

Health Index Generation based on Compressed Sensing and Logistic Regression for Remaining Useful Life Prediction

Christian Knoebel¹, Daniel Strommenger², Johannes Reuter¹ and Clemens Guehmann²

¹ *Institute of System Dynamics, University of Applied Sciences Konstanz, Konstanz, Germany*
cknoebel, jreuter@htwg-konstanz.de

² *Chair of Electronic Measurement and Diagnostic Technology, Technische Universität Berlin, Berlin, Germany*
daniel.strommenger, clemens.guehmann@tu-berlin.de

ABSTRACT

To accurately depict the current health state of the system, extracting suitable features from acquired data is crucial in data driven condition monitoring applications. Generally, analogue sensor data is sampled at rates far exceeding the Nyquist-rate, resulting in considerable amounts of noise and redundancies, imposing high computational demands on to the subsequent feature processing chain (filtering, dimensionality reduction, rating). To overcome such problems, Compressed Sensing can be used to sample directly to a compressed space, provided the signal at hand and the employed compression system meet certain criteria. Theory states, that during the compression step enough information is conserved, such that a reconstruction of the original signal is possible with high probability. However, the approach proposed in this paper does not rely on reconstructed data for condition monitoring purposes, but uses directly the compressed signal representation as feature vector. It is hence assumed that enough information is conveyed by the compression for condition monitoring purposes. To fuse the compressed coefficients into one health index with reasonable range (e. g. 0,1), a logistic regression approach is employed. To demonstrate the health index generation procedure, run-to-failure data of artificially aged electromagnetic actuators is used, providing the necessary reference states for parametrisation of the logistic regression model. From the Nyquist sampled coil current measurement a time-domain signal is extracted at each monitoring time step to serve as ground truth which is compared to the Compressed Sensing based health index. As both signals are in good agreement, it can be concluded that the underlying wear-out phenomena are captured by the proposed approach founding a promising basis for remaining useful life prediction.

Christian Knoebel et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1. INTRODUCTION

Main objective of feature extraction in diagnostics and prognostics is to identify and isolate fault or degradation specific information in signals of technical applications (Jardine, Daming Lin, & Banjevic, 2006) (e. g. vibration, current, voltage, flow, ...). This usually involves A/D-conversion at appropriate sampling rates and further processing via time, time-frequency or frequency-domain methods (Lei et al., 2018). The output of these methods contains in general an already more condensed form of the desired information but also unwanted redundancies. A subsequent compression or dimensionality reduction has to be employed to get rid of those redundancies and preserve only essential information. Crucial for the whole process to be successful is feature choice and feature quality which in turn is considerably influenced by the system designer's experience and expertise. Furthermore, not only the type of methods being used to generate and select appropriate features is important, but also how those methods scale to huge amounts of sampling data and how they make use of limited resources (e. g. CPU, network, storage, ...). As in many technical applications the acquired signals contain only a small portion of useful information, a lot of data is handled and stored that will be discarded anyway. To overcome this problem, an approach termed Compressed Sensing (CS, sometimes Compressive Sampling) has been simultaneously proposed by (Candes, Romberg, & Tao, 2006) and (Donoho, 2006), which allows to perform data compression and sampling in one step, leading to noticeably lower computational costs and data traffic. Its appealing mathematical properties and its simplicity led to a fast adoption in various research fields. Especially in image compression (e. g. medical imaging), radar and object tracking interest keeps increasing. Its fielding in condition based maintenance is a novel and very promising approach, which aims to replace the classical and aforementioned feature processing (sampling, feature calculation, dimensionality reduction, rating) by just one "analogue-to-information" step consequently

reducing computational costs and complexity. Major drawbacks of already existing studies employing CS for condition monitoring are twofold: (1) features are extracted from the reconstructed signal hence just relocating computation time (Wang, Xiang, Mo, & He, 2015; Liu, Zhang, & Xu, 2017) and (2) feature extraction methods are applied directly to the CS coefficients right after compression (H. Ahmed, Wong, & Nandi, 2018). To the best of our knowledge, the approach presented in (Knoebel, Wenzl, Reuter, & Gühmann, 2017) and in the following paper, is the first attempt to directly use compressively sampled data for health index generation without further feature extraction.

The remainder of this paper is structured as follows: in II the application example is outlined followed by a brief introduction to CS. Furthermore, its usage as feature generator as well as the health index generation process based on logistic regression is covered. Results are presented and discussed in III and IV. The paper is concluded in section V where also an outlook on work in progress is given.

