UNIVERSIDADE FEDERAL DE MINAS GERAIS
School of Engineering
Graduate Program in Electrical Engineering

MURILO CESAR OSORIO CAMARGOS FILHO

New Data-driven Methodologies for Fault Prognostics Using Evolving Fuzzy Models

Supervisor: Prof. Dr. Reinaldo Martinez Palhares

Belo Horizonte, Minas Gerais

2020

MURILO CESAR OSORIO CAMARGOS FILHO

New Data-driven Methodologies for Fault Prognostics Using Evolving Fuzzy Models

Thesis presented to the Graduate Program in Electrical Engineering (PPGEE) of the Federal Univesity of Minas Gerais (UFMG) in partial fulfillment of the requirements to obtain the degree of Doctor in Electrical Engineering.

Supervisor: Prof. Dr. Reinaldo Martinez Palhares

Belo Horizonte, Minas Gerais

2020

Camargos Filho, Murilo Cesar Osorio.
C172n New data-driven methodologies for fault prognostics using evolving fuzzy models [recurso eletrônico] / Murilo Cesar Osorio Camargos Filho. – 2020.
1 recurso online (104 f. : il., color.) : pdf.

Orientador: Reinaldo Martinez Palhares.

Tese (doutorado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Apêndices: f. 101-104.

Bibliografia: f. 82-99.
Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia elétrica - Teses. 2. Falha de sistema (Engenharia) - Teses. 3. Sistemas difusos- Teses. 4. Sistemas não lineares - Teses. I. Palhares, Reinaldo Martinez. II. Universidade Federal de Minas Gerais. Escola de Engenharia. III. Título.

CDU: 621.3(043)

Ficha catalográfica: Biblioteca Profº Mário Werneck, Escola de Engenharia da UFMG

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Escola de Engenharia
Programa de Pós-Graduação em Engenharia Elétrica

"New Data-driven Methodologies for Fault Prognostics using Evolving Fuzzy Models"

Murilo Cesar Osorio Camargos Filho

Tese de Doutorado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Doutor em Engenharia Elétrica.

Aprovada em 15 de dezembro de 2020.

Por:

Prof. Dr. Reinaldo Martínez Palhares
DELT (UFMG) - Orientador

Prof. Dr. Fernando Antônio Campos Gomide
DCA (UNICAMP)

Prof. Dr. Celso José Munaro
DEE (UFES)

Prof. Dr. Marcos Flávio Silveira Vasconcelos D'Angelo
DCC (UNIMONTES)

Prof. Dr. Walmir Matos Caminhas
DELT (UFMG)

Prof. Dr. Bruno Otávio Soares Teixeira
DELT (UFMG)

To my family.

AGRADECIMENTOS

Essa tese, antes de tudo, é fruto do incentivo de meus familiares e amigos que estiveram comigo nos momentos mais difíceis e compreenderam os períodos distantes. Agradeço, em especial, aos meus pais e à minha irmã que criaram o ambiente para que eu pudesse prosperar e à minha namorada, Brenda, que me apoiou incondicionalmente em todas as aventuras. Agradeço também aos meus primos Egydio e Adriano que, através de discussões nada produtivas, tornaram a experiência mais leve e agradável. Também devo mencionar meu amigo Léo, que dividiu sua casa e me ajudou em todas as adversidades durante minha estadia em BH.

Agradeço ao professor Reinaldo pela orientação honesta e rigorosa, pela prontidão em me ajudar a transpor todas as barreiras apresentadas, por compartilhar comigo suas ideias e posicionamentos, pelos conselhos e exemplos em situações pessoais e profissionais, e por incentivar a colaboração e a amizade no laboratório. Todos esses detalhes foram cruciais para a minha formação.

Devo agradecer ainda aos professores Renê Veloso, Steve Lacerda, Renato Dourado, João Batista e Rosivaldo (Gohnan), Luciana Cosme e Laércio Ives. Pois foram essenciais em minha formação e, de alguma forma, responsáveis pelo desejo de me tornar um pesquisador. Gostaria de agradecer imensamente ao professor Marcos Flávio (Marcão) por ter acompanhado toda minha trajetória acadêmica e profissional, se tornando um amigo e parceiro em diversas contribuições.

Sou extremamente grato aos colegas do D!FCOM, Fúlvia, Matheus, Márcia, Mariella, Gabriela, Everthon, Rodrigo, Rosi, Pedro, Luiz, Iury, por terem me acolhido e colaborado com minha formação por meio de intermináveis debates sobre os mais diversos temas, acadêmicos ou não. Agradeço, especialmente, aos amigos Pedro, Luiz e Iury que me ajudaram a transpor todas as dificuldades e encontraram formas de contribuir ativamente com meu trabalho; espero ter contribuído com o de vocês também.

Agradeço também à Petrobras pelo apoio financeiro durante essa etapa, mas principalmente por possibilitar a minha participação junto a um grupo de profissionais de excelência. Ao professor Benjamin por compartilhar sua imensa experiência em todos os aspectos relacionados à academia e ao mercado. Ao professor Walmir por despertar em todos o espírito incansável e a ânsia de colocar em prática as ideias. Ao professor Bomfim pelos conselhos e exemplos dados, nem sempre diretamente, que me influenciaram pessoalmente e profissionalmente. Ao professor Paim, que impactou diretamente este trabalho através de suas contribuições científicas. Aos colegas de trabalho, Nayron, Jefferson e Denis, por tornarem a experiência mais rica e agradável.

"Since all models are wrong the scientist must be alert to what is importantly wrong. It is inappropriate to be concerned about mice when there are tigers abroad."
(George Box, 1976)

