

Received 8 August 2022, accepted 22 September 2022, date of publication 3 October 2022, date of current version 14 October 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3211258

## RESEARCH ARTICLE

# A Hybrid Model-Based Approach on Prognostics for Railway HVAC

ANTONIO GÁLVEZ<sup>1,2</sup>, (Student Member, IEEE), DIEGO GALAR<sup>1,2</sup>,  
AND DAMMIKA SENEVIRATNE<sup>1</sup>

<sup>1</sup>Tecnalia, Basque Research and Technology Alliance (BRTA), 48170 Derio, Spain

<sup>2</sup>Division of Operation and Maintenance Engineering, Department of Civil, Environmental and Natural Resources Engineering, Luleå University of Technology, 971 87 Luleå, Sweden

Corresponding author: Antonio Gálvez (antonio.galvez@associated.ltu.se)

**ABSTRACT** Prognostics and health management (PHM) of systems usually depends on appropriate prior knowledge and sufficient condition monitoring (CM) data on critical components' degradation process to appropriately estimate the remaining useful life (RUL). A failure of complex or critical systems such as heating, ventilation, and air conditioning (HVAC) systems installed in a passenger train carriage may adversely affect people or the environment. Critical systems must meet restrictive regulations and standards, and this usually results in an early replacement of components. Therefore, the CM datasets lack data on advanced stages of degradation, and this has a significant impact on developing robust diagnostics and prognostics processes; therefore, it is difficult to find PHM implemented in HVAC systems. This paper proposes a methodology for implementing a hybrid model-based approach (HyMA) to overcome the limited representativeness of the training dataset for developing a prognostic model. The proposed methodology is evaluated building an HyMA which fuses information from a physics-based model with a deep learning algorithm to implement a prognostics process for a complex and critical system. The physics-based model of the HVAC system is used to generate run-to-failure data. This model is built and validated using information and data on the real asset; the failures are modelled according to expert knowledge and an experimental test to evaluate the behaviour of the HVAC system while working, with the air filter at different levels of degradation. In addition to using the sensors located in the real system, we model virtual sensors to observe parameters related to system components' health. The run-to-failure datasets generated are normalized and directly used as inputs to a deep convolutional neural network (CNN) for RUL estimation. The effectiveness of the proposed methodology and approach is evaluated on datasets containing the air filter's run-to-failure data. The experimental results show remarkable accuracy in the RUL estimation, thereby suggesting the proposed HyMA and methodology offer a promising approach for PHM.

**INDEX TERMS** Prognostics and health management, hybrid modelling, deep learning, HVAC system, railway.

## I. INTRODUCTION

The maintenance domain has evolved through a series of industrial revolutions. The 4th Industrial Revolution is based on data collection and analysis, not only for specific maintenance practices but also for more general objectives, such as zero-defect manufacturing and services, key drivers of performance. Companies are increasingly trying to control the global process from the supplier to the final customer. The goal is to use performance evaluation, production and

process deviation prediction, and decision support to identify defects and their causes and react before failure. When companies understand their processes, they will be able to reduce downtime and maximize production [1].

Wang argued [2] that key factors in reaching zero-defect product quality include monitoring the health state of facilities and equipment and optimizing decision-making with huge datasets. This is especially important in transport, energy, and chemistry sectors where safety is more important than reliability or efficiency; such industries monitor the health state of critical components to optimize decision-making under required conditions. The monitoring data are

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaojun Steven Li<sup>1</sup>.

used to continuously improve maintenance planning. The maintenance strategy that relies on monitoring the condition of an asset in real-time to determine a maintenance action is condition-based maintenance (CBM). The standard EN 13306:2017 describes this as a maintenance strategy that allows companies to extend the life cycle of assets while ensuring assets' behaviour and function under required conditions of safety, reliability, and effectiveness [3].

CBM predicts latent faults in advance and dynamically changes maintenance plans based on prognostics. However, faults or abnormal equipment behaviours can suddenly appear; these are detected in diagnostics processes, which together with prognostics, show the current state of a system [4], [5]. Thus, CBM includes both diagnostics and prognostics.

Diagnostics is used when a failure or an unusual behaviour is detected. Failure modes and effect analysis (FMEA) is used to trace the relationship between a failure and the data acquired from the system [6]. Diagnostics detects, isolates, and localizes a faulty component based on a failure model (FM) [7], [8]. Diagnostics for heating, ventilation, and air conditioning (HVAC) systems, the system of interest in this paper, is developed in [8], [9], [10], [11], and [12].

Prognostics is performed by assessing changes in the behaviour of components or systems over time to predict their remaining useful life (RUL) and end of life (EoL). The information obtained in diagnostics is considered in prognostics, as the accumulated degradation is evaluated to estimate the RUL and predict the future health state. Researchers have developed algorithms to predict RUL for different applications [13], [14], [15], [16], [17].

CBM is used in prognostics and health management (PHM), an engineering discipline which studies the health state of equipment and predicts its future evolution with the integration of aspects such as logistics, security, reliability, mission criticality and cost-effectiveness; thus, PHM goes beyond CBM [18], [19].

There are three main methods for implementing CBM [20]: physical model-based approaches [21], data-driven approaches or surrogate models [22], and hybrid model-based approaches (HyMA) [23].

Physical model-based approaches are based on mathematical models of the physical system; if the system degradation is accurately modelled, these approaches tend to be more effective than other approaches [20]. The models incorporate such characteristics as material properties, thermodynamics, and mechanical responses, thus requiring extensive prior knowledge of physical systems. Yet a detailed model is usually difficult to develop because some key parameters are unavailable in practice, especially in complex systems or processes [24], [25]. Model-based solutions have been developed by several researchers [26], [27], [28], [29]. Nevertheless, detailed physics-based models on deploying diagnostics for HVAC system are not being used, however they are being used for covering fault modeling in prognostics of HVAC systems [63].

Data-driven approaches do not require expertise to model system degradation; they are built using mathematical models and weight parameters and trained using historical data collected by sensors installed in the physical system. Therefore, data-driven approaches are more practical and agile than model-based approaches for deploying CBM in complex systems or processes. However, because they only depend on historical or online data and do not consider system complexity, they miss the relations between the data and the physical world [30], [31]. Data-driven approaches can be divided into two categories [20]. The first includes artificial intelligence (AI) approaches: neural networks (NNs) and fuzzy logic. The second includes statistical approaches; common techniques are support vector machines (SVMs), linear regression, hidden Markov model, and Gaussian process regression. Data-driven approaches for CBM in HVAC systems are discussed by [10], [32], [33], and [34]. These researchers obtained a remarkable results in their approaches and it is well known that data-driven models are easy to be developed on a cost-effective way; nevertheless, the robustness of these approaches for HVAC systems is sometimes questionable due to lack of data available while operating in faulty state.

