A Novel Taxonomy and Approaches for the Identification of Frequently Occurring Regularities in Degradation Processes of Engineering Systems

Fabian Mauthe[1], Christopher Braun[2,3], Julian Raible[2,3], Peter Zeiler[1], and Marco F. Huber[2,3]

[1] *Institute for Technical Reliability and Prognostics IZP, Esslingen University of Applied Sciences, Goeppingen, 73037, Germany*
fabian.mauthe@hs-esslingen.de, peter.zeiler@hs-esslingen.de

[2] *Institute of Industrial Manufacturing and Management IFF, University of Stuttgart, Stuttgart, 70569, Germany*
christopher.braun@iff.uni-stuttgart.de, julian.raible@iff.uni-stuttgart.de

[3] *Fraunhofer Institute of Manufacturing Engineering and Automation IPA, Stuttgart, 70569, Germany*
marco.huber@ieee.org

Abstract

The trend is shifting toward hybrid methods that incorporate prior knowledge into data-driven methods to address challenges in diagnostics and prognostics such as limited data, interpretability, and complex system behavior. While system-specific prior knowledge facilitates accurate, physically plausible modeling, the resulting hybrid model is typically tightly coupled to an individual engineering system. In contrast, general prior knowledge—such as fundamental physical laws or broadly applicable degradation knowledge—supports scalable, transferable models across various engineering systems. This opens the door to more adaptable approaches for diagnostics and prognostics, but the potential remains underexplored. To address this, a taxonomy is proposed that defines prior knowledge as frequently occurring regularities with four levels of validity, enabling hybrid methods to be characterized by their expected transferability. Two approaches are introduced and applied, both aimed at systematically identifying such regularities: one driven by expert knowledge, the other by data. Expert interviews further validate both the taxonomy and the identified regularities, establishing a foundation for developing transferable hybrid methods between various engineering systems.

1. Introduction

Prognostics and Health Management (PHM), though well-established, continues to underpin system reliability through monitoring, diagnosis, and the prediction of future health conditions (Zio, 2022). In contrast to diagnostics, which identifies current faults, prognostics relies on modeling complex, often nonlinear degradation processes that evolve under uncertainty. Accurate modeling of these dynamics is critical for estimating the Remaining Useful Life (RUL) of Engineering Systems (ESs), making prognostics one of the core challenges in PHM (Kordestani, Saif, Orchard, Razavi-Far, & Khorasani, 2021). To address this challenge, three main approaches are commonly employed: data-driven, physics model-based, and hybrid methods. While data-driven and physics model-based methods have long been established, hybrid methods are receiving increasing attention, as they provide a means to overcome the limitations inherent in both purely data-driven and physics-based modeling (Hagmeyer, Zeiler, & Huber, 2022). For instance, models must be interpretable, trustworthy, and robust under uncertainty—qualities that are difficult to guarantee when relying solely on data, which are often limited. In contrast, physics-based models offer high precision but are often impractical due to the complexity and limited observability of real-world degradation processes (Eker, Camci, & Jennions, 2016). In response, at the forefront of hybrid methods, Physics-informed Machine Learning (PIML) (Karniadakis et al., 2021) effectively integrates prior knowledge about the underlying physical system into data-driven methods to enhance the accuracy and generalizability of models, striking a balance between the flexibility of Machine Learning (ML) and the rigor of physics-based modeling.

Leveraging both data and prior knowledge, however partial, PIML offers a flexible and informed approach to improve modeling degradation (Deng, Nguyen, Medjaher, Gogu, & Morio, 2023). When this prior knowledge is specific to a given ES, such as the degradation behavior of a particular

https://doi.org/10.36001/IJPHM.2026.v17i1.4411

bearing, it enables accurate and physically plausible modeling for that context. Nevertheless, while effective in specific contexts, this approach also presents certain challenges. Implementation is non-trivial, as it requires identifying, formalizing, and integrating prior knowledge in a usable form, which is often a complex and highly system-specific task. Furthermore, grounded in system-specific prior knowledge, models tend to be tightly coupled to that context, making them difficult to transfer to other ESs or use cases (Bajarunas, Baptista, Goebel, & Chao, 2024). In contrast, leveraging more general prior knowledge—knowledge valid across multiple ESs or degradation processes—offers a pathway toward degradation foundation models (Y.-F. Li, Wang, & Sun, 2024). These models can serve as transferable, off-the-shelf solutions for a wide range of prognostic tasks, reducing the need for extensive re-engineering per system.

Hagmeyer et al. (2022) have taken initial steps toward delineating prior knowledge transferable across various ESs, accompanied by potential approaches for its integration into data-driven methods. Furthermore, the work of Bajarunas et al. (2024) represents one of the few examples in PHM where prior knowledge on a wider scope is explicitly addressed by integrating general knowledge about known cause-and-effect relationships within ESs and common degradation trends to estimate health indices using unsupervised learning. A similar approach is proposed by Chen, Ma, Zhao, Zhai, and Mao (2022), which incorporates a positive increment recurrence relationship to ensure monotonicity, keeping the learning process consistent with physical degradation of bearings. Beyond these examples, however, such prior knowledge is rarely used—indicating a preference for model performance, regardless of the significant overhead associated with tailoring models to individual ESs. Hence, this gap remains to be addressed through a systematic identification of relevant prior knowledge and an assessment of its applicability across different systems.

In this paper, a taxonomy is proposed that defines prior knowledge as Frequently Occurring Regularities (FORs) with four levels of validity, essentially enabling the characterization of hybrid methods based on their expected transferability. In the scope of this paper, prior knowledge refers to information about the degradation process that, while insufficient for complete physics-based modeling, captures broadly applicable regularities, such as typical degradation curve shapes or the monotonic decline of health states. The universal characteristics of the regularities are essential to facilitate the development of transferable hybrid methods. Following the definition of FORs, the question naturally arises of what methodologies can be employed for their identification. Despite its potential, a notable gap remains in the literature regarding the thorough exploration of such regularities. To address this gap, two approaches are proposed: one driven by expert knowledge, the other by data, both aimed at systematically identifying FORs. Therefore, this paper contains the following contributions:

1. A taxonomy is introduced that defines prior knowledge as FORs, specifies four levels of validity with respect to applications and domains, and enables characterization of the expected transferability of hybrid methods incorporating the corresponding regularities.
2. An expertise-driven and a data-driven approach are proposed for systematically identifying FORs across various degradation processes and ESs.
3. The identification and discussion of six FORs representing cross-domain characteristics considered valid in multiple domains, based on the expertise-driven approach, and 27 FORs representing cross-application characteristics identified through the data-driven approach.

The structure of this paper is as follows: Section 2 introduces the taxonomy, including the definition of FORs. Section 3 explores the expertise-driven and data-driven approaches used to systematically identify FORs. In Section 4, the results from both approaches and the findings from expert interviews conducted to evaluate the taxonomy and the identified FORs are presented, followed by an in-depth discussion in Section 5. Finally, Section 6 concludes the paper with a summary of key findings, an outlook on future work, and final reflections.

2. TAXONOMY OF FREQUENTLY OCCURRING REGULARITIES

The underlying principle of hybrid methods is to combine data-driven models with domain knowledge to improve aspects such as overall performance, interpretability, generalization, and robustness (Karniadakis et al., 2021). Since degradation data are inherently tied to the process from which they originate, the transferability of hybrid methods for diagnostics and prognostics hinges on the availability of broadly applicable prior knowledge to bridge contexts. However, the explicit consideration of such prior knowledge is rarely addressed in the existing literature. Consequently, a taxonomy is developed that defines prior knowledge as FORs and specifies four levels of validity across applications and domains (see Figure 1).

Regularities refer to systematic or consistent patterns or behaviors observed in physical phenomena or data. These regularities may manifest as recurring relationships, trends, or dependencies among variables that adhere to underlying physical laws or principles. They become frequently occurring when the same pattern or behavior is consistently observed across various configurations, i.e., environmental and operational conditions. Such regularities are referred to as application-specific FORs, whereas regularities that are additionally considered valid for different variations of this type of application are referred to as cross-application FORs. Moreover, regularities that are frequently occurring with respect to

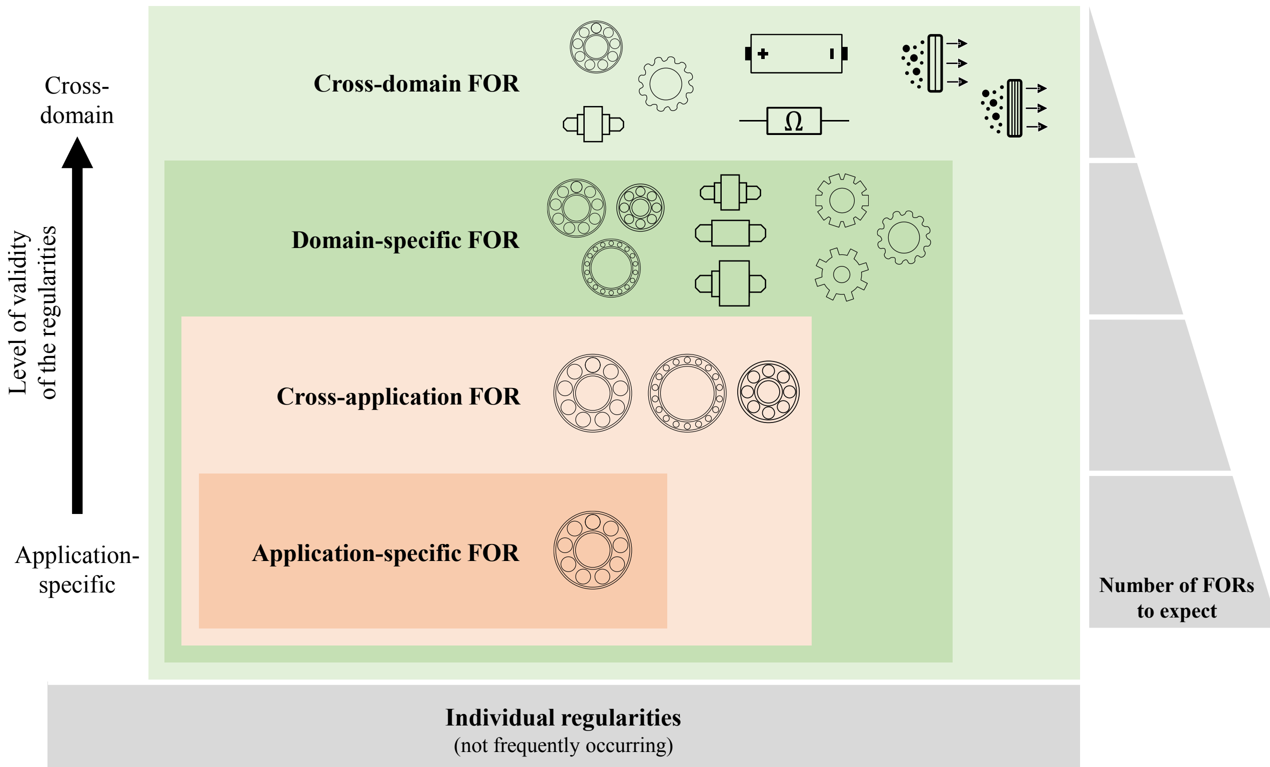

Figure 1. Illustration of the taxonomy along with an example for the distinction between the different levels of validity of the FORs. Application-specific: FORs occur under various configurations (i.e., variations such as distinct loads or rotational speeds, as well as different ESs in which the type of rolling bearing is installed) of the same type of rolling bearing. Cross-application: FORs are additionally valid for different types of rolling bearings. Domain-specific: The validity of FORs extends beyond rolling bearings to other applications within the same (mechanical) domain, e.g., shafts and gears. Cross-domain: FORs are additionally valid for different domains, e.g., electrical components and process technology.

different types of applications of one domain are considered domain-specific FORs, with cross-domain FORs characterizing regularities that apply within different domains.