RESUMO

Nesta tese são abordadas novas técnicas evolutivas baseadas em dados para o problema de prognóstico de falhas. Nesse tipo de problema, previsões acuradas de múltiplos passos à frente são essenciais para a determinação da vida útil remanescente (RUL, do inglês *remaining useful life*) de um determinado ativo. As soluções para prognóstico de falhas devem ser capazes de representar o comportamento tipicamente não-linear dos processos de degradação desses ativos e ser adaptável às particularidades de cada unidade. Nesse contexto, os sistemas nebulosos evolutivos são modelos capazes de representar tais comportamentos, além de serem capazes de lidar com o comportamento variante no tempo, também presente nesses problemas. Neste trabalho, propomos uma nova técnica de modelagem para sistemas nebulosos evolutivos que utiliza funções de pertinência Gaussianas multivariadas, tornando-a capaz de incorporar as complexas relações entre as variáveis de interesse, e um mecanismo de aprendizagem recursivo construído a partir de um fluxo de dados, ainda que dados históricos possam ser usados como ponto de partida. No mecanismo proposto, o conhecimento é gerido pelo monitoramento de limiares dinâmicos do erro de estimação. Além disso, é proposta uma metodologia para uso de tais técnicas em problemas de prognóstico de falhas, levando em consideração a propagação das incertezas do modelo em previsões de longo prazo. Três bases de dados bem estabelecidas são utilizadas para avaliar o modelo proposto em problemas de previsão de séries temporais e de prognóstico de falhas. Os experimentos indicam que o modelo proposto é competitivo em termos de precisão e número de parâmetros livres comparado a outros sistemas nebulosos evolutivos e pode se beneficiar da utilização tanto de dados históricos quanto de um fluxo de dados para estimar a RUL e sua incerteza. Além disso, na maioria dos cenários de teste, o modelo pode obter melhores desempenhos em relação às técnicas que não incorporam novos dados ou cuja base de conhecimento não é gerida com base nos erros de estimação.

Palavras-chave: Estimativa de RUL baseada em dados. Prognóstico de falhas. Sistemas nebulosos evolutivos. Modelos nebulosos Takagi-Sugeno.

ABSTRACT

This thesis addresses new data-driven evolving techniques to the problem of fault prognostics. In such problems, accurate predictions of multiple steps ahead are essential for the Remaining Useful Life (RUL) computation of a given asset. The fault prognostics' solutions must be able to model the typically nonlinear behavior of the degradation processes of these assets and be adaptable to each unit's particularities. In this context, the evolving fuzzy systems are models capable of representing such behaviors, in addition to being able to deal with time varying behavior, also present in these problems. In this work, we proposed a new modeling technique for evolving fuzzy systems that use multivariate Gaussian membership functions, making it able to incorporate the complex relationships between the variables of interest, and a recursive learning mechanism built upon a data stream, although historical data can be used as a starting point. In the proposed mechanism, the knowledge is managed through the monitoring of dynamic estimation error thresholds. Moreover, a methodology is proposed to use such techniques in fault prognostics problems, taking into account the model uncertainty propagation in long-term predictions. Three well-established data sets are used to evaluate the proposed model in problems of time-series prediction and fault prognostics. The experiments indicate that the proposed model is competitive in terms of precision and number of free parameters compared to other evolving fuzzy systems and can take advantage of both historical and stream data to estimate the RUL and its uncertainty. Furthermore, in most test scenarios, it may outperform other methods that do not manage new data incorporation or whose knowledge base management is not based on the estimation errors.

Keywords: Data-driven RUL estimation. Fault prognostics. Evolving fuzzy systems. Takagi–Sugeno fuzzy models.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

exTS	Evolving Extended Takagi-Sugeno
RUL	Remaining Useful Life
EOL	End of Life
PHM	Prognostics and Health Management
EBeTS	Error Based Evolving Takagi-Sugeno Model
CBM	Condition-based Maintenance
HI	Health Index
AI	Artificial Intelligence
TS	Takagi-Sugeno
UUT	Unit Under Test
FBeM	Fuzzy Set Based Evolving Modeling
MF	Membership Function
RLS	Recursive Least Squares
MSR	Most Similar Rule
MAPE	Mean Absolute Percentage Error
RA	Relative Accuracy
LCR	Last Created Rule
RMSE	Root Mean Squared Error
RMS	Root Mean Square
ANN	Artificial Neural Network
FT	Fault Threshold
PF	Particle Filter
ARMA	Autoregressive Moving Average
ANFIS	Adaptive Neuro-Fuzzy Inference System
NDEI	Non Dimensional Error Index
eNFN	Evolving Neo-Fuzzy Neural Network
eMG	Evolving Multivariable Gaussian
LSTM	Long Short-Term Memory
EFM	Evolving Fuzzy Model

LIST OF SYMBOLS AND NOTATION

$\mathbb{R}$	The set of real numbers
$\mathbb{R}^n$	The n-dimensional Euclidean space
$\mathbb{R}^{m\times n}$	The set of real matrices of order m by n
$\mathbb{N}$	The set of natural numbers, $\mathbb{N} = \{1, 2, \ldots\}$
$\mathbb{N}_{\leq k}$	The set $\{1, 2, \ldots, k\}$ for a given $k \in \mathbb{N}$
$\mathbf{x}^\top$ or $\mathbf{A}^\top$	Transpose of a vector $\mathbf{x}$ or a matrix $\mathbf{A}$
$(x_k)_{k=1}^M$	The sequence $\{x_1, \ldots, x_M\}$
x^+ or $\mathbf{x}^+$	A random variable or vector taking values in $\mathbb{R}$ or $\mathbb{R}^{n_x}$
$\mathbb{I}[E]$	Indicator function of the event E
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$
$\mathbf{I}_m$	Identity matrix of order m
$\mathbf{0}_{m\times n}$	Null matrix of order $m \times n$
$\inf\{\mathcal{S}\}$	The infimum of a sequence or set $\mathcal{S}$
ϕ	Fuzzy membership function defined by the map: $\mathbb{R} \to [0, 1]$
Φ	Fuzzy set $\{(u, \phi(u)) \,\vert\, u \in U\}$, U is the universe of discourse
$\mathrm{card}(\mathcal{S})$	Cardinality of a set $\mathcal{S}$
$\mathbb{E}[\cdot]$	Mathematical expectation
$\mathrm{Var}(\cdot)$	Variance
$\det(\mathbf{A})$	Determinant of a matrix $\mathbf{A}$
$H_{\mathcal{N}}(\cdot, \cdot)$	Hellinger distance between two Gaussian distributions
$\arg\max$	The argument that maximizes a function or set
$\in$	Denotes an element of

CONTENTS

1 INTRODUCTION

In the last few decades, the industry has been struggling to guarantee safety for critical systems. These systems need maintenance to prevent unexpected failures due to their deterioration or aging and to increase their reliability and availability (MA *et al.*, 2019b).