Hybrid model-based approaches combine data-driven and physical model-based approaches. This combination improves diagnostics and prognostics by overcoming the lack of historical data, thus improving the ability to detect failure modes (FMs) and reducing the appearance of hidden FMs, metaphorically known as "black swans" [35]. It can be expensive, difficult, or even impossible to install sensors in parts of a system that could be of interest for CBM. In these cases, soft sensors, also known as virtual sensors, can be defined in physics-based models. Soft sensors are modelled to generate additional information to improve fault detection and RUL estimation of monitored systems [36]. As a consequence, physics-based models can be used to generate synthetic data related to those situations or parts for which it is difficult to obtain data and system degradation in the required timeframe; this results in complete and large datasets that allow predictive maintenance through data-driven models. Moreover, emerging deep learning approaches have been successfully applied in the field of big data; these include convolutional neural networks (CNNs) [37], deep belief networks (DBNs) [38], recurrent neural networks (RNNs) [39], and long short term memory (LSTM) networks [40]. Zhang et al. proposed a novel bidirectional gated recurrent unit with a temporal self-attention mechanism to predict RUL; specifically, each considered time instance is assigned a self-learned weight according to the degree of significance [41]. Liu et al. presented a prediction model called an improved multi-stage long short term memory network with clustering; it combines the advantages of clustering analysis and the LSTM model [42]. Caceres et al. proposed a probabilistic Bayesian recurrent neural network (RNN) for RUL prognostics considering epistemic and aleatory uncertainties [43]. Deep learning has been used for different subsystem of HVAC system in PHM. Guo et al. [64] performed a deep

learning-based fault diagnostics of variable refrigerant flow system. Sun et al. [65] used deep learning techniques for developing a gradual fault diagnostics approach for air source heat pump system.

This paper proposes a hybrid-model approach which combines a physics-based model and a data-driven model. The physics-based model is used to generate run-to-failure data; these synthetic data are combined with real data and used to train, validate, and test a deep convolution neural network. The architecture used in the hybrid-model approach was first presented in an excellent article [37] and obtained higher prognostic accuracy than other traditional machine learning methods. We validate the proposed HyMA using run-to-failure data generated by the physics-based model; therefore, intrusive experiments are not performed in this study.

Some of the most remarkable recent CBM advancements have been for HVAC systems. Yet it is difficult to find research where RUL estimation models are developed for HVAC systems installed in high-speed passenger train carriages, even though a failure in this system affects people's safety and could affect the environment. This paper begins to fill the gap in the research.

The remainder of the paper proceeds as follows. Section 2 describes the methodology proposed for fusing physics-based and data-driven models. It explains the HVAC system, the physics-based model, the modelled failure, and the architecture of the deep learning model. Section 3 describes the experimental study, including the generation of data, the preparation of the dataset to be input for deep CNN, and the parameters for implementing the CNN model. Section 4 discusses the results. Section 5 closes the paper with conclusions and suggestions for future research.

## II. PROPOSED HYBRID MODEL APPROACH

Railway engineering systems have strict regulations for reliability, availability, maintainability, and safety (RAMS) during their life cycle, as specified in the standard EN 50126-1, 2017 [44]. Consequently, the lifetime of critical components is not maximized because maintainers usually replace them in early degradation stages for safety, environmental, and economic reasons. Only those with a low criticality are allowed to operate until failure. Data cannot be acquired by sensors in faulty stages of most components, and this complicates the acquisition of run-to-failure data. Thus, a combination of physical model-based and data-driven approaches is required. The hybrid model can overcome the lack of data to improve the detectability of failure modes, reduce the hidden failure modes, i.e., "black swan losses", and assess their effects within the timeframe.

We propose combining a physics-based model with a deep learning architecture to obtain an accurate HyMA. The HyMA is developed and simulated using MATLAB R2021b. An overview of the methodology used to combine data obtained by a physics-based model and sensors installed in the real system is presented in Figure 1.

The physics-based model is used to generate run-to-failure data. The model contains sensors installed in the real system and virtual sensors, which depend on the data gathered in the real system. The responses of the sensors defined in the model are recorded in a dataset which contains the output of these simulations. The model is simulated using real data acquired from the sensors and synthetic data acquired from the virtual sensors. The parameter of interest is the degradation of the air filter in terms of mass of dust. Every simulation contains run-to-failure data and generates timeseries data on every selected signal labelled with the RUL values. Therefore, the data related to a simulation contain timeseries data from every sensor selected. The raw signals are normalized to accelerate the training process in deep learning tasks [45]. Then, the datasets for training and testing are prepared to be inputs of the proposed network.

### A. PHYSICS-BASED MODEL OF THE HVAC SYSTEM

The general mission of an HVAC system is to maintain acceptable indoor air quality and thermal comfort through suitable ventilation with filtration while remaining within reasonable operation and maintenance costs. The HVAC system of interest (an HVAC in a passenger train carriage) was

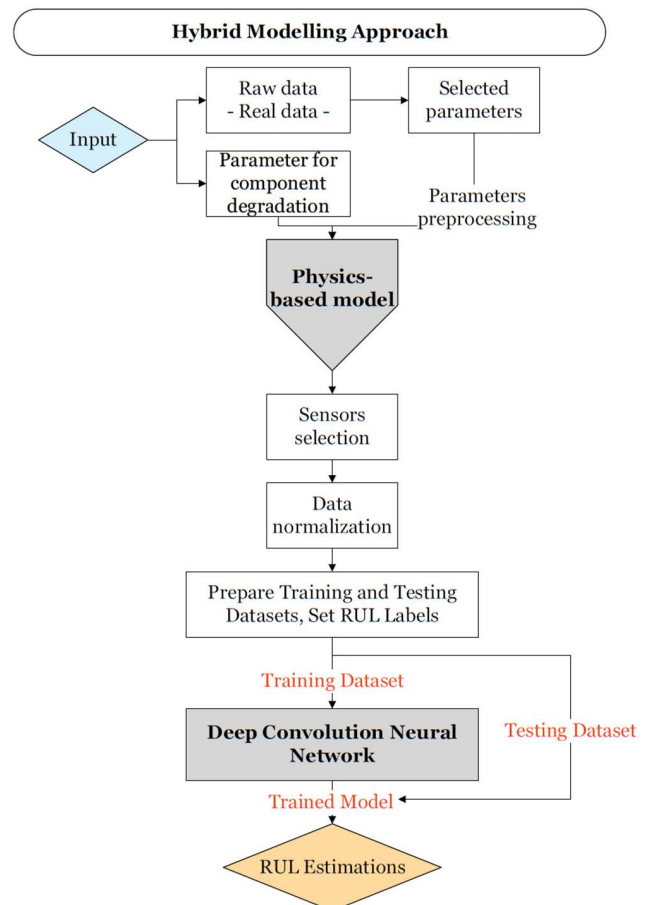


FIGURE 1. Methodology used to build the proposed HyMA.

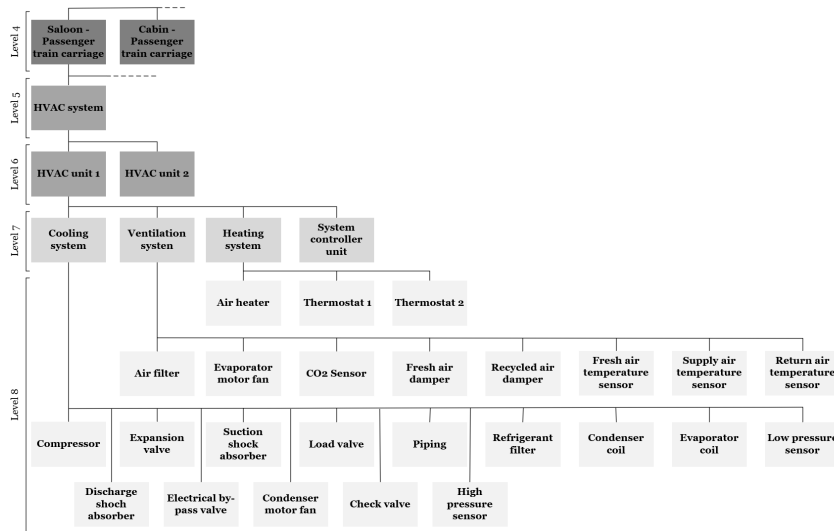


FIGURE 2. Main components considered in the studied HVAC system.

designed to satisfy the comfort conditions established in the standard EN 14750-1, 2006 [46]. Accordingly, the standard was used as an information resource to develop the physics-based model.