As illustrated in Figure 1, the proposed taxonomy categorizes FORs according to their level of validity, ranging from application-specific FORs, which have the narrowest scope, to cross-domain FORs, which possess the broadest scope. Consequently, only regularities with cross-application characteristics can be considered for domain-specific FORs, and this principle similarly applies to domain-specific and cross-domain FORs. The term application refers to basic elements integrated with respect to the system under consideration, such as bearings and gears or resistors and capacitors, which form the fundamental components of ESs. The term domain refers to the domain from which the application originates. The domains associated with the previous examples would then be mechanical or electrical components. This definition specifies the lowest level of FORs, which essentially establishes the subsequent levels of the taxonomy, thus facilitating the adoption of the concept of FOR and setting the groundwork for the transferability of hybrid methods.

To gain more practical insight into the taxonomy's use and to clarify the corresponding validity levels, Figure 1 shows the following example:

- **Application-specific FOR**: Observations of FORs occur under various configurations of the same type of rolling bearing, such as varying in load, rotational speed, and the ESs in which the bearings are installed.
- **Cross-application FOR**: The observed FORs are additionally valid for different types of rolling bearings. Here, for example, a rolling bearing with different dimensions or with a different shape of rolling element.
- **Domain-specific FOR**: The validity of FORs extends beyond rolling bearings to other applications within the same domain. For example, this includes components such as shafts and gears from the mechanical domain.
- **Cross-domain FOR**: The observed FORs are additionally valid for different domains. In this example, this relates to the domains of electrical components and process technology.

This example demonstrates that the applicability of a FOR broadens with successive validity levels. Hence, at higher validity levels, fewer FORs are expected to be identifiable, as ESs become increasingly diverse (see the right-hand side of Figure 1). This applies particularly to FORs at the domain-specific and cross-domain levels, as the types of applications can vary significantly within these levels. For the cross-application level, the FOR must apply to at least two different variations of the same type of application. For the domain-specific level, the FOR must also apply to at least two different variations of another type of application. Although applications from the same domain are still being considered, these can already vary significantly. For the FOR to be valid at the cross-domain level, it must apply to at least one further domain, which means that the applications vary considerably more. Consequently, FORs that demonstrate validity across a wide range of applications and domains hold greater potential for the development of transferable hybrid methods. Nevertheless, cross-application FORs, such as those applicable to different types of bearings, remain highly relevant in practice, offering valuable contributions to the prognosis of bearing degradation using hybrid methods.

In general, the concept of FORs offers broad flexibility with respect to the type of prior knowledge being described, ranging from simple univariate trend behaviors, e.g., the irreversibility of degradation, to more complex multivariate dependencies, such as partial differential equations capturing the dynamics of the system being modeled. It is the definition of application that enables establishing a reference point with respect to the usage of the taxonomy. Hence, in the context of this paper, applications are restricted to components. Current research in the area of PHM frequently focuses solely on single components or applications, as shown by the works of Lei et al. (2020), Braig and Zeiler (2023), and Lei et al. (2018). Thus, the authors hold the view that this conceptual constraint does not represent an obstacle for the taxonomy's practical applicability.

Building on the in-depth explanation, the taxonomy must undergo systematic evaluation to ensure its successful adoption. For this purpose, expert interviews are conducted to assess both the taxonomy and the FORs identified through the expertise- and data-driven approaches.

3. APPROACHES FOR IDENTIFYING FREQUENTLY OCCURRING REGULARITIES

To address the question raised in Section 1 concerning the identification of FORs, two approaches—an expertise-driven and a data-driven one—are proposed. The expertise-driven approach systematizes unstructured knowledge on diverse degradation processes, incorporating the authors' experience, to derive candidate FORs. Examining degradation data is crucial for diagnostics and prognostics. Thus, the data-driven approach involves a systematic analysis of degradation data with the goal of identifying shared characteristics. In both approaches, a range of applications from various domains are considered, enabling the identification of FORs with varying levels of validity. The following section outlines both approaches in detail.

3.1. Expertise-driven Analysis of Degradation Processes

In the field of diagnostics and prognostics within PHM, expert knowledge—accumulated through years of research, practice, and PHM-specific observations—is frequently available in addition to data on the degradation process (Kordestani et al., 2021). Despite the broadly diversified nature of this expert knowledge, much of it remains unstructured and, as a result, is not readily applicable for integration into data-driven methods. Therefore, the objective of the expertise-driven approach is to systematically consolidate and process this expert knowledge in such a way that FORs can be effectively derived from it. To this end, the expertise-driven approach employs the following methodology. A systematic analysis of the literature on degradation processes is conducted, informed by the authors' experience, to identify FORs. Due to this procedure, objectivity cannot be guaranteed. Therefore, the results, i.e., the FORs identified (see Section 4.1), must be additionally evaluated as objectively as possible. Consequently, the expert interviews in Section 4.3 include a part dedicated to the evaluation of these FORs. This methodology enables (a) the identification of FORs derived from unstructured expert knowledge and (b) their assessment via expert interviews, resulting in an evaluated set of FORs.

The expertise-driven approach focuses on regularities, such as recurring patterns and characteristic signal progressions observed in degradation data. Such regularities include, for example, monotonic trends in degradation signals, condition-dependent signal changes or steps, and recurring degradation behaviors. The expertise-driven approach places particular emphasis on identifying regularities within degradation processes or data, which are often recognized through expert observation rather than computational analysis, such as the data-driven analysis in the following section. In addition, with the aforementioned focus of the expertise-driven approach, the aim is to uncover regularities that exhibit wide-ranging validity, i.e., FORs of cross-application and cross-domain characteristics, respectively. In that case, the identified FORs could be used across diverse degradation processes.

Regarding the expertise-driven approach, literature on degradation processes and the authors' experiences in diagnostics and prognostics serve as sources of expert knowledge. The PHM literature includes numerous systematic reviews and survey papers (e.g., Deng et al. (2023), Zio (2022), Kordestani et al. (2021), and Lei et al. (2018)), which are examined to derive regularities, i.e., recurring patterns and characteristic signal progressions. The identified candidate FORs are corroborated with the authors' experience. Alongside review pa-

Algorithm 1: Pseudocode of the proposed algorithm underlying the data-driven analysis for the identification of FORs from degradation datasets.

Input: Degradation dataset $\mathcal{D} = \{r_i\}_{i=1}^{N}$, which contains runs $r_i \in \mathbb{R}^{T_i \times M}$ consisting of time series $s_j^{(i)} \in \mathbb{R}^{T_i}$ corresponding to signals $j = 1 \ldots M$.

Output: Identified FORs, represented by one of eight distinct curve progressions l.

for *each signal* j **do**
  Retrieve time series $s_j^{(i)}$ from all runs r_i of dataset $\mathcal{D}$
  for *each time series* $s_j^{(i)}$ **do**
    Extract feature trajectory $f_k^{(i)}$ for every feature k
    for *each feature trajectory* $f_k^{(i)}$ **do**
      Fit bounded test function for every curve progression l using nonlinear least squares
      Compute its corresponding pseudo R^2 score
  for *each feature* k *and curve progression* l **do**
    Compute median pseudo R^2 score
    if *median pseudo* R^2 *score* > 0.9 **then**
      FOR identified regarding curve progression l for feature k of signal j

pers, the literature explicitly focusing on applications or domains pertinent to diagnostics and prognostics is also examined. This involves using the taxonomy of publicly available degradation datasets discussed in the work by Mauthe, Steinmann, Neu, and Zeiler (2025) as a reference for relevant applications and domains. Consequently, the domains primarily identified are electrical components, mechanical components, production systems, and process technology. Within these domains, relevant applications include batteries and capacitors, rolling bearings and gears, manufacturing processes such as milling, and filtration applications. Literature suitable for the examination includes, among others, the works of Shrivastava, Naidu, Sharma, Panigrahi, and Garg (2023) and Makdessi et al. (2015) on batteries and capacitors; of D. Wang, Tsui, and Miao (2018) and H. Zhou et al. (2022) on rolling bearings and gears; of Y. Zhou, Liu, Yu, Liu, and Quan (2022) and He, Shi, and Xuan (2022) on milling; and of Hagmeyer and Zeiler (2023) and Thomas, Penicot, Contal, Leclerc, and Vendel (2001) on filtration processes.

3.2. Data-driven Analysis of Degradation Processes

Identifying common degradation signal characteristics, such as monotonic trends or typical curve shapes, across different degradation datasets holds substantial value, as it enables the abstraction of generalizable degradation behavior. This simplified yet broadly applicable knowledge can serve as a foundation for developing methods that are not tightly coupled to a specific application, thereby facilitating transferability across different ESs. Motivated by this, the following approach systematically analyzes degradation datasets to potentially uncover regularities of different levels of validity regarding the taxonomy (see Section 2). The following section provides a detailed explanation of the algorithm underlying the data-driven analysis (see Algorithm 1).

The starting point is a degradation dataset $\mathcal{D}$ of N runs r_i of variable duration T_i representing multivariate time series as

$$\mathcal{D} = \{r_i \mid i = 1, \ldots, N\} .$$

Runs r_i are either run-to-failure or run-to-threshold experiments, where differences in duration arise from inherent variability in degradation as well as different configurations, such as varying environmental and operational conditions. Following this, each run containing measurements of M signals is defined as

$$r_i = \begin{bmatrix} s_{1,1}^{(i)} & s_{1,2}^{(i)} & \cdots & s_{1,M}^{(i)} \\ s_{2,1}^{(i)} & s_{2,2}^{(i)} & \cdots & s_{2,M}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ s_{T_i,1}^{(i)} & s_{T_i,2}^{(i)} & \cdots & s_{T_i,M}^{(i)} \end{bmatrix} \in \mathbb{R}^{T_i \times M}.$$

For each run r_i, the time series of signal j is denoted as

$$s_j^{(i)} = \begin{bmatrix} s_{1,j}^{(i)} & s_{2,j}^{(i)} & \cdots & s_{T_i,j}^{(i)} \end{bmatrix}^{\top} \in \mathbb{R}^{T_i}, \quad j = 1, \ldots, M.$$

Following this, the collection of all time series of signal j across all runs in dataset $\mathcal{D}$ is defined as

$$S_j^{\mathcal{D}} = \{s_j^{(i)}\} ,$$

which provides the foundation for analyzing the signals with the aim of identifying systematic or consistent patterns or behaviors that reflect the progression of degradation. For each signal $s_j^{(i)}$, L features are extracted in a sliding window manner, using a set of functions $g_{\text{extract},k}$. Here, $g_{\text{extract},k}$ computes the k-th feature for a given window of length w, with $k = 1, \ldots, L$. This produces $N_w^{(i)} = T_i - w + 1$ windows. Computing L features in each window results in a feature matrix for signal $s_j^{(i)}$:

$$F_j^{(i)} = \begin{bmatrix} f_{1,1}^{(i)} & f_{1,2}^{(i)} & \cdots & f_{1,L}^{(i)} \\ f_{2,1}^{(i)} & f_{2,2}^{(i)} & \cdots & f_{2,L}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ f_{N_w^{(i)},1}^{(i)} & f_{N_w^{(i)},2}^{(i)} & \cdots & f_{N_w^{(i)},L}^{(i)} \end{bmatrix} \in \mathbb{R}^{N_w^{(i)} \times L},$$

where each column $f_k^{(i)}$ represents the trajectory of one feature, showing its evolution over time. The collection of feature matrices across all runs for sensor j is then defined as

$$F_j^{\mathcal{D}} = \{F_j^{(i)}\} .$$

Table 1. Test functions and identifiable curve progressions are provided with bounds that constrain the parameter space so that fitted curves remain within the intended progression. In addition, representative examples of applications across various domains are included, for which the associated degradation signals exhibit the corresponding progression. These examples provide the rationale for selecting the presented curve progressions and motivate the systematic analysis of datasets to identify applications exhibiting similar characteristics.