2. METHODS

In the following section the application example is explained and a brief introduction to CS and logistic regression is given. Furthermore, their joint usage for feature extraction is outlined.

2.1. Application

The health index generation approach is applied to run-to-failure datasets of translational electromagnetic actuators (TEA). A schematic cross section drawing of such an actuator is depicted in Fig. 1. When a constant voltage U_c is applied to the coil, magnetic flux and hence electromagnetic force builds up, resulting in the plunger to move in z (Fig. 2). Once the supply voltage is switched off, a spring retracts the plunger to its initial position. During this back and forth movement wear-out phenomena occur in the tribologi-

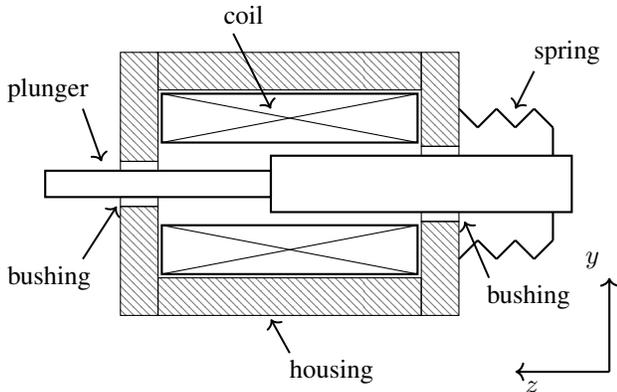


Figure 1. Cross section drawing of a translational electromagnetic actuator.

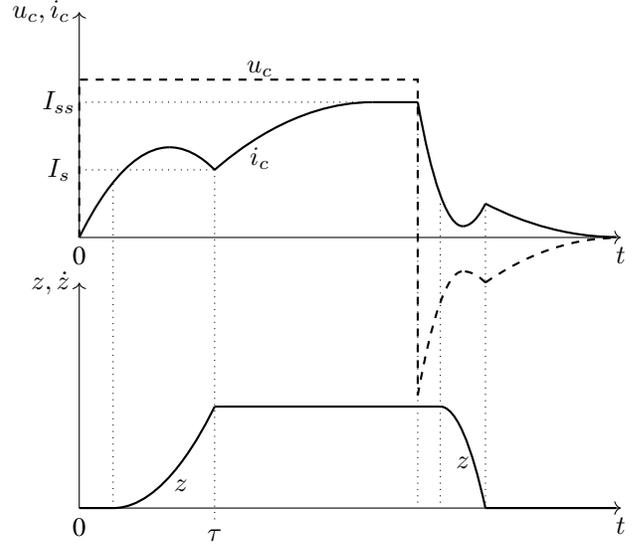


Figure 2. Dynamic actuator behaviour for the operating condition applied voltage. Depicted signals are: voltage u_c/V , coil current i_c/A and plunger displacement z/mm .

cal bushing/plunger pairing, leading to an increase in friction forces. This effect is intensified by flux leakages and manufacturing imperfections, causing a misalignment of the whole assembly and hence unit-to-unit variability. The time it takes the plunger to reach its final position is called switching time τ . It is generally used as a performance indicator by engineers and technicians, since it is directly related to an increase in friction and hence rising switching times. The exact value of τ can either be evaluated by using an oscilloscope or by a custom tailored algorithm working on Nyquist-sampled coil current measurements for finding the characteristic constriction representing the end of plunger movement. The aforementioned datasets each contain $I = 126000$ coil current and voltage measurements of $J = 10$ artificially aged units. For each unit a vector of ground truth data $\tau^j(t) = \tau_0^j, \dots, \tau_I^j$ exists, containing the determined switching time τ_i^j at time instant t_i . The actuator is said to be defective once the sequence $\tau^j(t)$ crosses a predefined threshold within its specified lifetime. During the test all units were exposed to the same environmental conditions, i.e. ambient temperature and applied voltage $U_c = 24V$.

