduce the uncertainty, reduce threats of damage, prolong the smooth operation period and extend availability.

Among the most widely used classification methods, e.g. neural networks, random forests, decision trees, logistic regression and SVMs, we applied the latter three and found that SVMs are the most suitable for the limited amount of data one can gather on the performance and effect of lubricant oils on bearing failure. The general classification generated by SeerWorks^{TM} Reliability has proven to be in line with the findings of FE-8 lab experiments, as shown by Table 4. According to Table 7, summarizing the classification with the binary classifiers generated by SVMs, it was found that these classifiers are consistent with each other: whenever a lubricant was classified as WEC (SIF) by their according classifier, it was classified as non-SIF (non-WEC) by the other corresponding classifier, as well as distinctively not being classified with the lubricant tests that were suspended. Additionally, comparing Tables 4 and 7, it was found that the SVM classifiers were in line with the findings of SeerWorks^{TM} Reliability and FE-8 bearing tests. Therefore, it is worthwhile to look into the coefficients of these binary classifiers that allowed us to identify the key lubricant compounds reliable for WEC and SIF, as

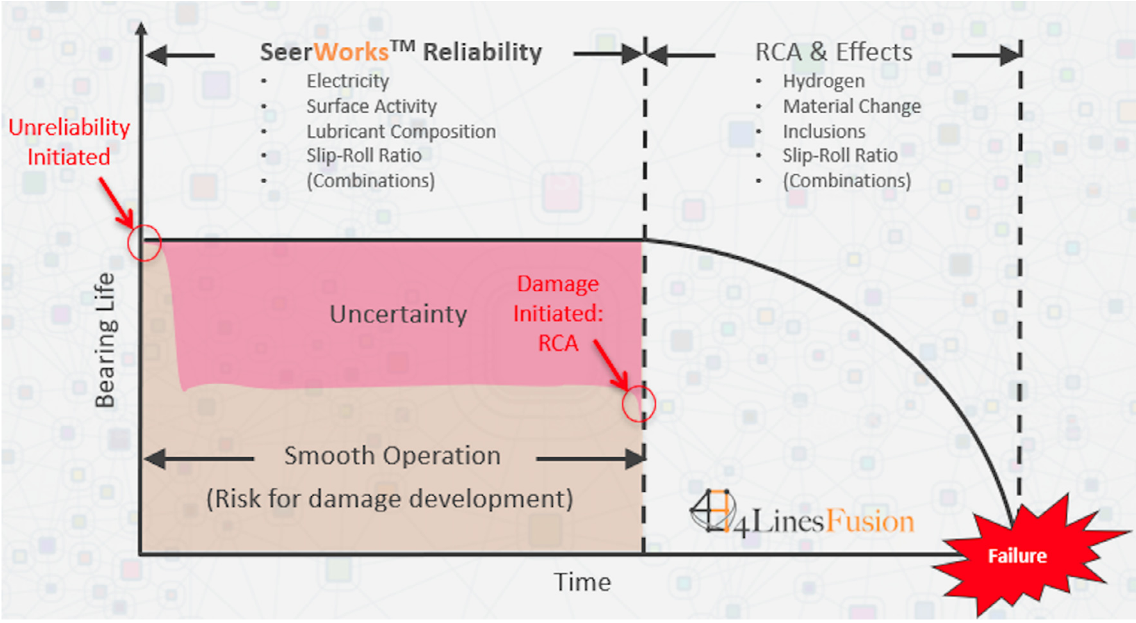

Fig. 7. Typical life-to-failure curve that is defined by a smooth operation interval (no/minor damage) and the interval that leads to failure where damage has been initiated (RCA, root-cause analysis).

Table 8 shows a distinct set of compounds for each class. As with most data processing tools, the results can be complex to interpret in order to create meaningful end-user knowledge in order to know what actions to take. Creating actionable knowledge is where the generalized classifications from SeerWorks^{TM} Reliability extends its role. The use of SVMs to verify the accuracy of SeerWorks^{TM} Reliability has shown that it is a very advanced and useful tool in monitoring bearing life in smooth operation.

Wear tests are inherently expensive to conduct and only provide one wear data point at a time. Thus, it is crucial to use methods that are able to accommodate small amounts of data points; as shown in the present study, SVM have shown to be very useful for this purpose. When large amounts of data points are available, different methods such as machine learning techniques are more favorable [31]. An additional analysis on the classification system of SeerWorks^{TM} Reliability was performed by Azzam et al. [32]; they applied state-of-the-art ANNs to isolate critical components in 700 oil samples and investigated their relationship with the WEC risk classes that were determined using SeerWorks^{TM} Reliability. The lubricant additive compounds were blinded and labeled by identification numbers to provide neutral dimensionless values for ANN processing. Firstly, the ANN processing was to evaluate the WEC risk classification levels based on the oil identity and percentage of constituting compounds and secondly, to identify oil compounds that affect the WEC risk in gearbox bearings. The key outcome from this study was the ability of ANN models to identify 8 out of 21 oil compounds as highly influential on WEC risk. Several ANN models were developed which gradually increased the classification accuracy on test oils to 99.8% by altering the networks architecture.