Test Function $f(x)$	**Curve Progression**	**Bounds** (lower, upper) a	b	c	d	**Representative Examples**
$a \cdot x + b$	Linearly increasing	0 1	$-\infty$ ∞	- -	- -	Tool wear (Y. Zhou et al., 2022), internal resistance of a battery cell (Diao et al., 2022)
	Linearly decreasing	-1 0	$-\infty$ ∞	- -	- -	Fuel cell voltage (BenChikha et al., 2022)
$a \cdot \frac{1}{1+e^{b\cdot(x-c)}}$	Sigmoidally increasing	0 1	15 ∞	0.1 0.9	- -	Tool wear (Colantonio et al., 2021)
	Sigmoidally decreasing	0 1	$-\infty$ -15	0.1 0.9	- -	Battery capacity (Johnen et al., 2021)
$a \cdot e^{b\cdot(x-c)} + d$	Progressively increasing	0 ∞	0 ∞	$-\infty$ ∞	$-\infty$ ∞	Crack growth (Castillo et al., 2010), differential pressure in filtration processes (Eker et al., 2019)
	Progressively decreasing	$-\infty$ 0	0 ∞	$-\infty$ ∞	$-\infty$ ∞	Battery capacity (Diao et al., 2022)
	Degressively increasing	$-\infty$ 0	$-\infty$ 0	$-\infty$ ∞	$-\infty$ ∞	Internal resistance of a battery cell (Lin et al., 2025)
	Degressively decreasing	0 ∞	$-\infty$ 0	$-\infty$ ∞	$-\infty$ ∞	Battery voltage (Lin et al., 2025)

This serves as the basis for identifying FORs by examining the progression of each feature trajectory to uncover characteristic trends. The analysis is performed on feature trajectories normalized to $[0, 1]$ by fitting test functions (see Table 1) using nonlinear least-squares optimization. The Trust Region Reflective (TRF) algorithm (Branch, Coleman, & Li, 1999) is employed for the minimization task due to its robustness in handling bounded optimization problems. The number of steps per optimization is limited to $N_P \cdot 100$, with N_P being the number of parameters of the respective test function. Since the $N_{CP} = 8$ curve progressions to be identified are known, the test functions are initialized with favorable parameter values for the first attempt. If no optimal solution is obtained after a maximum of four additional attempts, with parameters randomly initialized within the bounds of the respective test function, the optimization is terminated. With the coefficient of determination (R^2 score), the goodness-of-fit regarding a feature trajectory $f_k^{(i)}$ and each fitted test function are rated. In the case of nonlinear regression, fits can be arbitrarily poor and may produce negative values of the coefficient of determination, which is then typically referred to as a pseudo R^2 score. Hence, each feature trajectory $f_k^{(i)}$ of feature matrix $F_j^{(i)}$ is attributed a total of $N_{CP} = 8$ pseudo R^2 scores, defined as

$$R^{2\ (i)}_{k,l} \in [-\infty, 1]\ , \quad l = 1, \ldots, N_{CP}\ ,$$

resulting in a total of $N \cdot N_{CP} \cdot L$ pseudo R^2 scores regarding collection $F_j^{\mathcal{D}}$. Then, for each curve progression, the median pseudo R^2 score across all runs is computed, yielding $N_{CP} \cdot L$ scores. Those greater than or equal to 0.9 are considered FORs. Hence, the algorithm underlying the data-driven analysis essentially enables the identification of eight distinct characteristic trends for each feature extracted from the signal being analyzed in the degradation dataset. Identified trends correspond to application-specific FORs.

Hereafter, implementation-specific aspects of the data-driven analysis are described step by step according to the algorithm. To ensure reliable analysis of trend behavior, runs that are right- or left-censored are excluded, as they do not permit examination of the full progression of the underlying degradation processes. Additionally, runs containing fewer than 20% of the median number of samples across all runs are also discarded. This threshold prevents curve fitting on feature trajectories with very few points (close to the number of parameters of the test functions), which could otherwise produce inflated pseudo R^2 scores. Moreover, each dataset may require specific preprocessing steps due to the high heterogeneity of publicly available degradation datasets. For example, irregularly sampled signals are resampled using linear interpolation, as the set of functions $g_{\mathrm{extract},k}$ requires a constant sampling rate to produce meaningful features.

For feature extraction, the Time Series Feature Extraction Library (TSFEL) toolbox (Barandas et al., 2020) is employed due to its versatility and suitability for time-series analysis. In this work, a specific subset of $L = 47$ features from the statistical, temporal, and spectral domains is considered. Features that cannot be meaningfully applied in a sliding-window manner, or that are not relevant for capturing degradation trends (e.g., the human range energy ratio), are omitted. A complete list of the selected features is provided in Table 3 of Appendix A. The behavior of functions $g_{\mathrm{extract},k}$ depends on whether the data are recorded continuously or periodically, a distinction the algorithm is designed to handle. Continuously recorded signals allow for overlap as the window moves across the signal. In this case, the window size w is set to 10% of the median number of samples across all runs in the dataset $\mathcal{D}$. Periodically recorded signals, however, necessitate a fixed window size and mandate that the sliding of the window occurs on a per-snapshot basis, due to the predefined temporal intervals between successive recordings.

To identify characteristic trends indicative of degradation, test functions are fitted to each feature trajectory. This enables the evaluation of whether the temporal evolution of each feature follows a typical curve progression, whose selection is informed by degradation modeling. Degradation signals typically follow a monotonic trend, either increasing or decreasing over time (Dersion, Goglio, Bajarunas, & Arias-Chao, 2025; Fink et al., 2025; Meeker, Escobar, & Pascual, 2022). These trends are commonly represented in degradation modeling by linear, sigmoidal, progressive, or degressive curve progressions (Meeker et al., 2022; Colantonio et al., 2021; Johnen et al., 2021). Table 1 lists representative examples of applications whose degradation signals follow the corresponding curve progressions. These examples motivate the selection of these curve progressions, as identifying the same trends in other applications would indicate characteristic trends that are valid across a wide range of applications.

As already mentioned, the median pseudo R^2 score is used to determine, based on the overall trend of the feature trajectory, whether feature k corresponds to a FOR regarding $S_j^{\mathcal{D}}$. Only fits with values close to 1 are considered relevant, for which the metric remains a reliable indicator of fit quality, even in the context of nonlinear test functions. The threshold of 0.9 was heuristically determined by examining the fitted functions in relation to the feature trajectories used for fitting. For pseudo R^2 scores ≥ 0.9, feature trajectories were clearly representative of their respective curve progression, as determined by careful visual assessment. This value discards features that do not closely follow one of the eight distinct curve progressions while allowing sufficient flexibility to identify valid trends, i.e., a perfect fit is not required for a feature to be classified as a FOR.

4. RESULTS

This section presents the FORs identified using the approaches previously introduced, along with their corresponding validity levels as defined by the proposed taxonomy. Furthermore, the findings from expert interviews, conducted to provide an additional layer of evaluation by critically reviewing the taxonomy and the identified FORs, are reported.

4.1. Expertise-driven Regularities Identified

In real-world scenarios, the degradation processes are typically irreversible, i.e., an ES is unable to restore its original health state or functionality without maintenance (Lei et al., 2018). Therefore, the progression of degradation signals or corresponding health indicators has either a gradient ≥ 0 or a gradient ≤ 0 depending on the degradation process, according to Dersion et al. (2025), Fink et al. (2025), Lei et al. (2018), and Sadoughi, Lu, and Hu (2019). This refers to the entire progression. Temporary deviations due to noise or varying environmental conditions are not taken into account. The same holds for application-specific deviations such as changing lubrication conditions or running-in effects. An expert cannot reliably determine the exact progression of signals or health indicators (e.g., linear or exponential) based solely on visual inspection. However, the gradients of the progression of degradation signals or health indicators are observable, from which two FORs are identified:

- **Increasing progression**: The degradation signal or health indicator has a gradient ≥ 0. Examples are shown by H. Zhou et al. (2022), Kordestani et al. (2021), D. Wang et al. (2018), and Lei et al. (2018) for peaks of vibration signals; by Lei et al. (2018) and Castillo et al. (2010) for crack length progression; by Meghoe, Loendersloot, and Tinga (2020), He et al. (2022) and Y. Zhou et al. (2022) for mechanical wear; and by Eker et al. (2016) for increasing differential pressure in a filtration process.
- **Decreasing progression**: The degradation signal or health indicator has a gradient ≤ 0. Examples are shown by Pan, Yang, Wang, and Chen (2020), Severson et al. (2019), Shrivastava et al. (2023), and Lin et al. (2025) for battery capacity; by E et al. (2025) and Makdessi et al. (2015) for capacitor capacity; and by Diao et al. (2022) and BenChikha et al. (2022) for voltage degradation in fuel cells.

These two FORs will be more generalized by representing them as a Health Index (HI). Using a predefined threshold for a degradation signal or a health indicator, the HI can be normalized for both progressions so that $\mathrm{HI} \in [0, 1]$. Where $\mathrm{HI} = 1$ means that the system is in a completely healthy condition and $\mathrm{HI} = 0$ means it has reached the threshold, i.e., is regarded as failed (Dersion et al., 2025; Berghout & Benbouzid, 2022). Then, for the time-dependent $\mathrm{HI}(t)$ follows

$$\frac{\mathrm{d}\,\mathrm{HI}(t)}{\mathrm{d}t} \leq 0\,. \tag{1}$$

As Eq. (1) shows, such FORs can be mathematically well formulated and applied as constraints, for instance, in physics-based regularization.

In addition to the overall trend of the signal, experts can often detect patterns and correlations within its progression over time. These finer-grained insights are particularly important when knowledge of the overall trend alone is insufficient to characterize the system's behavior. With this understanding, the following FORs are identified:

- **Known health stages**: The degradation signal or health indicator exhibits different stages that correlate with the actual health state. Some signals can be subdivided into distinct stages, each corresponding to a known degradation level. Examples are shown by Lei et al. (2018) and D. Wang et al. (2018) for a steplike progression in bearing degradation and by Eker et al. (2016) for different stages through the changing of filter mechanisms in a filtration process.
- **Correlation to operating condition**: The correlation between operating conditions and the progression of degradation signals or health indicators is well established. Operating conditions, such as loading, have a significant impact on the degradation process, as shown for different domains and datasets, respectively, by Lei et al. (2018) and Bajarunas et al. (2024).

In addition to the FORs associated with the progression of degradation, two further FORs are identified:

- **Boundaries**: Predictions in diagnostics and prognostics must fit within known boundaries, such as adherence to existing physical laws and known system boundaries. For example, the RUL cannot have negative values (Hoenig, Hagmeyer, & Zeiler, 2019; Saxena, Celaya, Saha, Saha, & Goebel, 2010).
- **Problem structure**: Evaluating established structures in the degradation process, such as characteristic points in the degradation signal where progression changes, allows subtasks to be defined and appropriate health indicators to be derived. This has been demonstrated for rotating systems by Zhu, Nostrand, Spiegel, and Morton (2014).

The FORs *increasing progression* and *decreasing progression* offer the potential, through their representation in Eq. (1), to be used directly without additional consideration of the respective degradation processes. Furthermore, these FORs only necessitate the respective degradation signal or health indicator, thereby ensuring their validity across various applications and domains. The FORs *known health stages*, *correlation to operating condition*, *boundaries*, and *problem structure* have to be specified for the respective degradation processes under consideration. For example, this means that the respective health stages or boundaries have to be defined individually. However, this can be performed across various applications and in different domains. Consequently, the six identified FORs represent cross-domain FORs.

Table 2. Summary of all applications analyzed by means of the data-driven approach proposed in Section 3.2, structured according to the originating domains and types of application. For each type of application, the signals that were used for feature extraction and subsequent testing regarding the consistency with the eight identifiable characteristic curve progressions are listed.