The HVAC installed in a passenger train carriage is separated into cooling subsystems, heating subsystems, ventilation subsystems, and vehicle thermal networking system.

Therefore, the HVAC system modelled is a system-of-systems (SoS), as systems interact with their surrounding systems to perform the required functions [47], [48].

The standard ISO 14224-2016 is widely used to define the SoS taxonomy. This research considers the HVAC system to include from taxonomy level 5, known as section/system, to taxonomy level 8, defined as component/maintenance item.

Figure 2 illustrates the taxonomy of the studied HVAC system and the most relevant elements considered in this research. The passenger train car studied is a passenger saloon with an HVAC system composed of two HVAC units. This means that almost all components are duplicated.

Level 7 contains the subsystems of the HVAC system considered here. Some components, such as contactors, circuit breakers, electronic control board and control panel, are not represented in Figure 2 because their FMs are not analyzed in the research. Level 8 includes components whose interactions are modeled based on the principles of thermodynamics, fluid mechanics and heat transfer. The physics-based model also includes the thermal network of the vehicle, in this case, and the physics of the interactions between the high-speed passenger train and the environment.

The physics-based model used to generate run-to-failure data was developed and validated in previous research [49]. It was also previously used to generate synthetic data to build a data-driven model for multiple fault detection; the model was trained, validated, and tested using real data and synthetic data. The methodology proposed to combine the two types

TABLE 1. Condition monitoring sensors.

Indices	Variable	Units
S1	Pressure after compressor 1	bar
S2	Pressure after compressor 2	bar
S3	Pressure before compressor 1	bar
S4	Pressure before compressor 2	bar
S5	CO <sub>2</sub> level	ppm
S6	Vehicle temperature	°C
S7	Impulsion temperature	°C

TABLE 2. Virtual sensors.

Indices	Variable	Units
S1v	Temperature after compressor 1	°C
S2v	Temperature after compressor 2	°C
S3v	Temperature before compressor 1	°C
S4v	Temperature before compressor 2	°C
S5v	Pressure after air filter	Pa
S6v	Pressure before air filter	Pa
S7v	Real heat transfer	W
S8v	Mass flow rate	kg/s

of models and fuse the data sources obtained a remarkable accuracy [50].

In the train’s HVAC system, the temperature and the concentration of CO<sub>2</sub> are managed by two ventilation subsystems, two cooling subsystems, and two heating subsystems. Figure 3 contains a simple scheme of the modelled HVAC system; Table 1 contains the set of sensors used in the real system, and Table 2 shows the set of virtual sensors. Sensors that measure the control variables of various components, such as damper positions, operational state of compressors or heaters, and so on, are not mentioned.

1) FAULT MODELLING AND MODEL SYNCHRONIZATION

The real system uses the sensors listed in Table 1 to detect failures. The model presented in this research includes air filter degradation over the timeframe of interest.

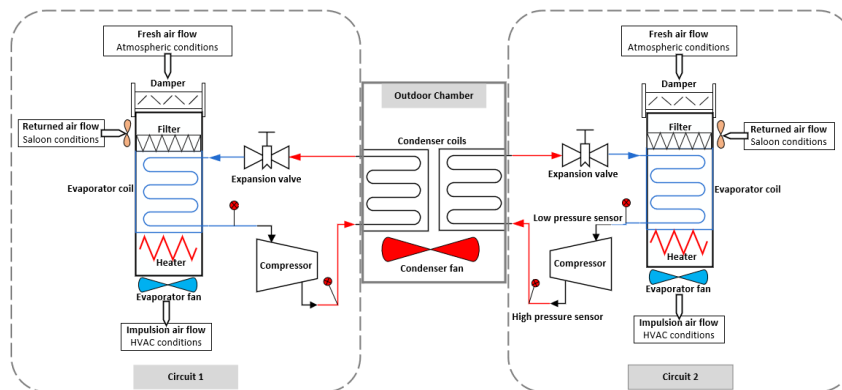


FIGURE 3. Scheme of the modelled HVAC system.

A component can be in a faulty state for a number of different reasons; thus, an FM is related to a cause. The FMA of complex systems commonly contains FMs that result from different causes but have similar effects in the same sensors.

The physics-based model includes the virtual sensors listed in Table 2. Soft sensing is important when there is an insufficient number of real sensors or there are relevant parameters to monitor (i.e., those for which there are insufficient data); the main requirement is that these sensors provide a response from sensor measurements.

Although the studied railway company does not have run-to-failure data for the HVAC system installed in the passenger train carriage, the CBM department previously performed experiments to assess the response of the HVAC system while increasing the mass of dust fed into the air filter. This makes it possible to model the degradation of the air filter. The timeframe used here is selected because the maintenance department had recorded the weight of the filter after its replacement and the number of working hours. By considering the air filter's weight in healthy state, we can determine the relations between degradation and time. Thus, the life cycle has an exponential degradation.

The experiments assessing filter degradation (mentioned above) also evaluated the signals obtained by the sensors listed in Table 1, the pressure of the air before and after the filter, and the mass flow rate. The responses of these parameters and the mass of dust fed into the filter were recorded in a dataset. These data provide key information for the present work. The input of the fault modelling is the mass of dust; these data are generated in the timeframe based on the information provided by maintainers.

Since the physics-based model used in this research and its capability to generate data in a faulty state were already developed and validated [49], [50], we validate the parameters monitored with virtual sensors, including pressure after air filter, pressure before air filter, and mass flow rate, using the previously generated data.

The parametrization of the physics-based model is a key step in first synchronizing the model with the real system and then validating it. The uncertainty of the parameters and

observations makes synchronization a stochastic problem. As the ideal validation of a physics-based model implies obtaining the whole posterior distribution of the parameters and suggests a high computing burden, in most cases, the parameters that enable physics-based models to fit the system behaviour are estimated, and the values that obtain the best results are selected [51].

## B. DEEP CONVOLUTIONAL NEURAL NETWORK AS DATA-DRIVEN MODEL

Convolutional neural networks (CNNs) have commonly been used for spatial pattern analysis to learn spatial features, but CNNs are showing remarkable success in many other research and industrial applications, such as vegetation remote sensing [52], seismo-acoustic event classification [53], computer vision [54], RUL estimation [55], among others. Like all typical neural network-type models, CNNs are neuron-based. The neurons are distributed in layers and can learn hierarchical representations. CNN's unique network architecture reduces the complexity and overfitting of a neural network. The structure comprises a number of layers; the initial layer is the input layer, i.e., raw data, and the last layer is the output, e.g., RUL prediction. At least one convolutional layer is included as a hidden layer; the convolutional layer involves multiple filters with raw input data it generates features and exploits patterns. Convolutional layers include optimizable filters that modify the input or preceding hidden layers. The number and size of filters define the depth of a convolutional layer. The resulting transformations are processed by the following pooling layers which extract the most significant local features in a way that matches the output.