2.2. Compressed Sensing

One fundamental assumption CS is relying on is signal sparsity. Let $x \in \mathbb{R}^n$ be a sparse vector and Φ an appropriate sensing matrix of size $\mathbb{R}^{m \times n}$ with $m < n$, then the compressed representation of x is given by the underdetermined linear system

$$y = \Phi x \text{ with } y \in \mathbb{R}^m. \quad (1)$$

Ideally, \mathbf{x} has to be sparse with at most k non-zero coefficients for CS to work (denoted as k -sparsity). The size of Φ , i.e. the number of measurements m to be taken, hereby strongly depends on sparsity and not on signal length n . However, most signals of technical applications are not directly sparse in time-domain and hence have to be represented in an appropriate orthonormal basis $\Psi \in \mathbb{R}^{n \times n}$ where sparsity can be achieved. This in turn leads to the system given in (2) with \mathbf{c} being the coefficient vector of \mathbf{x} on Ψ .

$$\mathbf{y} = \Phi \Psi \mathbf{c} = \Theta \mathbf{c} \quad (2)$$

Usually, \mathbf{c} is "only" compressible, as an ideal Ψ perfectly sparsifying \mathbf{x} is often unknown but could be learned at the expense of additional computational costs (Han, Jiang, Sun, Wang, & Yang, 2018). I.e., the support $\text{supp}(\mathbf{c}) = \{i : c_i \neq 0\}$ of k -sparse and compressible vectors is inherently different, as they contain lots of small and near-zero coefficients obeying a decreasing behaviour when sorted in descending order (with some constant values α and γ).

$$|c_i| \leq \alpha i^{\frac{1}{2}-\gamma}, \forall i \leq n \quad (3)$$

Hence, the faster $|c_i|$ is decaying, the more compressible the signal is. By thresholding near zero coefficients and keeping only the k largest ones, a compressible vector can be represented by its k -sparse approximation. This approach is often used when the signals under consideration are not compressible in standard bases, but essentially one is performing transform coding and not *Compressed Sensing*. In contrast, the output of a proper CS-system would be a compressively sampled vector \mathbf{y} which in turn is used for reconstructing the original signal of interest. For this step to be successful, not only sparsity (compressibility) is a fundamental assumption of CS, but also some key properties of the measurement system itself, following the subsequent premises: (a) the information contained in the signal is preserved and (b) it is ensured that two sparse signals $\mathbf{x}_1, \mathbf{x}_2 \in \Sigma_{2k}$ (with $\mathbf{x}_1 \neq \mathbf{x}_2$ and the set of all k -sparse signals sharing the same support being $\Sigma_k = \{\mathbf{x} : \|\mathbf{x}\|_0 \leq k\}$) don't give the same reconstruction result (Eldar & Kutyniok, 2012). In this context incoherence is one important property, measured by the mutual coherence of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$. Here a low mutual coherence is beneficial as it reduces the amount of compressive measurements to be taken. Another property guaranteeing recovery of k -sparse signals even under noisy conditions, is the Restricted-Isometry-Property (RIP) of order k . It is satisfied by \mathbf{A} if there exists a $\delta_k \in (0, 1)$ such that

$$(1 - \delta_k) \|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1 + \delta_k) \|\mathbf{x}\|_2^2, \forall \mathbf{x} \in \Sigma_k. \quad (4)$$

Generally it is very difficult to prove any of the discussed criteria for arbitrary matrices as a combinatorial search over $\binom{m}{k}$ sub-matrices would be necessary. However, random matrices with entries drawn independently from $\mathcal{N}(0, \frac{1}{m})$ are known to meet the aforementioned requirements well and are hence

employed as sampling matrix Φ in this study. Discrete cosine transform (DCT) is used as sparsifying basis Ψ , since the mutual coherence of the resulting sensing matrix $\Theta = \Phi\Psi$ is known to be small and the monitoring data (measured coil current) is compressible in the DCT basis.

2.2.1. CS in the academic context

As already mentioned, the scope of CS is reconstructing the sparse representation \mathbf{x} or \mathbf{c} of an analogue signal from its compressively sampled measurements \mathbf{y} . I.e. one would have some sampling device that is able to directly acquire the signal of interest in its compressed form \mathbf{y} (where the critical task now would be to find \mathbf{x} , or \mathbf{c} respectively). As such real world sampling devices still show poor performance or are simply not yet available (Harms, Bajwa, & Calderbank, 2013), it is common practice in academia to artificially compress the previously Nyquist sampled analogue signals using a known finite-dimensional CS system (H. O. A. Ahmed & Nandi, 2019). The same procedure is adopted for this study, allowing the investigation of the proposed health index generation approach in a more controllable setting.