Domain	**Type of application (number of specific applications)**	**Signals**
Process technology	Filtration (2)	Differential pressure
Mechanical component	Bearing (5)	Vibration, temperature
Electrical component	Fuel cell (1)	Voltage
	Battery (6)	Discharge voltage, temperature
Manufacturing process	Milling (2)	Vibration, acoustic emission
Sum across all domains	**16**	

4.2. Data-driven Regularities Identified

In this work, 17 publicly available datasets from the overview[1] of Mauthe, Braun, Raible, Zeiler, and Huber (2024) regarding the task *prognosis* are analyzed. Two datasets arise from the same application, resulting in 16 different applications listed in Table 2. More detailed information regarding the datasets (e.g., number of runs, configurations defined by environmental and operational conditions, the analyzed signals, and their origin) and the FORs identified through the data-driven approach are presented in detail in Table 4 of Appendix B. The latter serve as the empirical basis for evaluating their respective levels of validity based on the taxonomy shown in Figure 1. In the following paragraphs, both application-specific FORs and cross-application FORs are described with respect to the types of application mentioned in Table 2. Then, potential domain-specific FORs and cross-domain FORs will be explored.

Filtration: The type of application *filtration* of the domain *process technology* comprises three datasets (PHM DATA CHALLENGE 2020 EUROPE - FILTRATION SYSTEM, KAGGLE - PREVENTIVE TO PREDICTIVE MAINTENANCE, and KAGGLE - PROGNOSIS BASED ON VARYING DATA QUALITY), which enable the analysis of the signal differential pressure. The last two datasets mentioned are recorded via the same filtration test bench and, as such, belong to the same application. Therefore, the identified FORs that match are considered application-specific. Both datasets' FORs match

[1] `https://arxiv.org/abs/2403.13694`

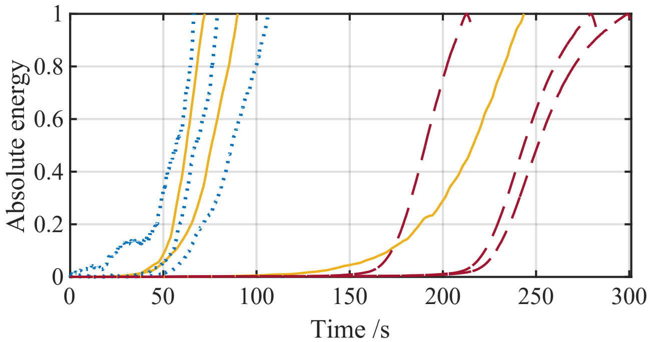

Figure 2. Temporal evolution of the feature absolute energy, extracted from the signal differential pressure, for three randomly selected runs of the filtration datasets. The features are normalized for comparison purposes due to the different units in the raw data. Dotted line: KAGGLE - PROGNOSIS BASED ON VARYING DATA QUALITY, solid line: KAGGLE - PREVENTIVE TO PREDICTIVE MAINTENANCE, and dashed line: PHM DATA CHALLENGE 2020 EUROPE - FILTRATION SYSTEM.

closely. Multiple features follow a progressively increasing trend (e.g., *absolute energy*, *area under the curve*, *autocorrelation*, *average power*, *max*, *min*, *mean*, *median*, and *root mean square*) but are also well approximated by a sigmoidally increasing trend. Considering the dataset PHM DATA CHALLENGE 2020 EUROPE - FILTRATION SYSTEM, the application-specific FORs identified are also described by sigmoidally increasing and progressively increasing trends, respectively. In addition, the feature *spectral distance* follows the trend of a sigmoidally decreasing, as well as progressively decreasing, function and, hence, yields another application-specific FOR.

Considering the three datasets originating from two different test benches simultaneously, several cross-application FORs can be determined. These are exactly the aforementioned features, as they appear in all datasets. All ten cross-application FORs of the *filtration* type of application in Table 4 of Appendix B are colored in orange. Regarding the progressively increasing and sigmoidally increasing trends, it can be inferred that the features experience a steep increase towards the end while tapering off shortly after. Both trends can be recognized in Figure 2, which depicts three randomly selected runs per dataset for the feature *absolute energy*.

Bearing: Regarding the type of application *bearing* within the domain *mechanical component*, five datasets were analyzed: PHM IEEE DATA CHALLENGE 2012 - FEMTO BEARING DATASET, NASA - BEARING DATASET, GITHUB - XJTU-SY BEARING DATASETS, MENDELEY - RUN-TO-FAILURE VIBRATION DATASET OF SELF-ALIGNING DOUBLE-ROW BALL, and ZENODO - BALL BEARINGS SUBJECTED TO TIME-VARYING LOAD AND SPEED CONDITIONS. Solely GITHUB - XJTU-SY BEARING DATASETS led to the identification of application-specific FORs regarding the signal vibration. Progressively increasing trends were obtained for a variety of features. Regarding the signal temperature that was available for both PHM IEEE DATA CHALLENGE 2012 - FEMTO BEARING DATASET and ZENODO - BALL BEARINGS SUBJECTED TO TIME-VARYING LOAD AND SPEED CONDITIONS, a degressively decreasing trend regarding the features *maximum frequency* and *spectral roll-off* was observed, marking cross-application FORs. These two are highlighted with blue in Table 4 of Appendix B.

Fuel Cell: The availability of fuel cell datasets was limited to one (PHM IEEE DATA CHALLENGE 2014 - FUEL CELL). The analysis regarding the signal voltage reveals that the statistical features *absolute energy*, *average power*, *mean*, *median*, and *root mean square* follow the trend of linearly decreasing, as well as degressively decreasing, functions. The same behavior can be attributed to the temporal features *area under the curve* and *autocorrelation*. *Spectral distance* is following a degressively increasing trend. Since only one dataset is available for this type of application, potential cross-application FORs cannot be identified.

Battery: The type of application *battery* regarding the domain *electrical component* contains six different datasets: NASA - RANDOMIZED BATTERY USAGE DATASET, NASA - HIRF BATTERY, NASA - LI-ION BATTERY AGING DATASET, MENDELEY - BATTERY DEGRADATION DATASET (FIXED CURRENT PROFILES AND ARBITRARY USES PROFILES), OXFORD BATTERY DEGRADATION DATASET, and ZENODO - DATA-DRIVEN CAPACITY ESTIMATION OF COMMERCIAL LITHIUM-ION BATTERIES FROM VOLTAGE RELAXATION. Multiple application-specific FORs were found for the datasets (with the exception of NASA - LI-ION BATTERY AGING DATASET and ZENODO - DATA-DRIVEN CAPACITY ESTIMATION OF COMMERCIAL LITHIUM-ION BATTERIES FROM VOLTAGE RELAXATION) for both signals (discharge voltage and temperature). Five of the six datasets are cycle-based datasets, meaning that charge and discharge cycles are performed alternately. Hence, the sliding window procedure described in Section 3.2 is performed on a cycle basis. The NASA - HIRF BATTERY dataset, however, is not a cycle-based dataset, providing an explanation for the strongly varying appearance of FORs in Table 4 of Appendix B and was therefore neglected in the comparison. The results are still reported.

Considering the datasets NASA - RANDOMIZED BATTERY USAGE DATASET and OXFORD BATTERY DEGRADATION DATASET, cross-application FORs are found regarding the voltage signal: The majority of features follow a linearly decreasing trend (e.g., *absolute energy*, *autocorrelation*, *signal distance*, and *spectral slope*). Also, linearly increasing, progressively increasing, and degressively decreasing trends can be observed for other features. Considering the temperature signal, further cross-application FORs can be determined with the majority of features following a linearly decreasing trend (e.g., *centroid*, *signal distance*, and *spectral slope*). All 15 cross-application FORs regarding the aforementioned two datasets are highlighted in purple in Table 4 of Appendix B.

Milling: Regarding the type of application *milling* within the domain *manufacturing process*, two datasets are analyzed with respect to vibration and acoustic emission signals: PHM DATA CHALLENGE 2010 - CNC MILLING MACHINE CUTTERS and NASA - MILLING DATASET. The PHM DATA CHALLENGE 2010 - CNC MILLING MACHINE CUTTERS revealed a variety of application-specific FORs regarding the vibration signal that are mainly progressively increasing trends. Both datasets have shown trend behavior for certain features regarding the acoustic emission. Overall, the two datasets show no consistent patterns, and thus no cross-application FORs could be identified for either signal.

A total of 27 cross-application FORs were identified across the examined datasets. Of these, the highest number is associated with the application *battery* (15), followed by *filtration* (10) and *bearing* (2). These findings carry significant potential as valuable building blocks for developing hybrid methods that are transferable across different types of the same application—an aspect of particular practical relevance, given that such components (i.e., batteries, filters, or bearings) are often monitored.

Domain-specific FORs: For the domain of *electrical component*, datasets from two types of applications were considered (see Table 2), allowing for the investigation of domain-specific FORs. Strictly speaking, since the application *fuel cell* includes only one dataset, no cross-application FORs can be determined—eliminating the basis for identifying domain-specific FORs. It is still worth mentioning, however, that the features *absolute energy*, *area under the curve*, and *autocorrelation* show a degressively decreasing trend for multiple *battery* datasets as well as the *fuel cell* dataset. The same holds for the feature *spectral distance*, which follows a degressively increasing trend. These findings provide strong indications of potential domain-specific FORs.

Cross-domain FORs: Considering Table 2, it is obvious that most domains include only datasets from one type of application, therefore making it impossible to identify multiple domain-specific FORs yet, which are the basis for cross-domain FORs. Nevertheless, it is worth mentioning that progressively increasing trends of multiple features are visible in different domains, like the vibration signals of *bearings* in the *mechanical component* domain, the vibration signals in *milling* within the *manufacturing process* domain, and the differential pressure signals in *filtration* within the *process technology* domain.

4.3. Expert Interviews

Conducting expert interviews aimed at two distinct goals: on the one hand, gathering unbiased insights from a variety of experts on the topic of degradation processes; on the other hand, evaluating the proposed taxonomy that hinges on the definition of prior knowledge as FORs. To that end, a total of ten participants from academia and industry were interviewed, who are active in different fields (each with expertise in different domains such as mechanical, electrical, production, and process technology) but share a background regarding reliability, physics-of-failure, degradation processes, and maintenance topics.

In line with the stated goals, the interviews followed a two-part structure and were conducted in a dialog format. In the first part, participants' backgrounds, knowledge, and experience in relation to regularities of degradation processes were investigated. Following this, participants were asked to derive criteria for assessing the level of validity of the regularities they had reported. To challenge their criteria, participants were tasked with evaluating three previously unknown, real or fictitious FORs according to the criteria they had proposed. The second part of the interview involved an evaluation of the taxonomy presented in Section 2, especially considering Figure 1. First, the taxonomy was explained in detail, including the rationale for defining prior knowledge as FORs. To rigorously test the usefulness of the proposed taxonomy, participants were then tasked with categorizing their previously reported regularities within the taxonomy. For this task, participants were provided with a blank version of Figure 1. Finally, the findings of Section 4.1 were evaluated.

During the first part, the participants named a wide variety of well-established regularities within their respective fields of expertise. For instance, those familiar with electrical components frequently mentioned characteristic degradation curves observed in specific applications, e.g., batteries and capacitors. Participants with expertise in mechanical components have often reported increasing power loss, temperature, and wear, which can be observed across different applications, e.g., bearings and gears. Across various applications and domains, participants frequently noted that system performance declines with degradation and that this phenomenon is often employed as a health indicator. When asked to define criteria for the validity of the regularities, the participants were often more critical, noting that the reported regularities are not consistently present and may occasionally deviate. Hence, most of the participants named criteria that are at the application level and a few that extend to the domain level. The criteria mentioned most frequently were operating and environmental conditions, as well as variations of an application for which the regularities are valid. These criteria correspond to application-specific and cross-application FORs within the proposed taxonomy, which can be interpreted as clear support for its validity. As a criterion for the domain-specific level, curve shapes of corresponding signals were mentioned by a few participants. The criteria defined by the participants could be used in part to categorize the three previously unknown FORs, depending on their validity.