As mentioned, CNNs are commonly used to learn abstract spatial features. Input data are usually prepared in a two-dimensional (2-D) format, but one-dimensional (1-D) and three-dimensional (3-D) formats can be also employed to learn spectral features and spatial features, respectively [56], [57].

### 1) INPUT SEQUENCE DATA AND CONVOLUTION LAYER

The raw data are processed and used to generate synthetic data. Processing includes data normalization and sliding

window operations. The input data sample is then generated. This study's input data are prepared in a 2-D sequence format. The data are processed and sorted, with the first dimension representing the number of selected signals and the second dimension representing the length of the time sequence of each signal. The signals used to build this prognostic model were collected by sensors located in different parts and subsystems of the HVAC system; this means the relations between the spatially neighbouring signals in the data sample are not notable. Therefore, the input and the signal maps are put in 2-D format and the convolution layers in 1-D. The network architecture selected for the study only applies 1-D convolution along the time sequence direction; thus, only the trends in one signal at a particular time are considered. This, in turn, means the order of signals and features does not affect the training process. The 1-D sequential data are assumed to be  $x = [x_1, x_2, x_3, \dots, x_N]$  where  $N$  denotes the length of the sequence.

Multiple filters of different lengths can be applied in convolution layers. Bigger numbers and larger sizes of filters lead to the ability to detect more complex patterns, generally resulting in both higher accuracy and a heavier computational burden [37], [58]. A balance must be reached in real cases; therefore, in this prognostics study, five convolutional layers are stacked successively for feature extraction with an increasing number and size of filters in subsequent layers.

## 2) ACTIVATION FUNCTIONS

The most commonly used activation functions in CNN are the rectified linear unit (ReLU) and the hyperbolic tangent function (Tanh). They are used to solve difficult problems. ReLU is an activation function that preserves the positive values and removes the negative values from the output of neurons, i.e., feature maps; this reduces the interdependence among parameters and speeds up the calculation [59]. The Tanh function ensures the output of neurons is within a value range of  $-1$  to  $1$ . The CNN in this study uses the ReLU activation function.

## 3) DROPOUT

Dropout is a regularized technique used when training NNs. This simple method helps minimize overfitting during training. Overfitting a model results in remarkable performance on the training dataset and poor performance on the testing dataset. Dropout is applied to avoid the extraction of the same features repeatedly and to reduce co-adaptation of units with the training data. In practice, randomly selected neurons (i.e., hidden neurons) are ignored during the training phase; thus, these neurons are not included in the forward propagation training process. Dropout is turned off during the testing phase; this implies that all the hidden neurons are activated in the testing process [60].

## 4) REGRESSION OUTPUT

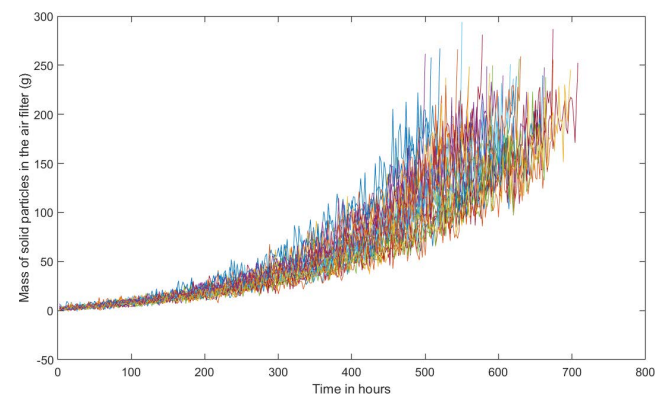
Regression is a predictive layer that involves predicting a numerical output given some input. Although some methods

require predicting more than one numeric value, these are known as multiple-output regressions. This study applies the most typical use of regression layers, the prediction of a single numeric value, i.e., the RUL.

## III. EXPERIMENTAL STUDY

### A. DEEP CONVOLUTIONAL NEURAL NETWORK AS DATA-DRIVEN MODEL

In this section, the proposed HyMA is demonstrated and evaluated on a synthetic dataset with run-to-failure data of an air filter installed in a train's HVAC system. Real working conditions recorded on-board the railway are used as input to the physics-based model. Data contain various scenarios of the real working conditions; we choose real data containing services longer than one hour and services for not just the ventilation mode but also the heating or cooling operation mode. These are chosen because of their relevance for practical applications.



**FIGURE 4.** Traces of mass of solid particles for the run-to-failure trajectories of the air filter used in the studied HVAC system.

These situations are subjected to the same failure mode. As mentioned, the run-to-failure data of the failure mode are modelled based on the previous experiment to assess the response of the HVAC system while increasing the mass of dust fed into the air filter and also on expert knowledge, where experts explained the initial degradation state of each filter can vary from 5-10% of the health index. Therefore, the failure mode modelled to generate the response of the system has a variability of the initial degradation state equivalent to 8% of the health index. Figure 4 contains an overview of the traces of the degradation imposed on the air filter. The traces define a growing abnormal (exponential) condition until filter failure, or, in terms of time, the end-of-life time ( $t_{EOL}$ ). The test developed to assess the behaviour of the HVAC system while working with the air filter at different levels of degradation provides the following information: (1) The more dust in the air filter, the less the air filter works as it should; this results in a growing exponential condition until filter failure. (2) The maximum level of mass of dust is in the range of 180-200 grams usually reached in a period of around 500-700 working hours or 21-29 days. Thus, the

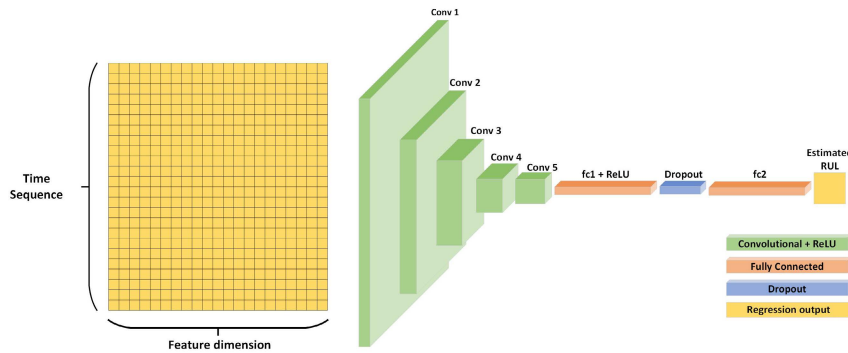


FIGURE 5. Deep CNN architecture for RUL estimation.

dataset generated for filter degradation contains 26 traces; the initial mass of dust at  $t_0$  is in the range of 0-16 grams, and it reaches the level of 200 grams of dust in the range of 497.9011-704.9124 working hours.

The dataset contains multivariate timeseries data of sensor readings and their corresponding RUL. In the dataset, the rows are the numbers of the sensors (see Tables 1 and 2). The length of each row is given by the length of the vector that contains the mass of dust over the timeframe; this differs from one observation to another. The operating conditions can also differ from observation to observation. Thus, each observation contains data generated under different conditions and degradation processes.

An overview of the traces generated as filter degradation is represented in Figure 4. This gives an overview of the initial and final mass of dust in the air filter at  $t_0$  and  $t_{EO}$ , respectively.

### B. DATA PRE-PROCESSING AND TIME SEQUENCE PROCESSING

Once the condition monitoring (CM) data are generated and given RUL labels, the next step is to create a data-driven model. The proposed data-driven model is a deep CNN which is expected to approximate the system dynamics based on previous observations, control variables, and model parameters [61].