Maintenance policies are continually evolving to become more and more cheap and reliable. The earliest policy consists of unplanned actions in which the replacement of faulty components happens after a breakdown. As industrial systems become more complex, waiting for a breakdown potentially increases the costs of maintenance; this leads to the creation of preventive policies in which a periodic time interval for maintenance is set regardless of the asset's health status. This kind of policy also represents a significant expense for industries since components are replaced unnecessarily. More efficient policies, such as Condition-based Maintenance (CBM), are developed to handle these situations (JARDINE; LIN; BANJEVIC, 2006). The costs associated with each maintenance policy is shown in Figure 1.1 that places CBM in the optimal point in terms of total costs, balancing operating costs with maintenance costs (COBLE, 2010; TOMS; TOMS, 2008). Detecting undesirable events in systems such as oil and gas wells can prevent environmental accidents, production losses, and maintenance costs while sparing human lives, as shown in Vargas *et al.* (2019).

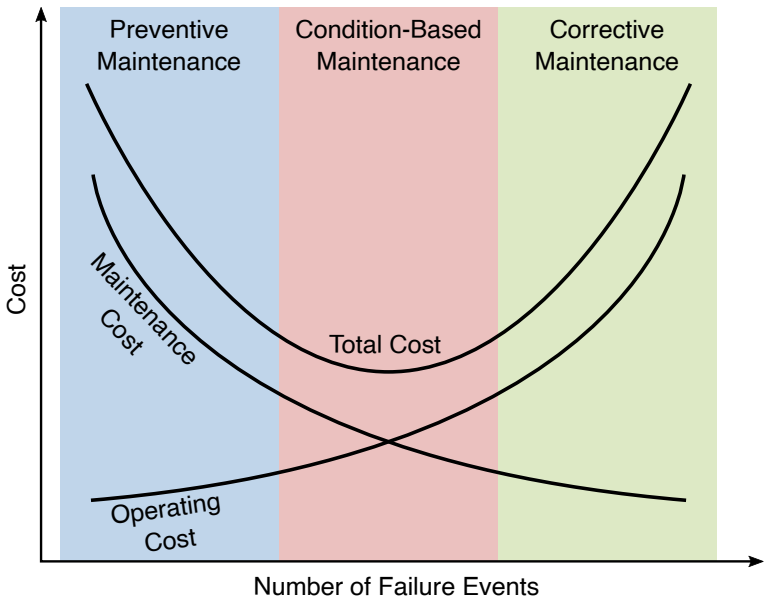


Figure 1.1 – Operating and maintenance cost chart.

Source: Coble (2010)

In the context of CBM, Prognostics and Health Management (PHM) enables the

use of real monitoring data to create relevant health indicators and trends (JOUIN *et al.*, 2016). This can aid in the system life-cycle support by reducing and eliminating inspections through early fault detection and prediction of impending faults (CHEN; YANG; HU, 2011). PHM can aid several systems such as: aircraft; industrial processes; water, gas and oil supply; manufacturing systems; transportation systems; (power) electrical and electronic systems; to become available, reliable, and safe by assisting the avoidance of catastrophic events (LY *et al.*, 2009). One primary task in CBM is health prognostics (LEI *et al.*, 2018), which consists in predicting the Remaining Useful Life (RUL) of machinery. Prognostics is usually composed of four processes, although some processes can be disregarded in some applications: data acquisition, Health Index (HI) construction, health stage division, and RUL prediction.

Different authors have tackled the RUL prediction problem by employing different methodologies. These methods are commonly classified into three categories: data-driven, model-based, and hybrid (KAN; TAN; MATHEW, 2015), which combines characteristics of the previous two. These categories can be divided into smaller groups in various ways (JAVED; GOURIVEAU; ZERHOUNI, 2017). The classification scheme adopted is the one proposed by Guo, Li & Li (2019) as depicted in Figure 1.2. Model-based approaches rely on mathematical models derived from the physics of components to assess their current and future health conditions (CUBILLO; PERINPANAYAGAM; ESPERON-MIGUEZ, 2016). When sufficiently complete, these models tend to outperform models in the other categories (LIAO; KÖTTIG, 2014; LUO *et al.*, 2003). However, their restricted application is the main drawback of these approaches since it is difficult to model the physics of damage in complex systems (LEI *et al.*, 2018; XU; WANG; XU, 2014).

1.1 Motivation

Data-driven approaches are divided into statistical and Artificial Intelligence (AI) models. Statistical models generally perform RUL prediction by fitting available observations into empirical models to be presented as a probability density function conditioned on these available observations (SI *et al.*, 2011). They are useful in describing uncertainties caused by different variability sources (LEI; LI; LIN, 2016). However, there are still some challenges related to data-driven statistical methods, in general, e.g., the ability to estimate the RUL with very few data, how to fuse multi-dimensional information of condition monitoring data, and models that can deal with multiple degradation stages (SI *et al.*, 2011).

AI approaches (PALHARES; YUAN; WANG, 2019), such as Artificial Neural Network (ANN), also deals with complex systems without relying on physical models of the assessed system, which is challenging to obtain. Their purpose is to learn how to produce the desired output of complex systems, possibly nonlinear, by reacting to given