During the second part, participants were presented with the proposed taxonomy (see Figure 1), which received broad support. They considered formalizing prior knowledge as FORs particularly valuable for standardizing research in the development of transferable hybrid methods. Hence, they viewed the taxonomy as a promising reference point, with its differentiation of levels of validity across different applications and domains. However, considerations were raised regarding potential difficulties in identifying cross-domain FORs. This observation emerged when participants were asked to name regularities for each level of validity (i.e., fill in a blank version of Figure 1). While doing this, many showed proficiency in recognizing cross-application FORs. While domain-specific FOR identification was prevalent, only a few cross-domain FORs were recognized. As a concrete example, the application most frequently referenced was the battery. The taxonomy was initially populated with a specific battery type, subsequently expanded to encompass various battery types at the cross-application level, and then refined to include batteries, capacitors, and transistors at the domain-specific level. Yet, the cross-domain level was frequently left blank. Another interesting example given analyzes a milling cutter type as an application. At the cross-application level, different types of milling cutters were mentioned, and at the domain-specific level, different machine tools, such as grinding wheels and turning tools, were mentioned. At the cross-domain level, tools from injection molding technology were mentioned in addition to machine tools. Despite acknowledging the potential of such FORs, participants noted that practical applications often diverge from known FORs. This aspect is crucial in FOR utilization. By the end of part two, all participants uniformly agreed on the FORs presented, which were identified via the expertise-driven approach. These were deemed valid and could be linked to the regularities discussed in part one. However, random deviations in practical applications of such FORs were noted and must be considered when applying FORs.

Overall, all participants offered positive and approving feedback. Although initially unfamiliar with the concept of FORs, participants swiftly developed an understanding and were subsequently able to reference regularities, several of which directly corresponded to those obtained via the expertise-driven approach, thus reinforcing the validity of the identified FORs. To help organize the regularities referenced by the participants, the proposed taxonomy was presented, enabling them to assign the regularities to the appropriate levels of validity. The necessity and idea, as well as the presented results of this work, were also rated very positively by the participants.

5. DISCUSSION

This section provides a detailed discussion of the results presented in the previous section. In addition, it discusses the limitations associated with the approaches dedicated to the FOR identification. Finally, conceptual examples are presented to illustrate the potential integration of FORs into hybrid methods as a means to convey the principle of transferability.

With its six identified cross-domain FORs, the expertise-driven approach provides regularities with the highest level of validity regarding the taxonomy. While some FORs (*increasing progression* and *decreasing progression*) can be directly utilized through their representation in Eq. (1), others require contextual adaptation to the respective degradation process. However, the applicability of all six FORs across diverse applications and domains highlights their potential as a basis for transferable hybrid methods. As already mentioned in Section 3.1, due to the notion of an expertise-driven approach, objectivity cannot be fully guaranteed. Consequently, a part of the expert interviews was dedicated to evaluating the identified FOR. All participants reached a consensus on the presented cross-domain FORs, which were considered valid. However, participants noted that random deviations can arise in practical applications and must be considered when applying these cross-domain FORs. Given the inherent subjectivity of the expertise-driven approach, considerable effort was made to enhance its objectivity by grounding it in a systematic review of the literature on degradation processes and interviewing ten experts to validate the results obtained by the authors. Nonetheless, some residual subjectivity may persist due to variations in expert interpretation and potential gaps or biases within the existing literature.

Having identified numerous application-specific FORs, along with 27 cross-application FORs (e.g., battery, filtration, and bearings), the data-driven approach proves to be well-suited for FOR identification across diverse applications. Nevertheless, its effectiveness is inevitably limited by factors such as data quality and quantity, while the subsequent analysis is inherently constrained by the scope of the available datasets. The heterogeneity of the analyzed datasets, coupled with a lack of standardization regarding the experiments conducted to acquire them, posed significant challenges, such as variations in sampling rates within single runs and across different runs, recordings with non-equidistant recording intervals, and missing values. This, in turn, impacted the degree of standardization achievable for the proposed algorithm, necessitating measures such as resampling based on linear interpolation, varying overlap regarding the sliding window, and removal of runs containing corrupt data. Owing to its internal sliding-window filtering, the algorithm inherently exhibits robustness to low-amplitude noise in real-world datasets. However, complex noise scenarios such as local noise artifacts could lead to misidentification of FORs. An additional methodological constraint concerns data of ESs affected by maintenance measures between collection intervals, which precludes their inclusion due to distortion of underlying degradation trends.

It is worth noting that the data-driven approach occasionally resulted in two FORs for a specific feature (e.g., sigmoidally increasing and progressively increasing), as both pseudo R^2 scores exceeded 0.9. However, only a single FOR is reported for that feature, acknowledging that its progression likely represents a mixture of both trends. The occurrence of such ambiguities represents a limitation of the data-driven approach. It should also be mentioned that some of the FORs identified relate to similar features. For example, if the feature *mean* is identified with an increasing progression, it is very likely that features such as *median* or *max* will also be identified with the same progression. Despite their similarities, it is still meaningful to consider these features as separate FORs, as they frequently have different pseudo R^2 values (see Table 4 in Appendix B).

The results of the data-driven approach confirm the FORs *increasing progression* and *decreasing progression* identified by the expertise-driven approach. The taxonomy is proposed alongside the assumption that as the level of validity increases, the expected number of FORs to be identified tends to decrease. This assumption is empirically supported by the results of both approaches, which identified 27 cross-application FORs using the data-driven approach and six cross-domain FORs using the expertise-driven approach. It is conceivable that within similar applications (cross-application), essentially the same underlying degradation mechanisms dominate, whereas across highly divergent applications (cross-domain), the mechanisms differ to such an extent that identified FORs may reflect statistical similarities rather than shared underlying degradation mechanisms.

Conceptual Examples on the Utilization of Identified Frequently Occurring Regularities

To illustrate the full pathway from the identification of FORs to their utilization, conceptual examples are outlined that follow two PIML principles—observational bias and learning bias (Karniadakis et al., 2021; Fink et al., 2025)—using the notation and regularities identified in this work. It should be noted that the actual utilization, i.e., the integration of identified regularities in hybrid methods, is outside the scope of this paper.

Observational bias leverages data that already embody the identified regularities or uses dedicated data augmentation strategies to imprint them into the training data distribution (Karniadakis et al., 2021). Hence, to introduce such bias, the identified FORs can be employed to generate synthetic data. Pre-training a model h_θ on these data provides an informed initialization of its weights. Fine-tuning on real data then starts from this advantageous starting point, yielding higher predictive accuracy than training from scratch (Hagmeyer et al., 2022). Specifically, auxiliary time series could be constructed that reflect generic degradation behavior corresponding to a given FOR. As indicated in Section 4.1, both monotonic increasing and monotonic decreasing degradation progressions can be generalized by representing them as $\mathrm{HI}(t)$, with $\mathrm{HI} \in [0,1]$ and with a monotonic decreasing gradient (see Eq. (1)). Time series $\widetilde{\mathrm{HI}}(t)$ can be generated to exhibit this behavior and subsequently employed to train hybrid models that can be transferred across domains. For example, in the case of filtration (see Section 4.2), auxiliary time series of differential pressure can be generated that exhibit progressively increasing trends, as identified by the data-driven approach. In this instance, pretraining h_θ on these data effectively enables transferring h_θ between different filtration systems.

The principle of learning bias is based on the idea that incorporating informed loss functions, constraints, or inference algorithms guides the learning process toward solutions that are consistent with prior knowledge or underlying physical laws (Karniadakis et al., 2021). A FOR can be incorporated as a constraint in the loss function $\mathcal{L}$ by adding an additional term that penalizes deviations from the expected behavior:

$$\mathcal{L} = \mathcal{L}_{\mathrm{data}} + w_{\mathrm{FOR}}\mathcal{L}_{\mathrm{FOR}} \ . \tag{2}$$

In addition to the data-driven loss $\mathcal{L}_{\mathrm{data}}$, which measures the difference between the model predictions and the observed data, an additional loss term $\mathcal{L}_{\mathrm{FOR}}$, weighted by w_{FOR}, enforces adherence to the FOR. This ensures that the model not only fits the training data but also respects the known regularities captured by the FOR, providing a principled way to embed prior knowledge into the learning process. One possible strategy is to represent a monotonically decreasing gradient, as described in Eq. (1). For example, by introducing a penalty on positive differences in the predicted health trajectory, such as

$$\mathcal{L}_{\mathrm{mono}} \ = \ \sum_k \left[\widehat{\mathrm{HI}}(t_{k+1}) - \widehat{\mathrm{HI}}(t_k)\right]_+^2 , \tag{3}$$

with $[x]_+ = \max(x, 0)$. Taking into account that this additional loss (i.e., monotonicity of $\mathrm{HI}(t)$) applies to all ESs, it follows that hybrid models incorporating $\mathcal{L}_{\mathrm{mono}}$ in their loss function can be transferred across domains. Furthermore, the loss function can be extended to encompass a loss term that enforces agreement between the shape of the predicted trajectory and the shape of an identified curve progression, i.e., a test function in Table 1 (see Section 4.2). For example, by quantifying the discrepancy between the prediction values $\widehat{f}(t)$ and the corresponding reference values $f(t)$, such as

$$\mathcal{L}_{\mathrm{shape}} \ = \ \sum_k \left(\widehat{f}(t_k) - f(t_k)\right)^2 . \tag{4}$$

The transferability of a hybrid model incorporating this additional loss term, $\mathcal{L}_{\text{shape}}$, depends on which applications adhere to the specific FOR underlying the reference values, potentially spanning from cross-application to cross-domain scenarios.

6. CONCLUSION AND OUTLOOK

Prior knowledge plays a pivotal role in advancing diagnostics and prognostics by means of hybrid methods. While current approaches often rely on system-specific knowledge, enabling accurate modeling but with limited applicability across ESs, general prior knowledge offers the potential for developing hybrid methods that are transferable across different ESs, eliminating or reducing the need for re-engineering. In light of this, a taxonomy is proposed, defining prior knowledge as FORs—recurring trends or relationships that adhere to underlying physical laws and are consistently observed across varying configurations, i.e., environmental and operational conditions. With FORs spanning four levels of validity across different applications and domains, this taxonomy offers a novel approach to characterizing the expected transferability of hybrid methods incorporating the respective regularities.

To identify FORs, an expertise-driven approach and a data-driven approach were developed. In the expertise-driven approach, FORs are derived from a systematic analysis of the literature on degradation processes, complemented by the authors' experience. Whereas, in the data-driven approach, publicly available degradation datasets form the basis for identifying FORs, aiming to infer characteristic trends indicative of degradation. Six cross-domain FORs and 27 cross-application FORs were identified with the expertise-driven and the data-driven approach, respectively. The proposed taxonomy and the identified FORs were evaluated via expert interviews, with participants expressing positive feedback.

By formalizing prior knowledge as FORs, this paper provides a novel, comprehensive contribution to the study of degradation processes. The proposed taxonomy provides a foundation for future research aimed at enhancing the transferability of hybrid methods in PHM by leveraging regularities of widespread applicability within the paradigm of PIML, potentially extending from cross-application to cross-domain scenarios. Future work building on the data-driven approach for identifying FORs could explore valuable extensions, such as analyzing a broader set of features or investigating additional curve progressions that may indicate degradation. In this context, a particularly promising direction is the integration of the continuous wavelet transform to obtain time–frequency representations. By simultaneously capturing temporal and spectral characteristics, this could complement and enhance the statistical, temporal, and frequency features already considered, enabling the discovery of subtle or overlapping degradation patterns that might otherwise remain hidden. Furthermore, the currently curve-fitting-based FOR determination can be extended to a rule-based one that uses a set of abstract rules that can be fitted within predefined ranges. In addition, specific noise scenarios, including local noise artifacts, can be taken into account to analyze the robustness of the data-driven approach. In general, the influence of varying operating conditions, which are currently considered implicitly in multiple configurations within a dataset (see Table 4 in Appendix B), can be investigated more thoroughly in future work. An analysis might be performed to assess how the number of configurations (i.e., operating conditions) affects the number of FORs identified in each configuration, allowing for an evaluation of the importance of the FORs. Following this, the parameters of the fitted curves of multiple experiments can be considered to investigate trends occurring along different operating and load conditions. These trends, in turn, potentially represent interesting prior knowledge that can be leveraged. Most importantly, the conceptual examples presented in Section 5 warrant exploration in an empirical study to determine whether the premise of increasing the transferability of hybrid methods by leveraging the identified FORs holds in practice. Nevertheless, further methods for the identification of FORs are still essential, as the suggested approaches are only one way of identifying FORs. In future work, alternative approaches, including learning-based ones, can also be investigated in order to extract FORs.