The multi-variate temporal data generated by the physics-based model contain measurements from 15 sensors, as shown in Tables 1 and 2. Although the control variables are not included in the sensor list, some sensors measure constant output in the air filter’s lifetime; hence, they do not provide the surrogate model valuable information for RUL estimation. Sensor measurements identified in Tables 1 and 2 as S5, S6, S7, S5v, S6v, S7v, and S8v are used as the raw input features.

The data from each sensor are normalized using the Z-score normalization method to have zero mean and unit variance:

$$x_{norm}^{i,j} = (x^{i,j} - \mu^j) / \sigma^j \quad (1)$$

where  $x^{i,j}$  is the original value of the  $i$ -th data point of the  $j$ -th sensor, and  $x_{norm}^{i,j}$  denotes the normalized value of  $x^{i,j}$ .

$\mu^j$  and  $\sigma^j$  denote the means and standard deviations of the original measurement data from the  $j$ -th sensor, respectively.  $\mu^j$  is used for centering, and  $\sigma^j$  is used for scaling the data; therefore, this standardization method does not produce normalized data with the exact same scale for every sensor.

In general, industrial applications cannot validate the precision of RUL estimation of a system at each time step without an accurate physics-based model [38]. Therefore, the suitability of the deep CNN is tested using a set of data from the generated CM datasets.

### 1) TIME SEQUENCE PROCESSING

Time sequence processing has huge potential for prediction performance. Generally, more information can be obtained from temporal sequence data than from a single time step in multi-variate data. Therefore, the data are prepared to use multi-variate temporal information by defining a time window.

The size of the time window is 19 single time-steps, and the timestamp is 1.01 seconds. The size is defined based on the method the HVAC system uses to operate cooling and heating subsystems. As shown in Figure 3, the HVAC system has two cooling and heating subsystems. When the cooling operational mode must work, the HVAC system switches the compressors off and on; every compressor is individually working for fewer than 20 seconds until reaching a comfortable temperature. The same occurs with the heating operational mode.

Once the time window is defined, all the sensor data within the window are collected to form a feature vector used as input for the CNN.

### C. PROPOSED NETWORK ARCHITECTURE

Deep NN models have shown an excellent ability to capture hidden complex information from raw input signals and to trace complex relations between inputs and target labels. As mentioned, a deep CNN is chosen in this study to find a mapping that relates the input to a target label. The main reasons for this choice are the remarkable accuracy obtained in other applications (mentioned above) and the simplicity of using multivariate timeseries data taken from sensor mea-

surements. Figure 5 shows the architecture of the deep CNN model built for RUL estimation.

First, the input data are prepared in 2-D format, but the convolutional operation is actually performed in 1-D, as mentioned previously. The dimension of the input is defined by the time sequence dimension and the number of selected features.

Second, five convolutional layers are used for feature extraction. The filter sizes are 3, 7, 9, 13, and 15, and the number of filters is 28, 56, 112, 224, and 448, respectively. The obtained output is the number of filters of feature maps, and the dimension of each feature map is the same as the original input sample.

Third, the output goes to a fully connected layer with 100 neurons; this step multiplies the input by the weight matrix and adds a bias vector. The output goes to a dropout layer that randomly sets input elements to zero, with the probability set to 0.5, before going to the last fully connected layer with one output. This last output is the input of the regression layer.

Fourth, the output of each convolutional layer is the input of a batch normalization layer. This technique allows the model to be trained with mini-batches instead of the full dataset, thus speeding up training and using higher learning rates.

All the layers up to the dropout layer use ReLU as the activation function. To improve the prognostic performance and the training process, we use the Adam optimizer. A reduction of the learning rate of 0.003 every 5 epochs and an initial learning rate of 0.2 are defined; the maximum number of epochs is set to 30, and the training process uses mini-batches with 5 observations at each iteration because fewer mini-batches implies lower computational resources.

TABLE 3. Parameters of the DCNN algorithm.

Parameter	Value
Filter size	3/7/ 9/13/15
Number of filters	28/56/112/224/448
Number of selected features	7
Initial mass of dust at t0	0-16 grams
t <sub>FOL</sub>	500-700 working hours
Training process	Adam optimizer
Hidden neurons	100
Dropout rate	0.5
Convolutional layers	5
Initial and final learning rate	0.02 – 0.002
Epoch number	30
Mini-batch size	5

### 1) DEEP CNN PROCEDURE

Figure 1 gives an overview of the methodology presented in this paper for combining data obtained from a physics-based model and with data obtained from sensors installed in the real system. The details of the deep CNN working process and how it estimates RUL are explained in this section.

The measurements of the sensors selected for RUL estimation of the air filter are indexed in Tables 1 and 2 as

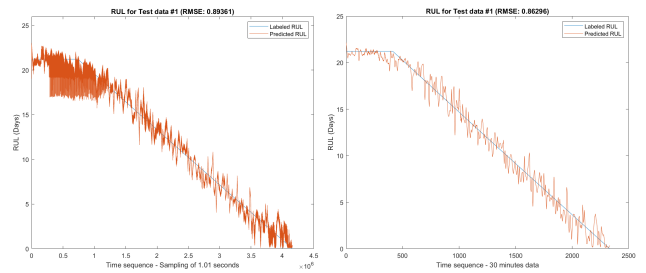


FIGURE 6. RUL estimations for testing dataset 1. Plot on left contains the RUL estimated every 1.01 seconds. Plot on right contains the mean value of 30 minutes of RUL estimations.

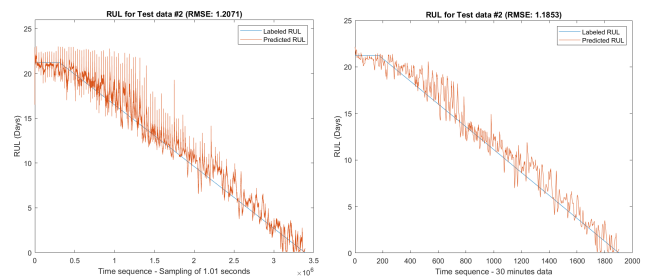


FIGURE 7. RUL estimations for testing dataset 2. Plot on left contains the RUL estimated every 1.01 seconds. Plot on right contains the mean value of 30 minutes of RUL estimations.

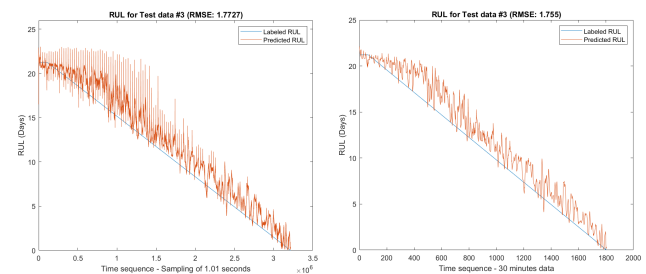


FIGURE 8. RUL estimations for testing dataset 3. Plot on left contains the RUL estimated every 1.01 seconds. Plot on right contains the mean value of 30 minutes of RUL estimations.

S5, S6, S7, S5v, S6v, S7v, and S8v. The corresponding data are normalized using the Z-score normalization method to have zero mean and unit variance. Next, datasets for training and testing processes are prepared. The datasets are in 2-D format and contain the time sequence information within the defined time window length for each sample. Therefore, signal processing experience or expertise on prognostics is not needed in the proposed data-driven model.