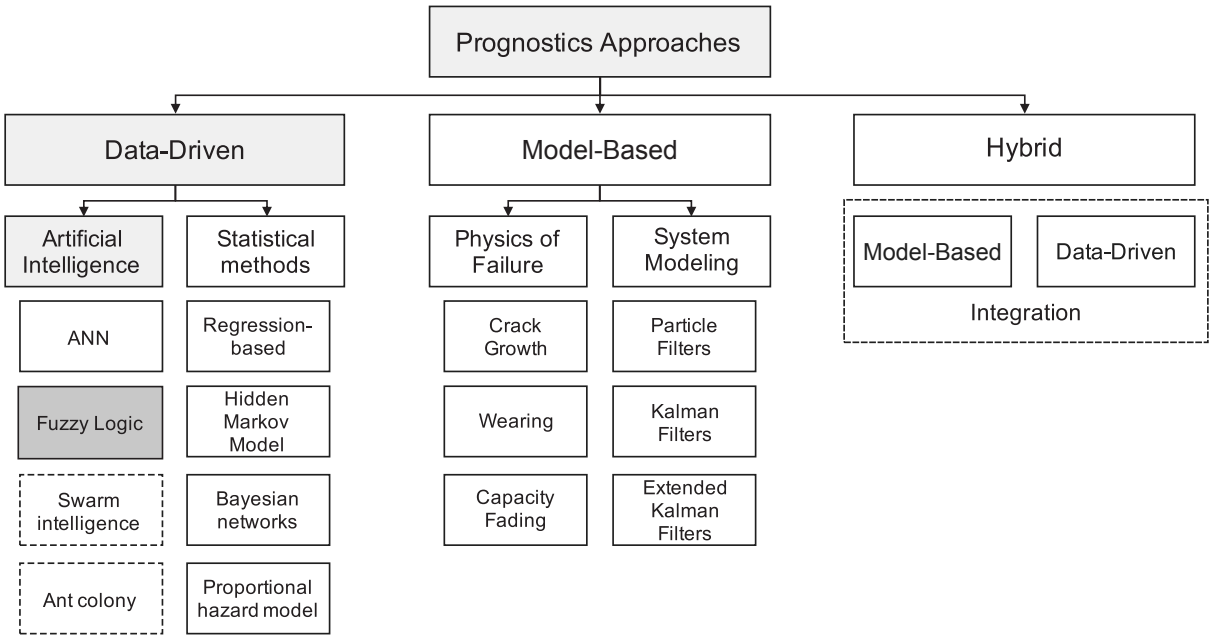


Figure 1.2 – Taxonomy of prognostics approaches adopted in this work.

Source: Guo, Li & Li (2019)

inputs, i.e., to learn the input-output relationship (AN; KIM; CHOI, 2015). However, it is necessary to retrain the neural network when operating conditions change (PENG; DONG; ZUO, 2010). As almost all data-driven methods, AI also needs a large amount of high-quality data from different operation conditions to train its models. AI approaches can also be used with expert-based approaches, such as fuzzy inference systems; neuro-fuzzy predictors have been extensively used for fault prognostics. For instance, in Wu *et al.* (2018) the Adaptive Neuro-Fuzzy Inference System (ANFIS) is used to create an HI through multi-sensor information fusion for a Computer Numerical Control (CNC) machine. Cosme *et al.* (2018) use a genetic algorithm to tackle the problem of impoverishment in particle filters, improving its application to fault prognostics to both rotating bearings and milling machines.

The necessity of high-quality data for prognostics techniques is a known problem in the literature that is receiving more and more attention. In Chiachío *et al.* (2020), a prognostics approach is proposed to predict the RUL of complex degradation processes, such as the fatigue crack propagation using a stochastic degradation model based on Markov chains. In the previous paper, the authors have shown that it is possible to manage the prediction uncertainty as more data become available in real-time. Similarly, Cadini *et al.* (2018) proposed a method based on particle filters and neural networks for prognostics in an environment where high-quality data may be scarce. It employs the proposed algorithm to predict the RUL of Li-ion batteries.

In general, these approaches have fixed structures, despite being adaptive, i.e., they rely on the assumption of a time invariant environment with respect to a training

data set. This assumption often does not hold, making the aforementioned approaches not suitable for real-time prognostics where human intervention is not always possible to redefine the model's domain when needed. A way of tackling the time invariance assumption is to develop strategies based on multiple models. Cosme *et al.* (2019) present a technique called IMMF, that consists of the interacting multiple filters theory associated with an ANFIS model to predict the RUL of rotating bearings. Other way to tackle the time invariance assumption is to develop evolving models, whose structures are not fixed and their knowledge-base is built based on a data stream instead of a data set, allowing the learning of complex behavior from scratch (CORDOVIL *et al.*, 2020). The ability to model complex dynamics places the Evolving Fuzzy Model (EFM) as an interesting choice for prognostics applications in cases where it is not easy to describe or model time-varying and nonlinear characteristics of a system. However, there are few works in the fault prognostics field using these structures (EL-KOUJOK; GOURIVEAU; ZERHOUNI, 2011; GOURIVEAU; ZERHOUNI, 2012; RAMASSO; DENOEUX, 2014), and none of them addresses the uncertainty quantification problem. The Evolving Extended Takagi-Sugeno (exTS) method was used in El-Koujok, Gouriveau & Zerhouni (2011) to predict the health state of real aircraft engines and its capabilities for forecasting tasks were tested in Gouriveau & Zerhouni (2012) under different prediction architectures. Ramasso & Denoeux (2014) proposes the use of exTS algorithm along with an evidential Markovian classifier for prognostics in the CMAPPS turbofan data set.

1.2 Research objectives

In the context of PHM, predicting the RUL of a given asset along with its confidence bounds is of utmost importance. There are several ways to tackle this problem; however, this work's scope is limited to evolving data-driven methods. To meet the general objectives, the following specific aspects will be explored:

a) to propose methods capable of incorporating new data recursively, enabling its practical application;

b) to investigate the direct use of estimation errors as part of the learning mechanism in the process of model creation;

c) to investigate and propose techniques capable of quantifying the uncertainty inherent to the process of RUL prediction;

d) to apply the proposed methods in experimental data sets.

1.3 Contributions

A novel evolving data-driven fuzzy model called Error Based Evolving Takagi-Sugeno Model (EBeTS) is developed and employed to perform fault prognostics based

on both historical and Unit Under Test (UUT) data, i.e., a specific system for which prognostics is applied (SAXENA *et al.*, 2008a). The proposed method aims at mitigating some data-driven disadvantages while intensifying its advantages through an evolving learning mechanism that disregards large amounts of high-quality data to start operating – because of its evolving structure and adaptive parameters that are updated based on a control chart of the regression error. Since it is a rule based system, the degradation is modeled as a multiple stage phenomena, which is closer to real degradation processes. The consequent of the rules is designed to be a simpler representation of the system that can bear a physical meaning through expert knowledge input; the fuzzy aggregation of these simple local models can approximate even complex and nonlinear functions. In short, the contributions of this work are listed as follows:

a) a new evolving Takagi-Sugeno (TS) model denominated EBeTS, whose model creation criterion is based on the estimation errors rather than in premise variables alone, controlling the model complexity to reduce the estimation error;

b) an evolving prognostics approach that uses EBeTS to incorporate new data from the UUT instead of using only historical data;

c) a way of quantifying and propagating model uncertainties to provide confidence intervals for RUL prediction using TS models;

d) to validate the proposed approach prognostics capabilities, two experimental data sets were used: the accelerated ball bearings from PRONOSTIA[1] platform and the Li-ion batteries data set from NASA Ames Prognostics Center of Excellence (PCoE).