2.2.2. Compressive Feature Generation

The aforementioned key requirements ensure information preservation during the compression step and hence enable signal reconstruction with high probability (Knoebel et al., 2017). In contrast to other studies employing CS for condition monitoring purposes, reconstruction is not of interest here. Instead, the condition monitoring approach presented in this paper is based on the question whether compressed signal representations directly contain the desired failure and deterioration specific information, rendering reconstruction procedures and succeeding computationally expensive feature processing methods redundant (see Fig. 3 for a schematic diagram). More precisely, the approach is based on the mathematical properties of the sampling system Θ which acts as an isometry on the data. Key enabler is the strong relationship between RIP and the Johnson-Lindenstrauss-Lemma (Baraniuk, Davenport, DeVore, & Wakin, 2008) which basically states that a linear embedding of some points of interest in a lower dimensional space is approximately distance preserving if the entries of the transformation matrix are drawn

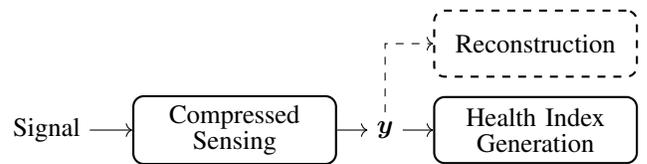


Figure 3. Flow chart of the proposed approach. Instead of using the compressively acquired information for reconstruction of the original signal, it is directly used as feature vector for health index generation.

from appropriate distributions. This holds for the mapping given in Eq. (2), indicating that failure and wear-out specific signal characteristics survive the compression step.

2.2.3. Reconstruction Algorithms

Signal reconstruction can e.g. be conducted via greedy search or combinatorial algorithms. For verifying the suitability of the selected measurement system, reconstruction is initially performed for some compressed measurements using ℓ_1 -minimisation (also known as basis pursuit)

$$\min_x \|x\|_1 \text{ s. t. } y = \Phi x. \quad (5)$$

For further information on algorithms and methods, the reader is referred to (Eldar & Kutyniok, 2012).

2.3. Health Index Generation

One major problem in data-based condition monitoring is the identification of an appropriate health index relating the current degradation state to a numeric value between 0 and 1 (degraded, healthy). Ideally, the health index should be monotonically increasing/decreasing so that it can be depicted by linear or nonlinear (quadratic or exponential) degradation models. Furthermore, physical interpretability is desirable as the health state can directly be related to the underlying wear-out phenomena. However, as the approach presented in this paper uses compressively sampled signal representations as features, the identification of a suitable health index is not straight forward. The underlying assumption is, that in time-domain the measured current profiles (see Fig. 2) of defective units are very similar – if not identical – and that these characteristics map to the compressed space making a distinction between new and deteriorated systems possible. To illustrate the assumption that compressed data conveys the necessary information to perform condition-monitoring, Fig. 4 shows the percentage change between two instances in component lifetime for each of the $n = 250$ CS coefficients. It can be observed that advancing deterioration leads to deviations from a baseline up to 500 to 1000 %, whereas a completely new system shows standard deviations below 10 % for the majority of CS coefficients. This (relative) change in coefficient magnitude can be used to capture the aforementioned failure and wear specific characteristics. To fuse all CS coefficients into one numeric value with reasonable range, a Logistic Regression (*LogReg*) approach – which has already been applied to similar problems (Spezzaferro, 1996; Caesarendra, Widodo, & Yang, 2010) – is used.

$$P(\mathbf{y}) = \frac{1}{1 + e^{-g(\mathbf{y})}} \quad (6)$$

LogReg relates a set of independent variables to a dichotomous dependent variable where the output has two states (1,0). It is based on the general logistic model where (6)

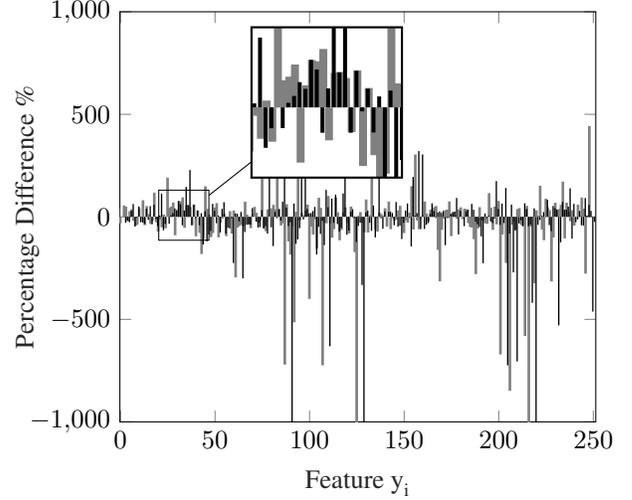


Figure 4. Percentage change in CS coefficients: new/ageing (black) and new/defective system (gray). It can be observed from the bar plot that all coefficients y_i undergo some change during the component lifetime.