The normalized datasets are labelled with the RUL values. The deep CNN uses the normalized training data as input and the RUL values as the target output of the network.

As mentioned, the Adam optimizer is used with predefined mini-batches and epochs. The maximum number of training epochs is 30, and for each epoch, the training dataset is randomly divided into five mini-batches to speed up the training process. The batch size affects the training performance [37]. It is usual to see mini-batches comprising hundreds of units in



the literature, but it is recommended to use smaller sizes when the datasets are extremely large. The number of mini-batches is set to five in this study based on iterative calculations achieving a balance between computational cost and training performance. Moreover, the learning rate is reduced over the training process: the initial learning rate is 0.2 for fast optimization, and the final learning rate is 0.002 for stable convergence. The weights in each layer of the CNN are optimized according to the mean loss function of each mini batch using the root mean square error (RMSE) values; see equation 2.

$$\text{RMSE} = \sqrt{\left(\sum_{i=1}^N (z_{fi} - z_{oi})^2\right) / N} \quad (2)$$

where  $z_{fi}$  is the predicted value,  $z_{oi}$  is the actual value, and  $N$  is the number of samples.

The testing data samples are loaded into the trained model for RUL estimation; then, the prognostic accuracy is obtained. The main parameters set in the DCNN algorithm are listed in Table 3.

#### IV. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

This section describes the prognostic performance of the proposed HyMA for RUL estimation. We use a computational system with Intel(R) Xeon(R) Gold 5120 CPU with 4 sockets and 14 cores per socket, and 1 TB RAM.

As mentioned, the synthetic data are generated by the physics-based model to obtain the air filter's run-to-failure data. The data are generated by using historical data taken from the real system as input in the physics-based model. The failure mode is modelled based on expert knowledge and data collected by sensors in a previous experiment studying the behaviour of the HVAC system when working, with different air filters being fed dust. Hence, the synthetic data contain information on the HVAC system's air filter at different levels of degradation in the selected timeframe.

The RUL prediction results of three testing datasets are represented in Figures 6, 7, and 8. Every figure contains two plots; the plot on the left contains the RUL predicted using all data collected by the sensor with the actual sampling time, and the plot on the right contains the RUL predicted by the HyMA, but in this case, every sample contains the mean value of 30 minutes of data, thus giving a smoother result. The figures show the RUL values estimated by the proposed HyMA are generally close to the labelled values. Prediction accuracy tends to increase in the regions where the HVAC system works close to failure. When the system is working close to failure, it can be identified by the proposed HyMA, thus providing better prognostics results.

The proposed HyMA is carefully designed to overcome the drawbacks of each model. The physics-based model is developed to infer information about system degradation, as data are lacking on advanced stages of degradation of some failures. Once the lack of data is overcome, a deep

learning model is used to estimate the RUL from both the CM data and the data generated by the physics-based model.

#### V. CONCLUSION AND FUTURE RESEARCH

This paper proposes an HyMA for the prognostics of an HVAC system installed in a passenger train carriage. The proposed HyMA is the combination of a physics-based model and a deep learning algorithm based on CNN for predicting the RUL of complex systems. The physics-based model and the informative features of the failure are modelled and calibrated using information and CM data from the real system and expert knowledge. The ability to generate supplementary operation conditions for different contexts allows us to compensate for the lack of monitoring data. The methodology presented in this paper uses the data generated as input to a deep CNN to develop the HyMA for RUL prediction.

We evaluate the performance of the HyMA on synthetic datasets which contain run-to-failure data of the air filter installed in an HVAC system. The synthetic data were generated by a previously developed and validated physics-based model, and the deep CNN obtained good experimental results.

The proposed HyMA enables the possibility of implementing prognostics models for critical components of complex and critical systems – those components for which it is usually difficult or impossible to obtain data when they are working in advanced stages of degradation. The CM data are not recorded in these situations, but these data are key for developing a robust prognostics model. A lack of data related to a critical component has been overcome in this paper, and a remarkable prognostics performance has been demonstrated by our HyMA. Thus, the proposed HyMA offers a promising direction for future research in PHM applications.

Future research should develop and evaluate the use of the proposed methodology for other critical components of the same system. Future work should also evaluate the HyMA for RUL estimation while the system works with various components close to failure; this would result in a robust HyMA that can be deployed in a real system to support decision-making.

#### REFERENCES

- [1] N. O. Omri, Z. A. Masry, S. Giampiccolo, N. Mairot, and N. Zerhouni, "Data management requirements for PHM implementation in SMEs," in *Proc. Prognostics Syst. Health Manag. Conf. (PHM-Paris)*, May 2019, pp. 232–238.
- [2] K.-S. Wang, "Towards zero-defect manufacturing (ZDM)—A data mining approach," *Adv. Manuf.*, vol. 1, no. 1, pp. 62–74, Mar. 2013.
- [3] *CEN—EN 13306 Maintenance—Maintenance Terminology*, Standard EN 13306, European Committee for Standardization, 2017.
- [4] Z. Shi and A. Chehade, "A dual-LSTM framework combining change point detection and remaining useful life prediction," *Rel. Eng. Syst. Saf.*, vol. 205, Jan. 2021, Art. no. 107257.
- [5] Y. Li, S. Peng, Y. Li, and W. Jiang, "A review of condition-based maintenance: Its prognostic and operational aspects," *Frontiers Eng. Manag.*, vol. 7, no. 3, pp. 323–334, Sep. 2020.