1.3.1 Published papers

Two papers that reflect the contributions of this work were published and one will be submitted to an international conference:

a) M.O. Camargos, I. Bessa, M.F.S.V. D'Angelo, L.B. Cosme, R.M. Palhares. "Data-driven prognostics of rolling element bearings using a novel Error Based Evolving Takagi-Sugeno Fuzzy Model". **Applied Soft Computing**, Elsevier BV, v. 96, p. 106628, 2020. DOI: 10.1016/j.asoc.2020.106628.

b) M.O. Camargos, I. Bessa, M.F.S.V. D'Angelo, R.M. Palhares. "Fault Prognostics of Rolling Bearings Using a Hybrid Approach". In: *21st IFAC World Congress* (2020), Berlim, Germany.

c) M.O. Camargos, I. Bessa, L.A.Q.C. Junior, P.H.S. Coutinho, R.M. Palhares. "Evolving fuzzy systems applied to Li-ion battery charge capacity prediction for fault prognostics". *To be submitted.*

[1] Provided for the PHM IEEE 2012 Prognostic Data Challenge.

The main contributions of this work, presented in Chapters 3–5, are developed in Camargos *et al.* (2020a). Moreover, in Camargos *et al.* (2020b), we propose a hybrid prognostics approach based on a novel degradation model for rolling bearings and use the results to extend the comparisons in Chapter 5.

1.4 Thesis outline

This thesis is organized in five chapters as follows. Chapter 1 presents the motivations for the research topic along with a brief literature review to points out some strengths and weaknesses of each type of fault prognostics algorithm. The research objectives and contributions are also presented in this chapter. Some background about fault prognostics steps, methodologies, and evaluation metrics is given in Chapter 2. Chapter 3 presents our main contribution: the new evolving TS model denominated EBeTS. Chapter 4 shows how to use the new proposed model for long-term predictions, making it applicable to prognostics tasks. The same Chapter 4 also shows the computation of confidence intervals for long-term predictions using TS models. In Chapter 5, the proposed model is tested in three experimental benchmarks; the first case study aims to test its ability to identify the input-output relation under different temperature time series conditions. The proposed approach is compared with two state-of-the-art EFMs, and showed competitive results in terms of accuracy and complexity metrics. In the second case study, experimental bearings degradation data from the PRONOSTIA platform (NECTOUX *et al.*, 2012) are used to validate the proposed prognostics approach. The proposed approach also shows improvement when compared to other state-of-the-art and popular methods for fault prognostics. The third experiment compares the performance of the proposed algorithm with additional state-of-the-art EFMs and a deep learning alternative for prognostics on NASA's Li-ion battery data set. Finally, Chapter 6 enumerates possible directions for continuity of this work.

2 BACKGROUND

Prognostics is an essential task in a CBM program; it deals with the problem of fault prediction before it occurs in an attempt to achieve zero-downtime performance (JARDINE; LIN; BANJEVIC, 2006). The International Standard Organization states that the goal of a prognostics process is the capability to predict the RUL (RASHID *et al.*, 2017). The prognostics program is generally composed of four processes: data acquisition, HI construction, health stage division, and RUL prediction (LEI *et al.*, 2018), as described in Section 2.1. In particular, the estimation of RUL can be done using several methods and tools that can be grouped into three main approaches: model-based prognostics, data-driven prognostics, and hybrid approaches as shown in Section 2.2 and in Figure 1.2. Finally, the techniques can be refined using available metrics, as listed in Section 2.3, to find bottlenecks in performance (SAXENA *et al.*, 2008a).

2.1 Prognostics steps

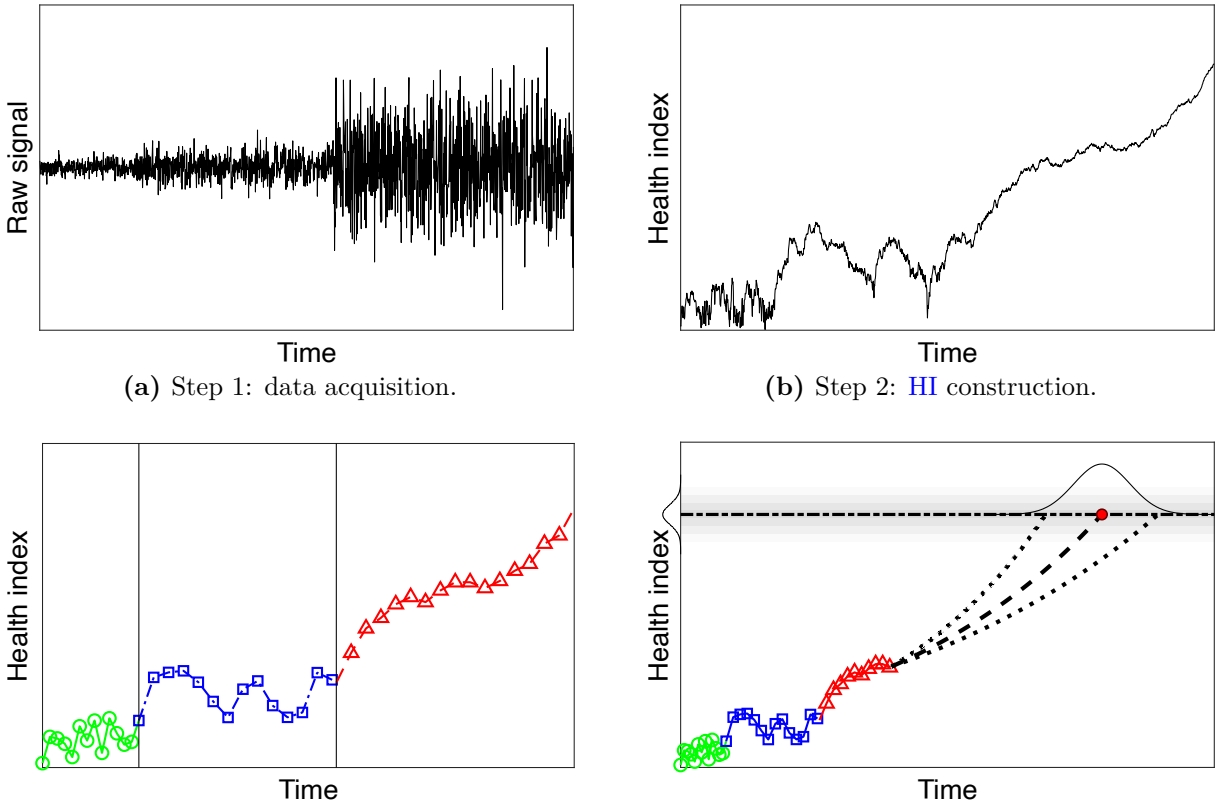
The four prognostics steps, namely data acquisition, HI construction, health stage division, and RUL prediction (LEI *et al.*, 2018), are shown in Figure 2.1 and described in the following subsections.