can be interpreted as the probability of the dependent variable equalling 1. In the present context, the independent variables correspond to the CS coefficients \mathbf{y} , and the dependent variable represents the probability $P(\mathbf{y})$ of the system under consideration being in a healthy state (i.e. the two states defective or healthy). Employing the logit-transform, which is denoted as the logarithm of the odds (7), the nonlinear problem of (6) is transformed to the linear one given in (8), where m is the size of the compressed vector \mathbf{y} .

$$O(h) := \frac{P(\mathbf{y})}{1 - P(\mathbf{y})} \quad (7)$$

$$\text{Logit}(h) = \ln(O(h)) = g(\mathbf{y}) = \beta_0 + \sum_{i=1}^m y_i \beta_i \quad (8)$$

By plugging $g(\mathbf{y}) = \beta_0 + \sum_{i=1}^m y_i \beta_i$ from (8) in (6), the desired health index can be calculated for each sample \mathbf{y} . Using an Expectation Maximization (*EM*) approach the regression weights β_0 and β_i of (8) are identified based on reference data. To account for unit-to-unit variability, a *LogReg* model for each component j is built based on the individual as-new and a generalized (fleet specific) worn out condition obtained from historical failure data \mathbf{y}_f . Once the model is fully parametrised, the current health index at monitoring instant t_i can be calculated for each component. As already described, an actuator is termed to be defective when the series $\tau^j(t)$ permanently crosses a predefined threshold value within its specified lifetime ($T_{tr} = 80\text{ms}$, $N = 2 \cdot 10^6$ working cycles). This condition is met by the devices 1, 2, 4 and 10 at different time instances t_i and working cycles n_i respectively, which leads to the following reference set $\mathbf{Y}_{aged} = [\mathbf{y}(\tau_i^1 > T_{tr}), \mathbf{y}(\tau_i^2 > T_{tr}), \mathbf{y}(\tau_i^4 > T_{tr}), \mathbf{y}(\tau_i^{10} > T_{tr})]$.

As the devices 3, 5, 6, 7, 8 and 9 reached their specified lifetime of $N = 2 \cdot 10^6$ cycles without crossing the threshold T_{tr} , they can be used to validate the HI generation process. For generating the necessary reference datasets the following requirements can be summarised:

- The initial state of component j in compressed feature space is defined by the first k measurements after initial operation $\mathbf{Y}_{new}^j = \mathbf{y}(\tau_{0\dots k}^j | \tau_{0\dots k}^j < T_{tr})$.
- The worn-out condition is assumed to be described in compressed feature space by a multivariate normal distribution $\mathbf{Y}'_{aged} \sim \mathcal{N}_p(\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_a)$ which is parametrised based on the fleet specific reference \mathbf{Y}_{aged} .
- $\mathbf{Y}_{new}^j, \mathbf{Y}_{aged} \in \mathbb{R}^{m' \times m}$ with $m' \geq m$.
- The training samples $\tilde{\mathbf{y}}_{i,new}^j$ and $\tilde{\mathbf{y}}_{i,aged}$ are drawn from \mathbf{Y}_{new}^j and \mathbf{Y}'_{aged} and are labelled with
 - 0.95 corresponding to 95 % health
 - and 0.05 corresponding to 5 % health.

The model fit is verified for each test object by checking the standardized Pearson residuals

$$r_{Ps_i} = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\phi}V(\hat{\mu}_i)(1 - h_i)}} \quad (9)$$

for normality with estimate value $\hat{\mu}_i$, measurement y_i and variance function $V(\cdot)$ with $\hat{\phi} = 1$ for binomial distributions. h_i is a distance measure for the respective data point towards the centroid of its data space. The pseudo-code for the whole health index generation process is given in Alg. 1. Please note, that the condition "Component is new" relates to the break-in phase described in the following section.

Algorithm 1: Health Index Generation

Data: Condition-Monitoring-Data $\mathbf{x}_{t_i}^j$.

Result: Health index $h_{t_i}^j$ from CS coefficients $\mathbf{y}_{t_i}^j$.