- [6] J. Huang, J.-X. You, H.-C. Liu, and M.-S. Song, "Failure mode and effect analysis improvement: A systematic literature review and future research agenda," *Rel. Eng. Syst. Saf.*, vol. 199, Jul. 2020, Art. no. 106885.
- [7] D. Galar and U. Kumar, "Diagnosis," in *Maintenance*. Cambridge, MA, USA: Academic, 2017, ch. 5, pp. 235–310.
- [8] W.-S. Yun, W.-H. Hong, and H. Seo, "A data-driven fault detection and diagnosis scheme for air handling units in building HVAC systems considering undefined states," *J. Building Eng.*, vol. 35, Mar. 2021, Art. no. 102111.
- [9] T. Li, Y. Zhao, C. Zhang, J. Luo, and X. Zhang, "A knowledge-guided and data-driven method for building HVAC systems fault diagnosis," *Building Environ.*, vol. 198, Jul. 2021, Art. no. 107850.
- [10] A. Montazeri and S. M. Kargar, "Fault detection and diagnosis in air handling using data-driven methods," *J. Building Eng.*, vol. 31, Sep. 2020, Art. no. 101388.
- [11] C. Zhang, Y. Zhao, Y. Zhao, T. Li, and X. Zhang, "Causal discovery and inference-based fault detection and diagnosis method for heating, ventilation and air conditioning systems," *Building Environ.*, vol. 212, Mar. 2022, Art. no. 108760.
- [12] K. Verbert, R. Babuška, and B. De Schutter, "Combining knowledge and historical data for system-level fault diagnosis of HVAC systems," *Eng. Appl. Artif. Intell.*, vol. 59, pp. 260–273, Mar. 2017. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S095219761630255X>
- [13] C. Chen, T. Xu, G. Wang, and B. Li, "Railway turnout system RUL prediction based on feature fusion and genetic programming," *Measurement*, vol. 151, Feb. 2020, Art. no. 107162.
- [14] O. Bektas, J. A. Jones, S. Sankararaman, I. Roychoudhury, and K. Goebel, "A neural network filtering approach for similarity-based remaining useful life estimation," *Int. J. Adv. Manuf. Technol.*, vol. 101, nos. 1–4, pp. 87–103, Mar. 2019.
- [15] Z. Guo, Y. Wan, and H. Ye, "An unsupervised fault-detection method for railway turnouts," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 11, pp. 8881–8901, Nov. 2020.
- [16] F. Wang, Z. Zhong, G. Wang, and T. Tang, "A penalized convolution model for oil leakage detection in electrohydraulic railway point systems," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2021.
- [17] H. Yu and H. Li, "Pump remaining useful life prediction based on multi-source fusion and monotonicity-constrained particle filtering," *Mech. Syst. Signal Process.*, vol. 170, May 2022, Art. no. 108851.
- [18] H. Meng and Y.-F. Li, "A review on prognostics and health management (PHM) methods of lithium-ion batteries," *Renew. Sustain. Energy Rev.*, vol. 116, Dec. 2019, Art. no. 109405.
- [19] N. Omri, Z. Al Masry, N. Mairouf, S. Giampiccolo, and N. Zerhouni, "Towards an adapted PHM approach: Data quality requirements methodology for fault detection applications," *Comput. Ind.*, vol. 127, May 2021, Art. no. 103414.
- [20] L. Liao and F. Köttig, "A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction," *Appl. Soft Comput.*, vol. 44, pp. 191–199, Jul. 2016.
- [21] M. Pecht and J. Gu, "Physics-of-failure-based prognostics for electronic products," *Trans. Inst. Meas. Control*, vol. 31, nos. 3–4, pp. 309–322, Jun. 2009.
- [22] A. Rivas, G. K. Delipei, and J. Hou, "Predictions of component remaining useful lifetime using Bayesian neural network," *Prog. Nucl. Energy*, vol. 146, Apr. 2022, Art. no. 104143.
- [23] U. Leturiondo, O. Salgado, and D. Galar, "Validation of a physics-based model of a rotating machine for synthetic data generation in hybrid diagnosis," *Struct. Health Monitor.*, vol. 16, no. 4, pp. 458–470, Jul. 2017.
- [24] H. Wang, Y. Chen, C. W. H. Chan, J. Qin, and J. Wang, "Online model-based fault detection and diagnosis strategy for vav air handling units," *Energy Buildings*, vol. 55, pp. 252–263, Dec. 2012.
- [25] D. An, N. H. Kim, and J. H. Choi, "Practical options for selecting data-driven or physics-based prognostics algorithms with reviews," *Rel. Eng., Syst. Safety*, vol. 133, pp. 223–236, Jan. 2015.
- [26] S. Bendapudi, J. Braun, and E. Groll, "A dynamic model of a vapor compression liquid chiller," in *Proc. Int. Refrig. Air Conditioning Conf.* West Lafayette, IN, USA: Purdue Univ., 2002.
- [27] J. Poon, P. Jain, I. C. Konstantakopoulos, C. Spanos, S. K. Panda, and S. R. Sanders, "Model-based fault detection and identification for switching power converters," *IEEE Trans. Power Electron.*, vol. 32, no. 2, pp. 1419–1430, Feb. 2017.
- [28] M. H. Hafizi and A. Izadian, "Model-based fault diagnosis of a DC–DC boost converters using hidden Markov model," in *Proc. IEEE Int. Conf. Electro-Inf. Technol. (EIT)*, Rapid City, SD, USA, May 2013, pp. 1–4.
- [29] S. Y. Chu, X. Zan, and A.-T. Avestruz, "Electromagnetic model-based foreign object detection for wireless power transfer," *IEEE Trans. Power Electron.*, vol. 37, no. 1, pp. 100–113, Jan. 2022.
- [30] S. Iyengar, S. Lee, D. Irwin, P. Shenoy, and B. Weil, "WattHome: A data-driven approach for energy efficiency analytics at city-scale," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, Jul. 2018, pp. 396–405.
- [31] W. J. N. Turner, A. Staino, and B. Basu, "Residential HVAC fault detection using a system identification approach," *Energy Buildings*, vol. 151, pp. 1–17, Sep. 2017.
- [32] M. Xin, W. Jiao, and L. Da-Zi, "Fault diagnosis of nuclear power plant based on simplified signed directed graph with principal component analysis and support vector machine," in *Proc. Chin. Autom. Congr. (CAC)*, Hangzhou, China, Nov. 2019, pp. 3082–3087.
- [33] K. Choi, S. M. Namburu, M. S. Azam, J. Luo, K. R. Pattipati, and A. Patterson-Hine, "Fault diagnosis in HVAC chillers," *IEEE Instrum. Meas. Mag.*, vol. 8, no. 3, pp. 24–32, Aug. 2005.
- [34] S. M. Namburu, M. S. Azam, J. Luo, K. Choi, and K. R. Pattipati, "Data-driven modeling, fault diagnosis and optimal sensor selection for HVAC chillers," *IEEE Trans. Autom. Sci. Eng.*, vol. 4, no. 3, pp. 469–473, Jul. 2003.
- [35] T. Aven, "On the meaning of a black swan in a risk context," *Saf. Sci.*, vol. 57, pp. 44–51, Aug. 2013.
- [36] F. Tahir, M. T. Islam, J. Mack, J. Robertson, and D. Lovett, "Process monitoring and fault detection on a hot-melt extrusion process using in-line Raman spectroscopy and a hybrid soft sensor," *Comput. Chem. Eng.*, vol. 125, pp. 400–414, Jun. 2019.
- [37] X. Li, Q. Ding, and J.-Q. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Rel. Eng. Syst. Saf.*, vol. 172, pp. 1–11, Apr. 2018.
- [38] Z. Chong, P. Lim, A. K. Qin, and K. C. Tan, "Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2306–2318, Apr. 2017.
- [39] S. Zhao, Y. Zhang, S. Wang, B. Zhou, and C. Cheng, "A recurrent neural network approach for remaining useful life prediction utilizing a novel trend features construction method," *Measurement*, vol. 146, pp. 279–288, Nov. 2019.
- [40] K. P. Tran, H. D. Nguyen, and S. Thomassey, "Anomaly detection using long short term memory networks and its applications in supply chain management," in *Proc. 9th IFAC Conf. Manuf. Modeling, Manag. Control MIM*, Berlin, Germany, 2019, pp. 2408–2412.
- [41] J. Zhang, Y. Jiang, S. Wu, X. Li, H. Luo, and S. Yin, "Prediction of remaining useful life based on bidirectional gated recurrent unit with temporal self-attention mechanism," *Rel. Eng. Syst. Saf.*, vol. 221, May 2022, Art. no. 108297.
- [42] J. Liu, F. Lei, C. Pan, D. Hu, and H. Zuo, "Prediction of remaining useful life of multi-stage aero-engine based on clustering and LSTM fusion," *Rel. Eng. Syst. Saf.*, vol. 214, Oct. 2021, Art. no. 107807.
- [43] J. Caceres, D. Gonzalez, T. Zhou, and E. L. Drogue, "A probabilistic Bayesian recurrent neural network for remaining useful life prognostics considering epistemic and aleatory uncertainties," *Struct. Control Health Monitor.*, vol. 28, no. 10, p. e2811, Oct. 2021.
- [44] *Railway Applications—The Specification and Demonstration of Reliability, Availability, Maintainability and Safety (RAMS)—Part 1: Generic RAMS Process*, Standard EN 50126-1:2017, 2017.
- [45] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. 32nd Int. Conf. Mach. Learn.*, Lille, France, 2015, pp. 448–456.
- [46] *Railway Applications—Air Conditioning for Urban and Suburban Rolling Stock—Part 1: Comfort Parameters*, Standard EN 14750-1:2006, 2006.
- [47] R. L. Ackoff, "Towards a system of systems concept," *Manag. Sci.*, vol. 17, no. 11, pp. 661–786, 1971.
- [48] S. Abdoli, S. Kara, and M. Hauschild, "System interaction, system of systems, and environmental impact of products," *CIRP Ann.*, vol. 68, no. 1, pp. 17–20, 2019.