2.1.1 Data acquisition

A prognostics program generally begins by measuring data relevant to the system health (LEI *et al.*, 2018). PHM programs can require information on thousands of parameters in the entire life cycle of the product, such as temperature, vibration, pressure, acoustic levels, voltage, and current (CHENG; AZARIAN; PECHT, 2010).

The set of monitored parameters will vary according to the application. In structural prognostics, for instance, vibration sensor data is commonly used (SCHWABACHER, 2005). Its use has been reported in several applications involving rolling bearings (LI *et al.*, 2019; AHMAD *et al.*, 2019; RAGAB *et al.*, 2019), such as in gear systems of helicopters (CHEN; VACHTSEVANOS; ORCHARD, 2012) and wind turbines (LEITE; ARAÚJO; ROSAS, 2018). For battery prognostics, it is common to use temperature, current, voltage, impedance, and capacity data (LIAO; KÖTTIG, 2014; KHUMPROM; YODO, 2019; KHALEGHI *et al.*, 2019; RAZAVI-FAR *et al.*, 2019).

Although new sensor and communication development are driving technologies that gradually improve industrial data acquisition devices, obtaining high-quality data that contains fault evolution until failure still faces significant limitations (SAXENA *et al.*, 2008b). Lei *et al.* (2018) enumerates some reasons for that:



(a) Step 1: data acquisition.

(b) Step 2: HI construction.

(c) Step 3: health stage division.

(d) Step 4: RUL prediction.

Figure 2.1 – Prognostics program steps.

a) capturing and storing run-to-failure data can be quite expensive since real degradation processes can take months or years;

b) most candidate components for prognostics are not allowed to fail because such failures can lead to catastrophic events;

c) institutions with the resources to capture run-to-failure data are unlikely to make this data available due to the application's nature and the effort it took to acquire these data.

2.1.2 HI construction

The second process of the program is the construction of an HI that is robust to measurement noise and can reveal the degradation process of the evaluated asset. Choosing the right HI is a crucial step because it can simplify the prognostics modeling, leading to more accurate predictions (LEI *et al.*, 2018).

It is possible to group the HIs into physical HI and virtual HI. Physical HIs are related to signals that can be directly used as a health indicator, in other words, signals that are directly related to the physics-of-failure. On the other hand, virtual HIs are, in general, combinations of physical HIs (HU *et al.*, 2012).

Among the physical HIs, several authors use Root Mean Square (RMS) and kurtosis to predict the RUL on vibrating mechanisms (SUTRISNO *et al.*, 2012; LEI; LI; LIN, 2016; HUANG *et al.*, 2017; LIAO; TIAN, 2013; ZHANG; SI; HU, 2015); in Lei, Li & Lin (2016), kurtosis is used to detect incipient faults and determine the first predicting time, while Zhang, Si & Hu (2015) uses a filtered version of the feature to predict the RUL. Despite its importance, features like RMS and kurtosis are only effective to indicate system health conditions at certain statuses (XI; SUN; KRISHNAPPA, 2000). To mitigate that, physical HIs can be mixed to create virtual HIs using principal component analysis, self-organizing maps, Mahalanobis distance, and other transformations (JIN *et al.*, 2019; WANG, 2012; QIU *et al.*, 2003; WANG *et al.*, 2016; WIDODO; YANG, 2011; XI *et al.*, 2014).

In order to analyze the HIs created, some metrics are proposed in the literature (JAVED *et al.*, 2015; ZHANG; ZHANG; XU, 2016). Since degradation in real-world applications is irreversible, it is reasonable to assume that good HIs are monotonically increasing or decreasing. However, this behavior can not always be met, justifying the creation of a function to measure how monotonic a given HI is. The function, known as "monotonicity", can be defined[1] as

$$\text{Mon}\left((x_k)_{k=1}^{M}\right) = \frac{1}{M-1}\left|\sum_{k=1}^{M-1}\mathbb{I}[x_{k+1} \geq x_k] - \sum_{k=1}^{M-1}\mathbb{I}[x_{k+1} < x_k]\right|, \tag{2.1}$$

where $(x_k)_{k=1}^{M}$ is the HI sequence with M elements and $\mathbb{I}[E]$ is an indicator function of the event E, i.e., its output is unitary if the event E happens and zero otherwise. Another metric that can be used to evaluate an HI consists in the correlation between HI and time because it is expected that the component will degrade gradually (YANG *et al.*, 2016). This metric is known as "trendability" and can be defined[2] as

$$\text{Tre}\left((x_k, t_k)_{k=1}^{M}\right) = \frac{M\left(\sum_{k=1}^{M} x_k t_k\right) - \left(\sum_{k=1}^{M} x_k\right)\left(\sum_{k=1}^{M} t_k\right)}{\sqrt{\left[M\sum_{k=1}^{M} x_k^2 - \left(\sum_{k=1}^{M} x_k\right)^2\right]\left[M\sum_{k=1}^{M} t_k^2 - \left(\sum_{k=1}^{M} t_k\right)^2\right]}}, \tag{2.2}$$

where $(x_k, t_k)_{k=1}^{M}$ is the HI with operation time sequence with M elements. Other metrics and variations can be found in Lei *et al.* (2018).

2.1.3 Health stage division

Having an HI constructed, in observance to the metrics of monotonicity and trendability, the third step of a prognostics program is the identification of distinct degradation stages. The constructed HI, in general, presents different tendencies according to the asset degradation level, as shown in Figure 2.2. Therefore, it is useful to divide the HI time series in order to apply the model more adequate to the degradation level.

[1] A definition of this metric that is robust to random fluctuations can be found in Camci *et al.* (2013).
[2] A definition of this metric that is sensible to nonlinear correlations can be found in Carino *et al.* (2015), Lei *et al.* (2016a).

Moreover, the problem of health stage division can be seen as a pattern classification problem where the patterns are HIs tendencies. However, it is not precisely the fault detection and diagnosis problem but it can be used to detect the incipient degradation of an asset and provide the instant to start RUL prediction, also known as the first predicting time (LEI *et al.*, 2018).