As a final comment, it is undisputable that digitization has penetrated all aspects of life. In the coming years, the digital transformation will further advance in waves and will present exciting opportunities for new technologies and innovations. While earlier digital waves have already occurred in automotive and logistics, followed by advancements in engineering, energy generation, healthcare and electrical engineering, future waves will follow in aerospace and in chemistry. Digitalization is not a new process and massive advancements in computing power, connectivity, and related technologies have solidified its presence over the past two decades; analog technologies will increasingly become a display of nostalgia – without compromising the essential need for human intelligence.

5. Conclusion

Proactive lubricant condition monitoring plays an important role in predictive maintenance. It allows the detection of miniscule system changes before damage has occurred and this is of much greater value than reactive measures. The proactive RaaS application presented in this study has proven to predict bearing failure due to white-etching cracks. It is also able to provide a classification system that allows the end-user to take immediate action to prevent further damage. In an industrial Analytics environment, this RaaS application is able to accommodate relatively small data sets to provide highly valuable information to the end-user. At its core, this technology is diametrical to the widely used artificial intelligence approaches. It is deterministic in nature and further supported by SVMs that confirmed the efficacy of the presented lubricant risk classification system.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by the Bundesministerium für Wirtschaft und Energie (BMWi, Berlin, Germany; Grant Number 0324082); the second author would like to acknowledge support from the National Research, Development and Innovation Fund (TUDFO/51757/2019-ITM, Thematic Excellence Program) as well as the University of Southampton, UK. We would further like to thank Baher Azzam from the Center of Wind Power Drives (CWD), RWTH Aachen, Germany, for discussions on data analysis; we are also grateful to Amanda Belanger (IBM, Ottawa, ON, Canada), Steven Yankowich (General Dynamics, Ottawa, ON), and the Telfer School of Management, University of Ottawa, for the business insights and future aspects of Industrial Analytics and as-a-Services business models.