Initialize CS System $\Theta = \Phi\Psi$;

Load historical failure data \mathbf{Y}_{aged} ;

foreach Component j **do**

while new Monitoring-Data at t_i is available **do**

 Compress Signal $\mathbf{y}_{t_i}^j \leftarrow \Theta \mathbf{x}_{t_i}^j$;

if Component is new **then**

 Parametrise *LogReg* model:

$\beta_0^j, \beta_i^j \leftarrow EM(\mathbf{Y}_{new}^j, \mathbf{Y}_{aged})$;

else

 Calculate HI:

$h_{t_i}^j \leftarrow \text{LogReg}(\mathbf{y}_{t_i}^j)$;

end

 Estimate RUL of component j based on $h_{t_i}^j$;

end

end

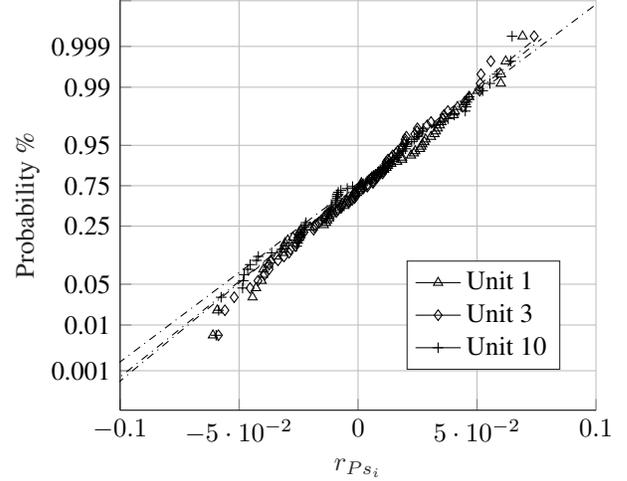


Figure 5. Normal Probability plot of the normalized Pearson residuals for the logistic models of units 1, 3 and 10.

3. RESULTS

For visualizing the proposed approach, datasets of three actuators are exemplarily selected. Hereby, two actuators (unit 1 and 10) did not reach their specified lifetime, whereas unit 3 is in a health state for the considered time span. Figure 5 shows the normal probability plot of the Pearson residuals of the logistic regression models of units 1, 10 (defective) and 3 (healthy). All three datasets show a slight tendency towards a skewed distribution on both ends of the data range, but match the normality assumption quite well indicating a good fit of the models. Figures 6, 7 and 8 show the *LogReg* based HIs (ranging from 1 to 0, i.e. 100% to 0% health) in comparison with the respective time series ground truth signals τ plotted over the whole $2 \cdot 10^6$ working cycles. The corresponding threshold value of $T_{tr} = 80ms$ is given as a dashed line for reference. In time-domain, a typical break-in behaviour can be observed for units 3 and 10 (depicted in boldface) which is characterised by an initial increase in switching time (up to a local maximum $\hat{\tau}$) with a succeeding drop to a local minimum. Despite unit 1 not showing the aforementioned behaviour, all calculated HIs are in a steady-state from start indicating normal/healthy operation. Once the time-domain signal starts to break away ($\tau(t_i) > \hat{\tau}$), the calculated HIs start to decrease as well. The ground truth signals of units 1 and 10 permanently cross the threshold value at $\approx 1.74 \cdot 10^6$ and $\approx 0.68 \cdot 10^6$ cycles indicating a failure. The corresponding HI of unit 10 reaches its minimum at $\approx 0.68 \cdot 10^6$ cycles as well, whereas the HI of unit 1 shows quite different behaviour and reaches a quasi steady-state at around $N \approx 0.9 \cdot 10^6$ cycles. Unit 3 behaves as expected and shows a reasonable agreement between HI and ground truth drift.

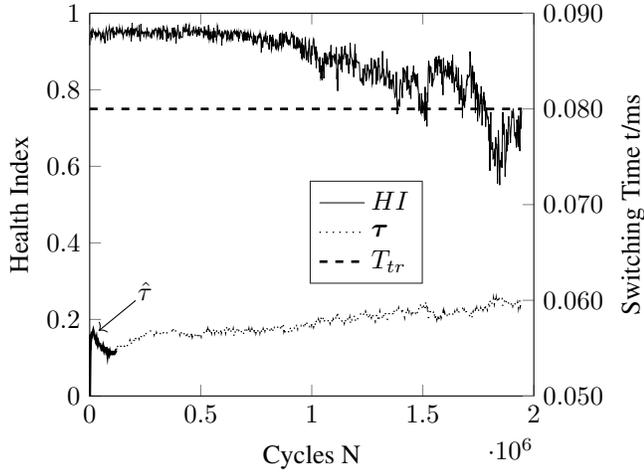


Figure 6. Health index HI , switching time τ and threshold T_{tr} of test unit 3. The break-in phase is depicted in boldface over the cycle range of 0 to $\approx 0.121 \cdot 10^6$.