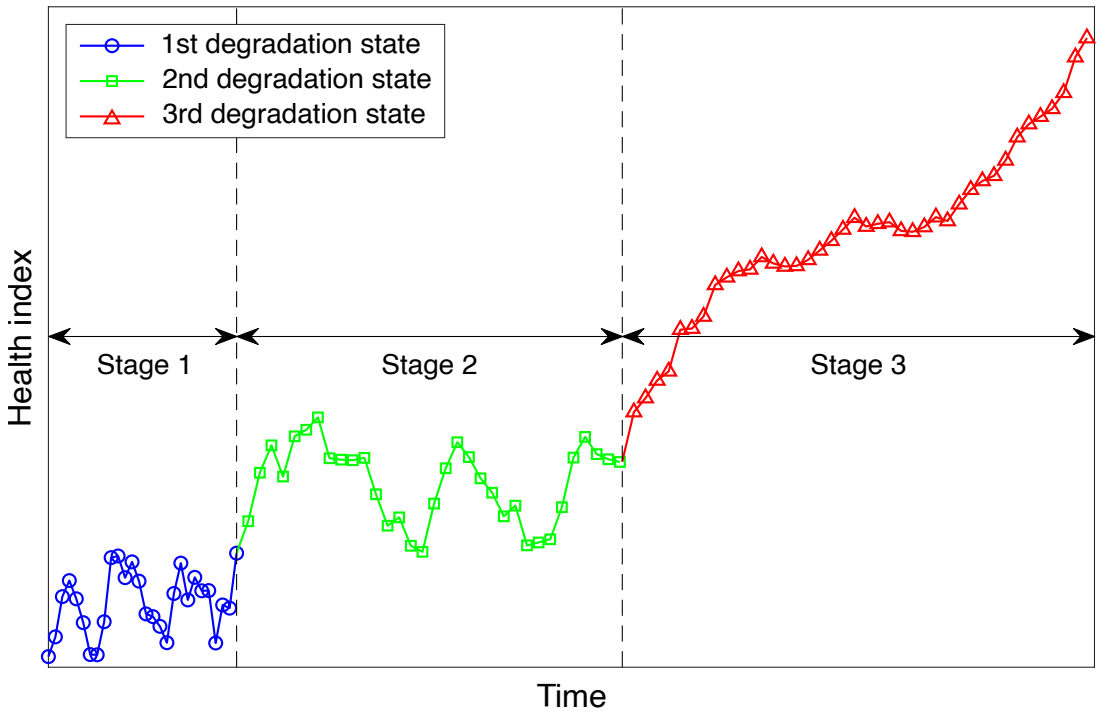


Figure 2.2 – Example of HI with multiple health stages.

Several methods can be used to detect the first predicting time. A simple strategy is to detect when the HI exceeds a Fault Threshold (FT). A few ways to define this FT are reported in the literature; some examples are the use of a 3σ interval as FT (WANG *et al.*, 2016) and monitoring some HI characteristics, such as the Hotelling T^2 statistics and Box-Cox transformations to define normalized FTs (ZHANG *et al.*, 2011; AJAMI; DANESHVAR, 2012). Trigger mechanisms are also proposed to increase random fluctuations robustness in detection; Li & He (2012) proposed to wait until a percentage of HI values exceeds the defined FT.

Other authors divide the degradation into multiple stages using clustering algorithms such as k-nearest neighbors, fuzzy c-means, and k-means to define the cluster prototypes (RAMASSO; ROMBAUT; ZERHOUNI, 2013; JAVED; GOURIVEAU; ZERHOUNI, 2015; SCANLON; KAVANAGH; BOLAND, 2013). It is also possible to use discrete state transition models, such as hidden markov models (SOUALHI *et al.*, 2014; LIU *et al.*, 2015), and AI classifiers, such as ANN, suport vector machines, and neuro-fuzzy systems (GUO *et al.*, 2016; SOUALHI; MEDJAHER; ZERHOUNI, 2015; ZURITA *et al.*, 2014).

2.1.4 RUL prediction

Predicting the RUL is an essential step in prognostics. It provides an estimation on how much time, from the current instant, remains before a possible fault occurrence, given the asset's age, condition, and past operation profile (JARDINE; LIN; BANJEVIC, 2006). Some authors define RUL from the HI point of view, i.e., RUL is the time left before the system's degradation state reaches a given FT (LI *et al.*, 2015; SI *et al.*, 2013), which is expressed by:

$$\hat{r}_k = \inf\{N \in \mathbb{N} \,|\, \hat{x}_{k+N} \geq \eta\}, \tag{2.3}$$

where $\hat{r}_k$ denotes the RUL computed at instant k given the observations of degradation state until k and $\mathbb{N}$ represents the natural numbers set; $\hat{x}_{k+N}$ represents the degradation state estimate, i.e., the HI, at time $k+N$ and η is the predefined FT.

More than an absolute estimation of the RUL, in this step, it is desired to provide a confidence interval in which the RUL lies, taking into account the inherent uncertainty of the program (TOBON-MEJIA *et al.*, 2012). The uncertainty in RUL estimation is shown in Figure 2.3, where the predicted degradation path reaches the FT before the correct path would. Another uncertainty present in the process of RUL prediction is related to the FT, which can be described as a probability distribution or a failure domain. However, in most of the works in the literature, including this one, the FT is represented by a constant line to simplify the RUL prediction process (LEI *et al.*, 2018). Therefore, (2.3) is a simplified version of the canonical RUL definition in Chiachío *et al.* (2015) where a failure domain is defined.