References

[1] A. Gilchrist, Industry 4.0: the Industrial Internet-of-Things, Apress, 2016.
[2] MarketsandMarkets, Industrial analytics market by component (software, service), analytics type (descriptive, diagnostic, predictive, and prescriptive), deployment model, organization size, industry vertical, and region - global forecast to 2022, 2017, (https://www.marketsandmarkets.com/Market-Reports/industrial-analytics-market-90118141.html). [Online; accessed 22-August-2019].
[3] C.P. Gatica, M. Koester, T. Gaukstern, E. Berlin, M. Meyer, An industrial analytics approach to predictive maintenance for machinery applications, in: 2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA), IEEE, 2016, pp. 1–4.
[4] IoT Analytics, Predictive Maintenance Report 2019-2024 (https://iot-analytics.com/product/predictive-maintenance-report-2019-2024/). [Online; accessed 22-August-2019].
[5] H.P. Jost, Tribology origin and future, Wear 136 (1) (1990) 1–17, doi:10.1016/0043-1648(90)90068-L.
[6] G. Stachowiak, A.W. Batchelor, Engineering tribology, Butterworth-Heinemann, 2013.
[7] ISO 281:2007, Rolling bearings – Dynamic load ratings and rating life, Standard, International Organization for Standardization, 2007.
[8] M.-H. Evans, An updated review: white etching cracks (WECs) and axial cracks in wind turbine gearbox bearings, Materials Science and Technology 32 (11) (2016) 1133–1169.
[9] J. Lai, K. Stadler, Investigation on the mechanisms of white etching crack (WEC) formation in rolling contact fatigue and identification of a root cause for bearing premature failure, Wear 364 (2016) 244–256.
[10] S.W. Ooi, A. Gola, R.H. Vegter, P. Yan, K. Stadler, Evolution of white-etching cracks and associated microstructural alterations during bearing tests, Materials Science and Technology 33 (14) (2017) 1657–1666.
[11] M. Paladugu, D.R. Lucas, R.S. Hyde, Effect of lubricants on bearing damage in rolling-sliding conditions: Evolution of white etching cracks, Wear 398 (2018) 165–177.
[12] H.A. Al-Tameemi, H. Long, R.S. Dwyer-Joyce, Initiation of sub-surface micro-cracks and white etching areas from debonding at non-metallic inclusions in wind turbine gearbox bearing, Wear 406 (2018) 22–32.
[13] F. Manieri, K. Stadler, G.E. Morales-Espejel, A. Kadiric, The origins of white etching cracks and their significance to rolling bearing failures, International Journal of Fatigue 120 (2019) 107–133.
[14] A.I. Vakis, V.A. Yastrebov, J. Scheibert, L. Nicola, D. Dini, C. Minfray, A. Almqvist, M. Paggi, S. Lee, G. Limbert, J.F. Molinari, G. Anciaux, R. Aghababaei, S. Echeverri Restrepo, A. Papangelo, A. Cammarata, P. Nicolini, C. Putignano, G. Carbone, S. Stupkiewicz, J. Lengiewicz, G. Costagliola, F. Bosia, R. Guarino, N.M. Pugno, M.H. Müser, M. Ciavarella, Modeling and simulation in tribology across scales: An overview, Tribology International 125 (2018) 169–199.
[15] A.R. Katritzky, V.S. Lobanov, M. Karelson, Qspr: the correlation and quantitative prediction of chemical and physical properties from structure, Chemical Society Reviews 24 (4) (1995) 279–287.
[16] M. Bockstedte, A. Kley, J. Neugebauer, M. Scheffler, Density-functional theory calculations for poly-atomic systems: electronic structure, static and elastic properties and ab initio molecular dynamics, Computer Physics Communications 107 (1-3) (1997) 187–222.
[17] 4LinesFusion, The 4linesfusion white etching crack (WEC) coefficient - k_{4LF}, 2015, (www.4linesfusion.com). [Online; accessed 18-June-2019].
[18] DIN 51810-3, Testing of lubricants – Mechanical-dynamic testing in the roller bearing test apparatus FE8 – Part 3: Test method for lubricating oils – applied test bearing: axial cylindrical roller bearing, Standard, Deutsches Institut fur Normung E.V. (DIN), 2016.
[19] T. Haque, S. Korres, J.T. Carey, P.W. Jacobs, J. Loos, J. Franke, Lubricant effects on white etching cracking failures in thrust bearing rig tests, Tribology Transactions 61 (6) (2018) 979–990.
[20] A.D. Richardson, M.-H. Evans, L. Wang, R.J.K. Wood, M. Ingram, B. Meuth, The evolution of white etching cracks (WECs) in rolling contact fatigue-tested 100cr6 steel, Tribology Letters 66 (1) (2017) 6.
[21] L. Breiman, J. Friedman, R. Olshen, C. Stone, Classification and regression trees, Wadsworth Int. Group 37 (15) (1984) 237–251.
[22] H. He, A. Ghodsi, Rare class classification by support vector machine, in: 2010 20th International Conference on Pattern Recognition, IEEE, 2010, pp. 548–551.
[23] C. Huang, L. Davis, J. Townshend, An assessment of support vector machines for land cover classification, International Journal of Remote Sensing 23 (4) (2002) 725–749.
[24] J.S.L. Senanayaka, S.T. Kandukuri, Huynh Van Khang, K.G. Robbersmyr, Early detection and classification of bearing faults using support vector machine algorithm, in: 2017 IEEE Workshop on Electrical Machines Design, Control and Diagnosis (WEMDCD), 2017, pp. 250–255, doi:10.1109/WEMDCD.2017.7947755.
[25] M. Sipper, R.S. Olson, J.H. Moore, Evolutionary computation: the next major transition of artificial intelligence?, 2017.
[26] C.A. Pena-Reyes, M. Sipper, Evolutionary computation in medicine: an overview, Artificial Intelligence in Medicine 19 (1) (2000) 1–23.
[27] V. Aguiar-Pulido, M. Gestal, M. Cruz-Monteagudo, J.R. Rabuñal, J. Dorado, C. Munteanu, Evolutionary computation and qsar research, Current computer-aided drug design 9 (2) (2013) 206–225.
[28] E. Eberbach, Toward a theory of evolutionary computation, BioSystems 82 (1) (2005) 1–19.
[29] K. Leahy, C. Gallagher, P. O'Donovan, D. O'Sullivan, Issues with data quality for wind turbine condition monitoring and reliability analyses, Energies 12 (2) (2019).
[30] M. Wei, M. Chen, D. Zhou, Multi-sensor information based remaining useful life prediction with anticipated performance, IEEE Transactions on Reliability 62 (1) (2013) 183–198, doi:10.1109/TR.2013.2241232.
[31] B. Azzam, F. Harzendorf, R. Schelenz, W. Holweger, G. Jacobs, Pattern discovery in white etching crack experimental data using machine learning techniques, Applied Sciences 9 (24) (2019) 5502.
[32] B. Azzam, White Etching Crack Bearing Failure: Identification of High Risk Lubricant Compounds using Artificial Neural Networks, Conference presentation, Center for Wind Power Drives (CWD), 2019.