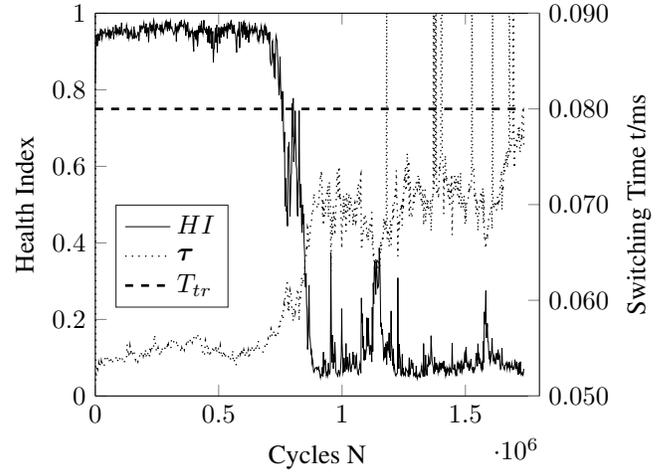


Figure 8. Health index HI , switching time τ and threshold T_{tr} of test unit 1.

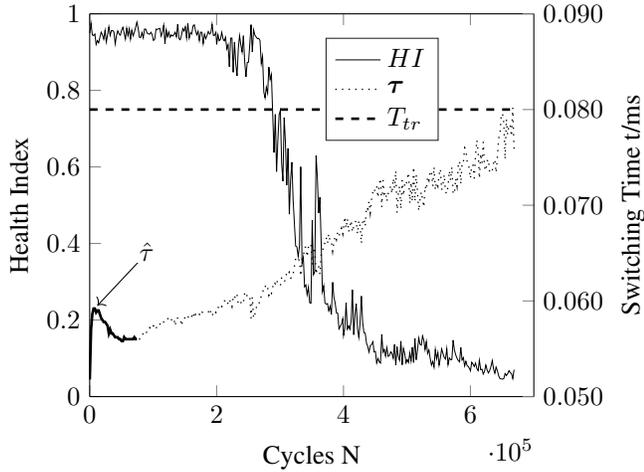


Figure 7. Health index HI , switching time τ and threshold T_{tr} of test unit 10. The break-in phase is depicted in boldface over the cycle range of 0 to $\approx 0.75 \cdot 10^5$.

4. DISCUSSION

By inspecting the results given in Fig. 7 and 8, it can be observed that the HI does not cover a range of (1,0) but instead one of approximately (0.95,0.05) even when the ground truth threshold is reached and the HI should be close to 0. However, this behaviour is expected, as values on both ends of the range are predetermined by the chosen training labels and thus limited to those values due to the mathematical properties of the logistic function/model (see assumption (d) of sec. 2.3). These characteristics also result in the HI showing steady-state behaviour during and after the break-in phase, although the ground truth values undergo some fluctuations. The missing break-in phase of unit 1 (Fig. 8) is most likely due to a properly aligned assembly which results in lower

friction forces, faster switching times and hence a different training dataset Y_{new}^1 compared to units 3 and 10. Generally, choosing the individual as-new condition Y_{new}^j is critical for the overall system performance, since the amount of fluctuation "dampening" during the break-in phase strongly depends on how long the unit is considered to be new. By labelling the break-in phase as initial/new system behaviour, a compromise between allowing a certain amount of drift and providing enough sensitivity to degradation can be achieved. This sensitivity of the HI generation approach towards changes in the system's health is important in two ways: (1) no time delay between wear-out phenomena and the HI is desirable and (2) healing effects – which can occur in real systems – have to be depicted. Furthermore, the applied methods (especially CS as key enabler) show good robustness against noise due to their mathematical properties. When comparing the generated HI of unit 1 given in Fig. 8 with its ground truth signal, this advantage is directly visible. The very prominent outliers of $\tau(t_i)$ of unit 1 in the range of $\approx 1.5 \cdot 10^6$ cycles are caused by noise failing the custom tailored algorithm to reliably find τ in the time-domain signal. To compensate for these algorithmically induced problems, a succeeding outlier treatment would be necessary imposing additional computational loads, counteracting the savings made at preceding processing steps. To evaluate the validity of the generated HIs, the normalized cross-correlation and correlation-coefficients of each ground-truth/HI pair were analysed, indicating a good agreement of both signals (see Tab. 1). As the Nyquist sampled time-domain information would not be available in a CS-based system (unless the original signal is reconstructed), the ability to qualitatively reproduce the ground truth signal with the HI is of central interest to meet the requirements outlined in the introduction to sec. 2.3. However, an in-depth evaluation can only be realised by using both, the HI and ground

truth signals as inputs for prognostic algorithms, where especially the performance regarding unit 1 is of interest.