- [49] D. Galar, A. Gálvez, and D. Seneviratne, "Development and synchronisation of a physics-based model for heating, ventilation and air conditioning system integrated into a hybrid model," *Int. J. Hydromechtron.*, vol. 1, no. 1, p. 1, 2021.
- [50] A. Gálvez, A. Diez-Oliván, D. Seneviratne, and D. Galar, "Fault detection and RUL estimation for railway HVAC systems using a hybrid model-based approach," *Sustainability*, vol. 13, no. 12, p. 6828, Jun. 2021.
- [51] Y. Tian, M. A. Chao, C. Kulkarni, K. Goebel, and O. Fink, "Real-time model calibration with deep reinforcement learning," *Mech. Syst. Signal Process.*, vol. 165, Feb. 2022, Art. no. 108284.
- [52] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 24–49, Mar. 2021.
- [53] L. Trani, G. A. Pagani, J. P. P. Zanetti, C. Chapeland, and L. Evers, "DeepQuake—An application of CNN for seismo-acoustic event classification in The Netherlands," *Comput. Geosci.*, vol. 159, Feb. 2022, Art. no. 104980.
- [54] Y. Liu, H. Pu, and D.-W. Sun, "Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices," *Trends Food Sci. Technol.*, vol. 113, pp. 193–204, Jul. 2021.
- [55] A. L. Ellefsen, E. Bjørlykhaug, V. Æsøy, S. Ushakov, and H. Zhang, "Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture," *Rel. Eng. Syst. Saf.*, vol. 183, pp. 240–251, Mar. 2019.
- [56] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community," *J. Appl. Remote Sens.*, vol. 11, no. 4, p. 1, Sep. 2017.
- [57] L. Zhou, C. Zhang, F. Liu, Z. Qiu, and Y. He, "Application of deep learning in food: A review," *Compr. Rev. Food Sci. Food Saf.*, vol. 18, no. 6, pp. 1793–1811, 2019.
- [58] G. S. Babu, P. Zhao, and X.-L. Li, "Deep convolutional neural network based regression approach for estimation of remaining useful life," in *Database Systems for Advanced Applications (Lecture Notes in Computer Science)*. Cham, Switzerland: Springer, 2016, pp. 214–228.
- [59] C.-C. Chen, Z. Liu, G. Yang, C.-C. Wu, and Q. Ye, "An improved fault diagnosis using 1D-convolutional neural network model," *Electronics*, vol. 10, no. 1, p. 59, Dec. 2020.
- [60] G. S. Nandini and A. P. S. Kumar, "Dropout technique for image classification based on extreme learning machine," *Global Transitions Proc.*, vol. 2, no. 1, pp. 111–116, Jun. 2021.
- [61] M. Arias Chao, C. Kulkarni, K. Goebel, and O. Fink, "Fusing physics-based and deep learning models for prognostics," *Rel. Eng. Syst. Saf.*, vol. 217, Jan. 2022, Art. no. 107961.
- [62] N. Omri, Z. Al Masry, N. Mairrot, S. Giampiccolo, and N. Zerhouni, "Industrial data management strategy towards an SME-oriented PHM," *J. Manuf. Syst.*, vol. 56, pp. 23–36, Jul. 2020.
- [63] V. Singh, J. Mathur, and A. Bhatia, "A comprehensive review: Fault detection, diagnostics, prognostics, and fault modelling in HVAC systems," *Int. J. Refrig.*, Aug. 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0140700722003073>
- [64] Y. Guo, Z. Tan, H. Chen, G. Li, J. Wang, R. Huang, J. Liu, and T. Ahmad, "Deep learning-based fault diagnosis of variable refrigerant flow air-conditioning system for building energy saving," *Appl. Energy*, vol. 225, pp. 732–745, Sep. 2018.
- [65] Z. Sun, H. Jin, J. Gu, Y. Huang, X. Wang, and X. Shen, "Gradual fault early stage diagnosis for air source heat pump system using deep learning techniques," *Int. J. Refrig.*, vol. 107, pp. 63–72, Nov. 2019.



**ANTONIO GÁLVEZ** (Student Member, IEEE) received the bachelor's degree in mechanical engineering from the University of Malaga (UMA), in 2016, and the master's degrees in industrial engineering and in industrial management engineering from Deusto University, in 2018 and 2019. He is currently pursuing the Ph.D. degree with the Division of Operation and Maintenance, Lulea University of Technology (LTU). He develops the related master thesis with the DeustoTech

Research Center in energy and environment. His master thesis was developed in ZF Chassis Components Toluca, Toluca de Lerdo, Mexico. He develops his doctoral thesis in an industrial environment in reliability and maintenance with the Industry and Transport Division, Tecnalia. The topic of his doctoral thesis is conducted to extend the useful life of critical components and systems by researching and applying novel techniques. He is a Guest Professor with the University of La Rioja (UNIR), where was part of the opponent board of different master thesis. He has published related to his doctoral thesis in various journals and international conferences.



**DIEGO GALAR** was involved in the SKF UTC Center, Luleå, focused on SMART bearings and also actively involved in national projects with the Swedish industry or funded by Swedish national agencies, such as Vinnova. He has been involved in the raw materials business of Scandinavia, especially with mining and oil and gas for Sweden and Norway, respectively. Indeed, LKAB, Boliden or STATOIL have been partners or funders of projects in the CBM field for specific equipment such as

loaders, dumpers, rotating equipment, and linear assets. In the international arena, he has been a Visiting Professor with the Polytechnic of Braganza, Portugal, the University of Valencia and NIU, USA, and the Universidad Pontificia Católica de Chile. He is currently a Visiting Professor with the University of Sunderland, U.K., the University of Maryland, USA, the University of Stavanger, Norway, and Chongqing University, China. He is also a Professor of condition monitoring with the Division of Operation and Maintenance Engineering, Luleå University of Technology (LTU), where he is coordinating several H2020 projects related to different aspects of cyber physical systems, Industry 4.0, the IoT or industrial Big Data. He is also the Principal Researcher with Tecnalia, Spain, heading the Maintenance and Reliability Research Group, Division of Industry and Transport. He has authored more than five hundred journals and conference papers, books, and technical reports in the field of maintenance. He is also as a member of editorial boards, scientific committees, chairing international journals and conferences and actively participating in national and international committees for standardization and Research and Development in the topics of reliability and maintenance.



**DAMMIKA SENEVIRATNE** received the Ph.D. degree in offshore technology from the University of Stavanger and the Postdoctoral degree from the Lulea University of Technology. He is currently a Senior Researcher with Tecnalia Research and Innovation, Spain. His research interests include condition monitoring, operation and maintenance, risk-based inspection planning, risk-based maintenance, and RAMS analysis.

• • •