ACKNOWLEDGMENT

The authors wish to acknowledge that the research findings presented in this paper originated from the project TheoMation, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant no. 514247199.

In addition, the authors thank the anonymous reviewers for their valuable feedback and constructive suggestions, which helped to improve the overall quality of this paper.

The authors would like to thank the following participants for their willingness and cooperation in the expert interviews conducted for this research:

- Thomas Adolf, Fraunhofer Institute for Manufacturing Engineering and Automation IPA
- Dr. rer. nat. Benjamin Adrian, Fraunhofer Institute for Industrial Mathematics ITWM
- Patrick Barylla, Fraunhofer Austria Research GmbH
- Prof. Dr.-Ing. André Böhm, Esslingen University of Applied Sciences
- Dr. rer. nat. Florian Heiland, IMS Gear SE & Co. KGaA
- Prof. Dr.-Ing. Dominik Lucke, Reutlingen University & Fraunhofer Institute for Manufacturing Engineering and Automation IPA
- Dr.-Ing. Valentin Meimann, MML Solutions GmbH
- Thomas Rittler, Festo SE & Co. KG

- Prof. Dr.-Ing. Jan Singer, Esslingen University of Applied Sciences
- Dr.-Ing. Marco Werschler, IMS Gear SE & Co. KGaA

REFERENCES

Agogino, A. and Goebel, K. (2007). *Milling data set.* Moffett Field, CA. Retrieved 20.03.24, from `https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/`

Aimiyekagbon, O. K. (2024). *Run-to-failure data set of ball bearings subjected to time-varying load and speed conditions.* Retrieved 29.03.25, from `https://zenodo.org/records/10868257`

Bajarunas, K., Baptista, M. L., Goebel, K., & Chao, M. A. (2024). Health index estimation through integration of general knowledge with unsupervised learning. *Reliability Engineering & System Safety*, *251*, 110352. doi: 10.1016/j.ress.2024.110352

Barandas, M., Folgado, D., Fernandes, L., Santos, S., Abreu, M., Bota, P., ... Gamboa, H. (2020). TSFEL: Time Series Feature Extraction Library. *SoftwareX*, *11*, 100456.

BenChikha, K., Kandidayeni, M., Amamou, A., Kelouwani, S., Agbossou, K., & Abdelghani, A. B. B. (2022). Fuel cell ageing prediction and remaining useful life forecasting. In *2022 ieee vehicle power and propulsion conference (vppc)* (pp. 1–6). Piscataway, NJ: IEEE. doi: 10.1109/VPPC55846.2022.10003313

Berghout, T., & Benbouzid, M. (2022). A systematic guide for predicting remaining useful life with machine learning. *Electronics*, *11*(7), 1125. doi: 10.3390/electronics11071125

Birkl, C., & Howey, D. (2017). *Oxford battery degradation dataset 1.* Retrieved 29.03.25, from `https://ora.ox.ac.uk/objects/uuid:03ba4b01-cfed-46d3-9b1a-7d4a7bdf6fac`

Bole, B. and Kulkarni, C. and Daigle, M. (2014). *Randomized battery usage data set.* Moffett Field, CA. Retrieved 20.03.24, from `https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/`

Braig, M., & Zeiler, P. (2023). Using data from similar systems for data-driven condition diagnosis and prognosis of engineering systems: A review and an outline of future research challenges. *IEEE Access*, *11*, 1506–1554. doi: 10.1109/ACCESS.2022.3233220

Branch, M. A., Coleman, T. F., & Li, Y. (1999). A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM Journal on Scientific Computing*, *21*(1), 1–23.

Castillo, C., Fernández-Canteli, A., Castillo, E., & Pinto, H. (2010). Building models for crack propagation under fatigue loads: application to macrocrack growth. *Fatigue & Fracture of Engineering Materials & Structures*, *33*(10), 619–632. doi: 10.1111/j.1460-2695.2010.01475.x

Chen, X., Ma, M., Zhao, Z., Zhai, Z., & Mao, Z. (2022). Physics-informed deep neural network for bearing prognosis with multisensory signals. *Journal of dynamics, monitoring and diagnostics*, 200–207.

Colantonio, L., Equeter, L., Dehombreux, P., & Ducobu, F. (2021). A systematic literature review of cutting tool wear monitoring in turning by using artificial intelligence techniques. *MACHINES*, *9*(12), 351. doi: 10.3390/machines9120351

Deng, W., Nguyen, K. T., Medjaher, K., Gogu, C., & Morio, J. (2023). Physics-informed machine learning in prognostics and health management: State of the art and challenges. *Applied Mathematical Modelling*, *124*, 325–352. doi: 10.1016/j.apm.2023.07.011

Dersion, P., Goglio, D., Bajarunas, K., & Arias-Chao, M. (2025). Analytical health indices: Towards reliability-informed deep learning for phm. *International Journal of Prognostics and Health Management*, *16*(2). doi: 10.36001/ijphm.2025.v16i2.4262

Diao, W., Kim, J., Azarian, M. H., & Pecht, M. (2022). Degradation modes and mechanisms analysis of lithium-ion batteries with knee points. *Electrochimica Acta*, *431*, 141143. doi: 10.1016/j.electacta.2022.141143

E, L., Wang, J., Yang, R., Wang, C., Li, H., & Xiong, R. (2025). A physics-informed neural network-based method for predicting degradation trajectories and remaining useful life of supercapacitors. *Green Energy and Intelligent Transportation*, *4*(3), 100291. doi: 10.1016/j.geits.2025.100291

Eker, O. F., Camci, F., & Jennions, I. K. (2016). Physics-based prognostic modelling of filter clogging phenomena. *Mechanical Systems and Signal Processing*, *75*, 395–412. doi: 10.1016/j.ymssp.2015.12.011

Eker, O. F., Camci, F., & Jennions, I. K. (2019). A new hybrid prognostic methodology. *International Journal of Prognostics and Health Management*, *10*(2). doi: 10.36001/ijphm.2019.v10i2.2727

FCLAB Federation. (2014). *Ieee phm 2014 data challenge - fuel cell.* FR CNRS 3539, France.

Fink, O., Nejjar, I., Sharma, V., Niresi, K. F., Sun, H., Dong, H., ... Zhao, M. (2025). From physics to machine learning and back: Part ii - learning and observational bias in phm. *arXiv*, *2509.21207v1*, 1–58. Retrieved from `https://arxiv.org/pdf/2509.21207`

Gabrielli, A., Battarra, M., Mucchi, E., & Dalpiaz, G. (2024). Physics-based prognostics of rolling-element bearings: The equivalent damaged volume algorithm. *Mechanical Systems and Signal Processing*, *215*, 111435. Retrieved from https://www.sciencedirect.com/science/article/pii/S0888327024003339 doi: https://doi.org/10.1016/j.ymssp.2024.111435

Hagmeyer, S., Mauthe, F., & Zeiler, P. (2021). Creation of publicly available data sets for prognostics and diagnostics addressing data scenarios relevant to industrial applications. *International Journal of Prognostics and Health Management*, *12*(2).

Hagmeyer, S., & Zeiler, P. (2023). A comparative study on methods for fusing data-driven and physics-based models for hybrid remaining useful life prediction of air filters. *IEEE Access*, *11*, 35737–35753. doi: 10.1109/ACCESS.2023.3265722

Hagmeyer, S., Zeiler, P., & Huber, M. F. (2022). On the integration of fundamental knowledge about degradation processes into data-driven diagnostics and prognostics using theory-guided data science. In *Phm society european conference* (Vol. 7, pp. 156–165).

He, Z., Shi, T., & Xuan, J. (2022). Milling tool wear prediction using multi-sensor feature fusion based on stacked sparse autoencoders. *Measurement*, *190*, 110719. doi: 10.1016/j.measurement.2022.110719

Hoenig, M., Hagmeyer, S., & Zeiler, P. (2019). Enhancing remaining useful lifetime prediction by an advanced ensemble method adapted to the specific characteristics of prognostics and health management. In *Proceedings of the 29th european safety and reliability conference (esrel)* (pp. 1155–1162). doi: 10.3850/978-981-11-2724-3_0204-cd

Johnen, M., Pitzen, S., Kamps, U., Kateri, M., Dechent, P., & Sauer, D. U. (2021). Modeling long-term capacity degradation of lithium-ion batteries. *Journal of Energy Storage*, *34*, 102011. doi: 10.1016/j.est.2020.102011

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, *3*(6), 422–440.

Kordestani, M., Saif, M., Orchard, M. E., Razavi-Far, R., & Khorasani, K. (2021). Failure prognosis and applications—a survey of recent literature. *IEEE Transactions on Reliability*, *70*(2), 728–748. doi: 10.1109/TR.2019.2930195

Kulkarni, C. and Hogge, E. and Quach, C. and Goebel, K. (2015). *Hirf battery data set.* Moffett Field, CA. Retrieved 20.03.24, from https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/

Lee, J. and Qiu, H. and Yu, G. and Lin, J. (2007). *Bearing data set.* Moffett Field, CA. Retrieved 20.03.24, from https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/

Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to rul prediction. *Mechanical Systems and Signal Processing*, *104*, 799–834. doi: 10.1016/j.ymssp.2017.11.016

Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., & Nandi, A. K. (2020). Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, *138*, 106587. doi: 10.1016/j.ymssp.2019.106587

Li, X., Lim, B. S., Zhou, J. H., Huang, S., Phua, S. J., Shaw, K. C., & Er, M. J. (2009). Fuzzy neural network modelling for tool wear estimation in dry milling operation. *Annual Conference of the PHM Society*, *1*(1).

Li, Y.-F., Wang, H., & Sun, M. (2024). Chatgpt-like large-scale foundation models for prognostics and health management: A survey and roadmaps. *Reliability Engineering & System Safety*, *243*, 109850.

Lin, C., Tuo, X., Wu, L., Zhang, G., Lyu, Z., & Zeng, X. (2025). Physics-informed machine learning for accurate soh estimation of lithium-ion batteries considering various temperatures and operating conditions. *Energy*, *318*, 134937. doi: 10.1016/j.energy.2025.134937

Lu, J., Xiong, R., Tian, J., Wang, C., Hsu, C.-W., Tsou, N.-T., … Li, J. (2022). *Battery degradation dataset (fixed current profiles & arbitrary uses profiles).* Retrieved 29.03.25, from https://data.mendeley.com/datasets/kw34hhw7xg/3

Makdessi, M., Sari, A., Venet, P., Aubard, G., Chevalier, F., Préseau, R., … Duwattez, J. (2015). Lifetime estimation of high-temperature high-voltage polymer film capacitor based on capacitance loss. *Microelectronics Reliability*, *55*(9-10), 2012–2016. doi: 10.1016/j.microrel.2015.06.099

Mauthe, F., Bakir, E. M., Scheerer, T., & Zeiler, P. (2022). *Prognosis based on varying data quality.* Kaggle. Retrieved from https://www.kaggle.com/datasets/prognosticshse/prognosis-based-on-varying-data-quality doi: 10.34740/kaggle/dsv/3814464

Mauthe, F., Braun, C., Raible, J., Zeiler, P., & Huber, M. F. (2024). *Overview of publicly available degradation data sets for tasks within prognostics and health management.* Retrieved from https://arxiv.org/abs/2403.13694

Mauthe, F., Steinmann, L., Neu, M., & Zeiler, P. (2025). Overview and analysis of publicly available degradation data sets for tasks within prognostics and health management. In E. B. Abrahamsen, T. Aven, F. Bouder, R. Flage, & M. Ylönen (Eds.), *Proceedings of the 35th european safety and reliability conference and the 33rd society for risk analysis europe conference* (pp. 945–

952). Singapore: Research Publishing. doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P7412-cd

Meeker, W. Q., Escobar, L. A., & Pascual, F. G. (2022). *Statistical methods for reliability data* (Second edition ed.). Hoboken, NJ: Wiley.