Table 1. Cross-correlation peak values and correlation-coefficients between τ and HI.

| – | Cross-correlation | Correlation-coefficient |
|------|------------------------------|-------------------------|
| Unit | $\rho_{\tau,HI}^j(\tau = 0)$ | $\kappa_{\tau,HI}^j$ |
| 3 | 0.64 | 0.74 |
| 10 | 0.79 | 0.89 |
| 1 | 0.81 | 0.95 |

5. CONCLUSION AND FUTURE WORK

In this paper an approach towards health index generation from compressively sampled data is proposed. By directly using the compressed representation of an analogue signal as features, a dedicated and computationally expensive feature extraction and selection procedures is rendered redundant. Using a logistic regression approach, the current condition of the system can be fused into a health index with reasonable range, allowing a comparison with the time-domain ground truth. As both signals are in good agreement, it can be concluded that the underlying wear-out phenomena can be captured by the constructed health index. Especially when considering applications with limited resources, this approach offers great potential by reducing the necessary amount of storage space, computational power and time. However, defining reasonable failure states in the compressive feature space – without the need for run-to-failure tests – is still problematic and under research. Due to the stochastic nature of the generated health index, current and future work focuses on implementing a Wiener process based remaining useful life prediction approach with concurrent state and parameter estimation (Si, Hu, Chen, & Wang, 2011). Applied to the same run-to-failure experiments as used in this study, prediction results are encouraging and reasonable.

REFERENCES

- Ahmed, H., Wong, M., & Nandi, A. K. (2018). Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features. *Mech Sys Signal Process*, 99, 459–477.
- Ahmed, H. O. A., & Nandi, A. K. (2019). Intelligent Condition Monitoring for Rotating Machinery Using Compressively-Sampled Data and Sub-space Learning Techniques. In *Proc of the 10th Int Conf on Rotor Dynamics (IFTOMM)* (Vol. 61, pp. 238–251). Springer.
- Baraniuk, R. G., Davenport, M. A., DeVore, R., & Wakin, M. B. (2008). A Simple Proof of the Restricted Isometry Property for Random Matrices. *Constructive Approximation*, 28(3), 253–263.
- Caesarendra, W., Widodo, A., & Yang, B.-S. (2010). Application of relevance vector machine and logistic regression for machine degradation assessment. *Mech Sys Signal Process*, 24(4), 1161–1171.
- Candes, E. J., Romberg, J. K., & Tao, T. (2006). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans Inform Theory*, 52(2), 489–509.
- Donoho, D. L. (2006). Compressed sensing. *IEEE Trans Inform Theory*, 52(4), 1289–1306.
- Eldar, Y. C., & Kutyniok, G. (2012). *Compressed Sensing: Theory and Applications*. Cambridge University Press.
- Han, T., Jiang, D., Sun, Y., Wang, N., & Yang, Y. (2018). Intelligent fault diagnosis method for rotating machinery via dictionary learning and sparse representation-based classification. *Measurement*, 118, 181–193.
- Harms, A., Bajwa, W. U., & Calderbank, R. (2013). A Constrained Random Demodulator for Sub-Nyquist Sampling. *IEEE Trans Signal Process*, 61(3), 707–723.
- Jardine, A., Daming Lin, & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech Sys Signal Process*, 20(7), 1483–1510.
- Knoebel, C., Wenzl, H., Reuter, J., & Gühmann, C. (2017). A Compressed Sensing Feature Extraction Approach for Diagnostics and Prognostics in Electromagnetic Solenoids. In *Proc Annu Conf Prognostics and Health Management Society (PHM2017)*.
- 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. *Mech Sys Signal Process*, 104, 799–834.
- Liu, Y., Zhang, G., & Xu, B. (2017). Compressive sparse principal component analysis for process supervisory monitoring and fault detection. *J Process Control*, 50, 1–10.
- Si, X.-S., Hu, C.-H., Chen, M.-Y., & Wang, W. (2011). An adaptive and nonlinear drift-based Wiener process for remaining useful life estimation. In *Progn and System Health Management Conf (PHM-Shenzhen), 2011* (pp. 1–5). IEEE.
- Spezzaferro, K. E. (1996). Applying logistic regression to maintenance data to establish inspection intervals. In *Ann Reliability and Maintainability Symp* (pp. 296–300). IEEE.
- Wang, Y., Xiang, J., Mo, Q., & He, S. (2015). Compressed sparse time–frequency feature representation via compressive sensing and its applications in fault diagnosis. *Measurement*, 68, 70–81.