Meghoe, A., Loendersloot, R., & Tinga, T. (2020). Rail wear and remaining life prediction using meta-models. *International Journal of Rail Transportation*, *8*(1), 1–26. doi: 10.1080/23248378.2019.1621780

Nectoux, P. and Gouriveau, R. and Medjaher, K. and Ramasso, E. and Morello, B. and Zerhouni, N. and Varnier, C. (2012). *Pronostia : An experimental platform for bearings accelerated degradation tests.* Denver, CO, USA. Retrieved 20.03.24, from `https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/`

Pan, R., Yang, D., Wang, Y., & Chen, Z. (2020). Performance degradation prediction of proton exchange membrane fuel cell using a hybrid prognostic approach. *International Journal of Hydrogen Energy*, *45*(55), 30994–31008. doi: 10.1016/j.ijhydene.2020.08.082

Sadoughi, M., Lu, H., & Hu, C. (2019). A deep learning approach for failure prognostics of rolling element bearings. In *2019 ieee international conference on prognostics and health management (icphm)* (pp. 1–7). Piscataway, NJ: IEEE. doi: 10.1109/ICPHM.2019.8819442

Saha, B. and Goebel, K. (2007). *Battery data set.* Moffett Field, CA. Retrieved 20.03.24, from `https://www.nasa.gov/intelligent-systems-division/discovery-and-systems-health/pcoe/pcoe-data-set-repository/`

Saxena, A., Celaya, J., Saha, B., Saha, S., & Goebel, K. (2010). Metrics for offline evaluation of prognostic performance. *International Journal of Prognostics and Health Management*, *1*(1). doi: 10.36001/ijphm.2010.v1i1.1336

Severson, K. A., Attia, P. M., Jin, N., Perkins, N., Jiang, B., Yang, Z., … Braatz, R. D. (2019). Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, *4*(5), 383–391. doi: 10.1038/s41560-019-0356-8

Shrivastava, P., Naidu, P. A., Sharma, S., Panigrahi, B. K., & Garg, A. (2023). Review on technological advancement of lithium-ion battery states estimation methods for electric vehicle applications. *Journal of Energy Storage*, *64*, 107159. doi: 10.1016/j.est.2023.107159

Thomas, D., Penicot, P., Contal, P., Leclerc, D., & Vendel, J. (2001). Clogging of fibrous filters by solid aerosol particles experimental and modelling study. *Chemical Engineering Science*, *56*(11), 3549–3561. doi: 10.1016/S0009-2509(01)00041-0

Wang, B., Lei, Y., Li, N., & Li, N. (2020). A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Transactions on Reliability*, *69*(1), 401-412. doi: 10.1109/TR.2018.2882682

Wang, D., Tsui, K.-L., & Miao, Q. (2018). Prognostics and health management: A review of vibration based bearing and gear health indicators. *IEEE Access*, *6*, 665–676. doi: 10.1109/ACCESS.2017.2774261

Zhou, H., Huang, X., Wen, G., Lei, Z., Dong, S., Zhang, P., & Chen, X. (2022). Construction of health indicators for condition monitoring of rotating machinery: A review of the research. *Expert Systems with Applications*, *203*, 117297. doi: 10.1016/j.eswa.2022.117297

Zhou, Y., Liu, C., Yu, X., Liu, B., & Quan, Y. (2022). Tool wear mechanism, monitoring and remaining useful life (rul) technology based on big data: a review. *SN Applied Sciences*, *4*(8). doi: 10.1007/s42452-022-05114-9

Zhu, J., Nostrand, T., Spiegel, C., & Morton, B. (2014). Survey of condition indicators for condition monitoring systems. *Annual Conference of the PHM Society*, *6*(1). doi: 10.36001/phmconf.2014.v6i1.2514

Zhu, Jiangong. (2022). *Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation.* Retrieved 29.03.25, from `https://zenodo.org/records/6405084`

Zio, E. (2022). Prognostics and health management (phm): Where are we and where do we (need to) go in theory and practice. *Reliability Engineering & System Safety*, *218*, 108119. doi: 10.1016/j.ress.2021.108119

BIOGRAPHIES

Fabian Mauthe received his B.Sc. degree in Mechanical Engineering and Mechatronics and his M.Sc. degree in Mechanical Engineering from Furtwangen University of Applied Science, Germany. He is currently pursuing the PhD degree in cooperation with the University of Stuttgart. He is also working as a Research Assistant with the Institute for Technical Reliability and Prognostics (IZP), Esslingen University of Applied Sciences. His research interests include diagnostics and prognostics, condition monitoring, data-driven methods, and hybrid methods.

Christopher Braun received his B.Sc. degree in Sportsmedical Engineering from University of Applied Sciences Koblenz and his M.Sc. degree in Applied Physics from the University of Koblenz. He is currently pursuing a PhD in cooperation with the University of Stuttgart, where he works as a Research Assistant at the Institute of Industrial Manufacturing and Management (IFF), while also working at the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA). Within the department of Cognitive Production Systems (IFF) and the research unit of Machine Vision and

Artificial Intelligence (IPA), respectively, his focus lies on leveraging the potential of artificial intelligence in an interdisciplinary fashion. Specifically, his research interests revolve around the integration of domain-specific knowledge and prior information into machine learning methodologies, aiming to enhance algorithmic capabilities and robustness within the realm of informed machine learning.

Julian Raible received the Dipl.-Ing. degree in mechanical engineering from the Technical University of Kaiserslautern. He is currently pursuing a PhD in cooperation with the University of Stuttgart, where he works as a Research Assistant at the Institute of Industrial Manufacturing and Management (IFF), while also working at the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA). Within the department of Cognitive Production Systems (IFF) and the research unit of Machine Vision and Artificial Intelligence (IPA), respectively, his research interests include Physics-informed Machine Learning and Prognostics and Health Management.

Peter Zeiler received the Dipl.-Ing. and Dr.-Ing. degrees in mechanical engineering from the University of Stuttgart. He worked as a Team Leader of Research and Development at Robert Bosch GmbH. Thereafter, he became the Head of the Department of Reliability Engineering and the Department of Drive Technology, Institute of Machine Components, University of Stuttgart. In 2017, he joined the Esslingen University of Applied Sciences, Germany. Since 2021, he has been associated with the Faculty of Engineering Design, Production Engineering and Automotive Engineering, University of Stuttgart. He is currently a Professor with the Faculty of Mechanical and Systems Engineering, Esslingen University of Applied Sciences, where he is the Head of the Institute for Technical Reliability and Prognostics (IZP). His research interests include prognostics and health management, methods for calculating reliability under operating conditions, modeling and simulation of reliability, and availability for complex systems. He is committed to Verein Deutscher Ingenieure (VDI) (English: Association of German Engineers). Then, he is the Chairperson of the Safety and Reliability Advisory Board, and the Technical Committees for Prognostics and Health Management and Monte Carlo Simulation. Furthermore, he is the Conference Chair of the German Conference Fachtagung Technische Zuverlässigkeit (TTZ) (English: Technical Reliability Conference).

Marco F. Huber received his diploma, Ph.D., and habilitation degrees in computer science from the Karlsruhe Institute of Technology (KIT), Germany, in 2006, 2009, and 2015, respectively. From June 2009 to May 2011, he was leading the research group Variable Image Acquisition and Processing of the Fraunhofer IOSB, Karlsruhe, Germany. Subsequently, he was Senior Researcher with AGT International, Darmstadt, Germany, until March 2015. From April 2015 to September 2018, he was responsible for product development and data science services of the Katana division at USU Software AG, Karlsruhe, Germany. At the same time he was adjunct professor of computer science with the KIT. Since October 2018 he is full professor with the University of Stuttgart. He further is scientific director for digitalization and artificial intelligence and head of the research unit Artificial Intelligence and Machine Vision with Fraunhofer IPA in Stuttgart, Germany. His research interests include machine learning, planning and decision making, machine vision, and robotics.

APPENDIX A

Table 3. Specific selection of features (47 in total) considered for extraction. Naming in accordance with TSFEL's convention.

Spectral	**Temporal**	**Statistical**
Fundamental frequency	Area under the curve	Absolute energy
Max power spectrum	Autocorrelation	Average power
Maximum frequency	Centroid	Entropy
Median frequency	Mean absolute diff	Interquartile range
Power bandwidth	Mean diff	Kurtosis
Spectral centroid	Median absolute diff	Max
Spectral decrease	Median diff	Mean
Spectral distance	Negative turning points	Mean absolute deviation
Spectral entropy	Neighbourhood peaks	Median
Spectral kurtosis	Positive turning points	Median absolute deviation
Spectral positive turning points	Signal distance	Min
Spectral roll-off	Slope	Peak to peak distance
Spectral roll-on	Sum absolute diff	Root mean square
Spectral skewness	Zero crossing rate	Skewness
Spectral slope		Standard deviation
Spectral spread		Variance
Spectral variation		

APPENDIX B

Table 4. Listing of the results regarding the data-driven identification of dataset-specific FORs based on 17 degradation datasets. The generalizability of the FOR identified strongly depends on the quantity of data, i.e., available number of runs (R), as well as the quality of data, i.e., the representativeness and completeness characterized by the number of different configurations (C) applied, such as environmental and operational conditions, as well as failure modes (FM) present in the respective dataset. R, C or FM are indicated as not applicable (n.a.), if no information was provided by the host of the dataset. The proposed approach enabled identifying eight distinct curve progressions: linearly increasing (li), linearly decreasing (ld), sigmoidally increasing (si), sigmoidally decreasing (sd), progressively increasing (pi), progressively decreasing (pd), degressively increasing (di), degressively decreasing (dd). The trajectories of the features were analyzed to determine if their trends could be modeled by one of these curve progressions by solving a nonlinear least squares problem. The resultant curves are considered an application-specific FOR, if the median pseudo-R^2 score across all runs accumulates to at least 0.9. If no application-specific FORs could be identified at all, it will be marked as n.a. accordingly. The color coding refers to cross-application FORs, with orange marking cross-application FORs within the *filtration* type of application, blue marking cross-application FORs within the *bearing* type of application and purple marking cross-application FORs within the *battery* type of application.

Dataset *(Domain, Application)*	**Runs, Configs, Failure Modes**	**Signal**	**Feature**	**Curve progression** li	ld	si	sd	pi	pd	di	dd
PHM Data Challenge 2020 Europe - Filtration System[2] *(Process technology, Filtration)*	R=32, C=8, FM=1	Differential Pressure	Absolute energy			1.0		0.96			
			Area under the curve			0.98		0.96			
			Autocorrelation			1.0		0.96			
			Average power			1.0		0.96			
			Max			0.99		0.93			
			Mean			0.98		0.96			
			Median			0.98		0.96			

Continued on next page

Table 4 – continued from previous page

Dataset *(Domain, Application)*	**Runs, Configs, Failure Modes**	**Signal**	**Feature**	li	ld	si	sd	pi	pd	di	dd
			Median frequency				0.94				
			Min			0.92		0.93			
			Root mean square			0.98		0.96			
			Spectral distance				0.99		0.93		
			Spectral entropy				0.95				
			Spectral kurtosis			0.95		0.94			
			Spectral skewness					0.93			
			Zero crossing rate				0.99				0.94
Kaggle - Preventive to Predictive Maintenance[3] *(Process technology, Filtration)*	R=100, C=7, FM=1	Differential pressure	Absolute energy			0.99		1.0			
			Area under the curve			0.93		1.0			
			Autocorrelation			0.99		1.0			
			Average power			0.99		1.0			
			Max			0.91		0.99			
			Mean			0.93		1.0			
			Mean absolute deviation					0.91			
			Mean diff					0.91			
			Median			0.93		1.0			
			Min			0.94		0.99			
			Peak to peak distance					0.92			
			Root mean square			0.93		1.0			
			Slope					0.91			
			Spectral distance				0.95		1.0		
			Spectral slope								0.91
			Standard deviation					0.92			
Kaggle - Prognosis based on Varying Data Quality[4] *(Process technology, Filtration)*	R=55, C=4, FM=1	Differential pressure	Absolute energy			0.97		0.99			
			Area under the curve			0.95		0.99			
			Autocorrelation			0.97		0.99			
			Average power			0.97		0.99			
			Entropy			0.96				0.91	
			Max			0.92		0.97			
			Mean			0.95		0.99			
			Median			0.93		0.99			
			Median frequency				0.98				
			Min			0.96		0.97			
			Root mean square			0.95		0.99			

Continued on next page

Table 4 – continued from previous page

Dataset *(Domain, Application)*	**Runs, Configs, Failure Modes**	**Signal**	**Feature**	**Curve progression** li	ld	si	sd	pi	pd	di	dd
			Spectral centroid				0.93				
			Spectral decrease				0.9		0.97		
			Spectral distance				0.95		0.98		
			Spectral kurtosis			0.93		0.95			
			Spectral skewness					0.94			
			Spectral slope				0.93				
			Zero crossing rate				0.99				
PHM IEEE Data Challenge 2012 - FEMTO Bearing Dataset[5] *(Mechanical component, Bearing)*	R=6, C=3, FM=n.a.	Vibration	n.a.								
	R=4, C=3, FM=n.a.	Temperature	Absolute energy			0.95				0.98	
			Area under the curve			0.96				0.98	
			Autocorrelation			0.95				0.98	
			Average power			0.95				0.98	
			Max			0.94				0.98	
			Maximum frequency								0.96
			Mean			0.96				0.98	
			Median			0.95				0.98	
			Min			0.92				0.97	
			Root mean square			0.96				0.98	
			Spectral roll-off								0.96
NASA - Bearing Dataset[6] *(Mechanical component, Bearing)*	R=4, C=1, FM=3	Vibration	n.a.								
GitHub - XJTU-SY Bearing Datasets[7] *(Mechanical component, Bearing)*	R=15, C=3, FM=n.a.	Vibration (vertical)	Absolute energy					0.92			
			Autocorrelation					0.92			
			Average power					0.92			
			Max					0.92			
			Mean absolute diff					0.94			
			Median absolute diff					0.95			
			Min						0.92		
			Peak to peak distance					0.92			
			Signal distance					0.95			

Continued on next page

Table 4 – continued from previous page

Dataset *(Domain, Application)*	**Runs, Configs, Failure Modes**	**Signal**	**Feature**	Curve progression li	ld	si	sd	pi	pd	di	dd
			Sum absolute diff					0.94			
			Variance					0.92			
		Vibration (horizontal)	Interquartile range					0.9			
			Mean absolute deviation					0.91			
			Mean absolute diff					0.93			
			Median absolute deviation					0.9			
			Median absolute diff					0.93			
			Root mean square					0.91			
			Signal distance					0.92			
			Standard deviation					0.91			
			Sum absolute diff					0.93			
Mendeley - Run-to-Failure Vibration Dataset of Self-Aligning Double-Row Ball[8] *(Mechanical component, Bearing)*	R=6, C=4, FM=n.a.	Vibration	n.a.								
Zenodo - Ball bearings subjected to time-varying load and speed conditions[9] *(Mechanical component, Bearing)*	R=17, C=11, FM=n.a.	Vibration	n.a.								
		Temperature	Maximum frequency								0.91
			Spectral roll-off								0.91
			Variance								0.95
PHM IEEE Data Challenge 2014 - Fuel Cell[10] *(Electrical component, Fuel cell)*	R=5, C=1, FM=n.a.	Voltage	Absolute energy		0.95						0.98
			Area under the curve		0.95						0.98
			Autocorrelation		0.95						0.98
			Average power		0.95						0.98
			Mean		0.95						0.98
			Median		0.95						0.99
			Root mean square		0.95						0.98
			Spectral distance							0.93	

Continued on next page

Table 4 – continued from previous page

Dataset *(Domain, Application)*	**Runs, Configs, Failure Modes**	**Signal**	**Feature**	**Curve progression** li	ld	si	sd	pi	pd	di	dd
NASA - Randomized Battery Usage Dataset[11] *(Electrical component, Battery)*	R=25, C=7, FM=n.a.	Voltage (discharge)	Absolute energy		0.97						0.98
			Area under the curve		0.97						0.98
			Autocorrelation		0.97						0.98
			Centroid		0.97						0.98
			Entropy							0.97	
			Fundamental frequency	0.97				0.95			
			Max power spectrum		0.93						0.98
			Mean		0.9						
			Mean absolute diff	0.96							
			Mean diff		0.96				0.96		
			Median absolute deviation							0.92	
			Root mean square		0.9						
			Signal distance		0.97						0.98
			Slope		0.98						0.91
			Spectral distance	0.95						0.99	
			Spectral slope		0.97				0.95		
		Temperature	Centroid		0.97						0.97
			Signal distance		0.97						0.99
			Spectral distance	0.93						0.97	
			Spectral slope		0.96				0.96		
NASA - HIRF Battery[12] *(Electrical component, Battery)*	R=44, C=varying, FM=n.a.	Voltage (discharge)	Max		0.92						0.94
			Spectral distance							0.9	
		Temperature	Absolute energy	0.98						0.99	
			Area under the curve	0.98						0.99	
			Autocorrelation	0.98						0.99	
			Average power	0.98						0.99	
			Max	0.96						0.98	
			Mean	0.98						0.99	
			Median	0.97						0.99	
			Min							0.91	
			Root mean square	0.98						0.99	
NASA - Li-ion Battery Aging Datasets[13] *(Electrical component, Battery)*	R=34, C=12, FM=n.a.	Voltage (discharge)	n.a.								
		Temperature	n.a.								

Continued on next page

Table 4 – continued from previous page

Dataset *(Domain, Application)*	**Runs, Configs, Failure Modes**	**Signal**	**Feature**	**Curve progression**							
				li	ld	si	sd	pi	pd	di	dd
Mendeley - Battery Degradation Dataset (Fixed Current Profiles and Arbitrary Uses Profiles)[14] *(Electrical Component, Battery)*	R=73, C=multiple, FM=n.a.	Voltage (discharge)	Absolute energy				0.93				
			Area under the curve				0.92				
			Autocorrelation				0.93				
			Centroid				0.91				
			Neighbourhood peaks				0.95				
			Signal distance				0.91				
			Spectral distance			0.99					
			Spectral positive turning points				0.91				
		Temperature	Centroid				0.91				
			Spectral distance			0.99					
Oxford Battery Degradation Dataset[15] *(Electrical Component, Battery)*	R=8, C=1, FM=n.a.	Voltage (discharge)	Absolute energy		0.98						1.0
			Area under the curve		0.98						1.0
			Autocorrelation		0.98						1.0
			Centroid		0.98						1.0
			Fundamental frequency	0.99				0.98		1.0	
			Mean absolute diff	0.99						1.0	
			Mean diff		0.99						1.0
			Negative turning points		0.91						0.91
			Positive turning points		0.91				0.9		0.91
			Signal distance		0.98						1.0
			Skewness		0.94				0.91		
			Slope		0.98				0.98		
			Spectral centroid							0.96	
			Spectral distance	0.97						1.0	
			Spectral entropy			0.91				0.96	
			Spectral kurtosis								0.93
			Spectral positive turning points		0.93						0.95
			Spectral skew-ness								0.95
			Spectral slope		0.99						1.0
			Sum absolute diff		0.9						0.91
		Temperature	Absolute energy		0.97						0.99
			Area under the curve		0.98						0.99
			Autocorrelation		0.97						0.99
			Centroid		0.98						1.0

Continued on next page

Table 4 – continued from previous page

Dataset *(Domain, Application)*	**Runs, Configs, Failure Modes**	**Signal**	**Feature**	**Curve progression**							
				li	ld	si	sd	pi	pd	di	dd
			Fundamental frequency	0.99				0.98		1.0	
			Max power spectrum		0.97						0.98
			Negative turning points		0.93						0.95
			Positive turning points		0.93						0.95
			Signal distance		0.98						1.0
			Spectral distance	0.97						1.0	
			Spectral positive turning points		0.94						0.96
			Spectral slope		0.99						1.0
			Sum absolute diff		0.92						0.93
Zenodo - Data-driven capacity estimation of commercial lithium-ion batteries from voltage relaxation[16] *(Electrical component, Battery)*	R=130, C=11, FM=n.a.	Voltage (discharge)	n.a.								
PHM Data Challenge 2010 - CNC milling machine cutters[17] *(Manufacturing process, Milling)*	R=6, C=1, FM=1	Vibration (workpiece)	Absolute energy					0.96			
			Area under the curve					0.97			
			Autocorrelation					0.96			
			Average power			0.9		0.96			
			Entropy	0.91				0.95			
			Interquartile range					0.96			
			Max					0.91			
			Mean					0.97			
			Mean absolute deviation					0.96			
			Mean absolute diff					0.94			
			Median					0.97			
			Median absolute deviation					0.96			
			Median absolute diff					0.95			
			Peak to peak distance					0.91			
			Root mean square					0.97			
			Standard deviation					0.96			

Continued on next page

Table 4 – continued from previous page

Dataset *(Domain, Application)*	**Runs, Configs, Failure Modes**	**Signal**	**Feature**	**Curve progression** li	ld	si	sd	pi	pd	di	dd
			Sum absolute diff					0.94			
			Variance			0.93		0.95			
		Acoustic emission (workpiece)	Median frequency								0.98
			Zero crossing rate								0.99
NASA - Milling Dataset[18] *(Manufacturing process, Milling)*	R=15, C=8, FM=1	Vibration (spindle)	n.a.								
		Acoustic emission (spindle)	Area under the curve	0.9							
			Mean	0.9							
			Median	0.92						0.93	
			Root mean square	0.9							

[2]Eker, Camci, and Jennions (2016)
[3]Hagmeyer, Mauthe, and Zeiler (2021); Signals in the original dataset are right-censored, in this work, however, the entire signals were used.
[4]Mauthe, Bakir, Scheerer, and Zeiler (2022)
[5]Nectoux, P. and Gouriveau, R. and Medjaher, K. and Ramasso, E. and Morello, B. and Zerhouni, N. and Varnier, C. (2012)
[6]Lee, J. and Qiu, H. and Yu, G. and Lin, J. (2007)
[7](B. Wang, Lei, Li, & Li, 2020)
[8]Gabrielli, Battarra, Mucchi, and Dalpiaz (2024)
[9]Aimiyekagbon, O. K (2024)
[10]FCLAB Federation (2014)
[11]Bole, B. and Kulkarni, C. and Daigle, M. (2014)
[12]Kulkarni, C. and Hogge, E. and Quach, C. and Goebel, K. (2015)
[13]Saha, B. and Goebel, K. (2007)
[14](Lu et al., 2022)
[15](Birkl & Howey, 2017)
[16]Zhu, Jiangong (2022)
[17]X. Li et al. (2009)
[18]Agogino, A. and Goebel, K. (2007)