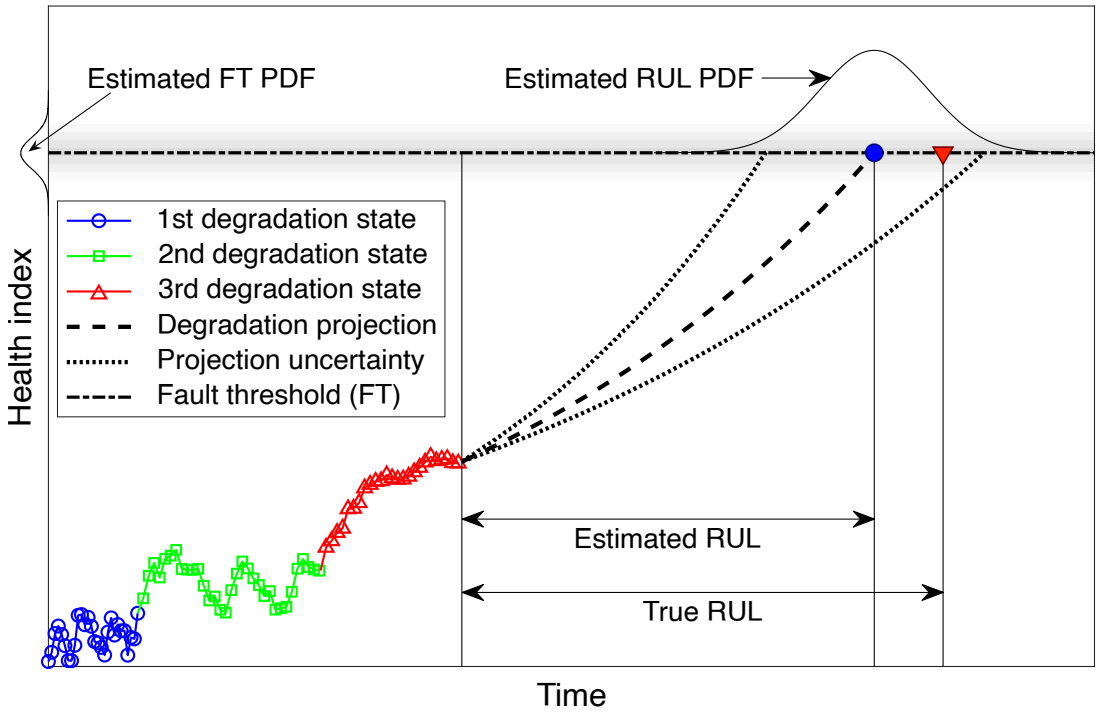


Figure 2.3 – Degradation stages and uncertainty in RUL prediction.

2.2 RUL estimation approaches

Several authors have tackled the RUL prediction problem by employing different approaches, but there is no consensus among authors about how these approaches should be divided. Some of them adopt granular classifications: in (SIKORSKA; HODKIEWICZ; MA, 2011), the approaches are grouped into knowledge-based models, life expectancy models, ANN, and physical models. Lei *et al.* (2018) proposed four groups: physics-based, statistical, AI, and hybrid. Other authors prefer to use a more general classification: in (JARDINE; LIN; BANJEVIC, 2006), the methods are classified into statistics, model-based, and AI. Despite the differences among classifications, in general, the approaches are grouped into physics-based, data-driven, and hybrid approaches (LEE *et al.*, 2014; KAN; TAN; MATHEW, 2015; JAVED; GOURIVEAU; ZERHOUNI, 2017). In this work, it is used the recent classification scheme proposed by Guo, Li & Li (2019) as depicted in Figure 1.2. In the next subsections, each category will be described, and some examples from the literature will be given.

2.2.1 Model-based approaches

Model-based techniques rely on mathematical models derived from the physics-of-failure of component to assess its current and future health condition (CUBILLO; PERINPANAYAGAM; ESPERON-MIGUEZ, 2016). It is the most accurate approach for prognostics, justifying its use in cost-justified applications, and it generally requires less training data than other approaches (HENG *et al.*, 2009). However, their restrict application is the main drawback of these approaches since it is difficult to model the physics of damage in complex systems (LEI *et al.*, 2018).

For instance, in rotating machinery, deriving the physical model of bearing degeneration is too complex (LIU *et al.*, 2017). Nonetheless, for crack growth analysis, it is possible to prove, through experiments, that the crack length is correlated to the number of cycles (PARIS; ERDOGAN, 1963). These experiments gave birth to the most used model for crack growth in prognostics, the Paris-Erdogan law:

$$\frac{da}{dN} = m\Delta K^b, \tag{2.4}$$

where a is the crack length, m and b are parameters depending on the bearing material, N is the number of cycles and K is an intensity coefficient that depends on the type and geometry of the crack (CUBILLO; PERINPANAYAGAM; ESPERON-MIGUEZ, 2016). For state propagation, a Particle Filter (PF) with a modified crack growth model, based on the Paris-Erdogan law, was used to predict the RUL of bearings in Liu *et al.* (2017). An improved exponential model based on the same law was proposed for RUL prediction in Li *et al.* (2015). In Nguyen, Liu & Zio (2019), an ensemble model that includes the Paris-Erdogan law was combined with a PF to predict the RUL. The law was also

used with PF for wind turbine bearing prognostics and insulated gate bipolar transistors prognostics (LU; CHRISTOU, 2019; WANG *et al.*, 2020).

Furthermore, prognostics is also frequently applied to Li-ion batteries. Several models are proposed to estimate the battery's state of charge and input it to the battery-management system in order to increase safety since Li-ion batteries can explode when overcharged (CHATURVEDI *et al.*, 2010). In Bartlett *et al.* (2016), a reduced-order electrochemical model is used to estimate the state of charge which is used to estimate the battery's capacity fade. It is also possible to use simpler models to estimate the capacity fade; some authors use a combination of exponential models:

$$C(k) = a \exp(bk) + c \exp(dk), \tag{2.5}$$

where C is the battery's capacity, k is the cycle number, the parameters a and b are related to the internal impedance, while d and c stand for aging rate, as proposed in He *et al.* (2011). The model described by (2.5) was used for RUL prediction in Downey *et al.* (2019) where a nonlinear least squares method with dynamic bounds estimates the models' parameters; in Walker, Rayman & White (2015) a PF is used to estimates the models' parameters.

2.2.2 Data-driven approaches

Data-driven approaches rely on the assumption that there exists available data from run-to-failure or past experiments to train degradation models in an offline stage and that data from the UUT will also be available to aid the prognostics up to the time where prediction starts (t_p). These models are not easy to explain because they are not related to any physical meaning, but they can be applied to other systems without knowing their complex physical models (LIAO; KÖTTIG, 2014). For this reason, the prognostics problem may benefit from an evolving learning mechanism that can adapt its parameters and change its structure according to new available data.

2.2.2.1 Statistical approaches

Statistical models generally perform RUL prediction by fitting available observations into empirical models to be presented as a probability density function conditioned on these available observations (SI *et al.*, 2011). They are useful in describing uncertainties caused by four variability sources: 1) temporal variability; 2) unit-to-unit variability; 3) nonlinear variability; and 4) measurement variability (LEI; LI; LIN, 2016). In Li *et al.* (2019), a Wiener process is used to predict the RUL of turbofan engines considering unit-to-unit variability. Gebraeel *et al.* (2005) consider temporal and unit-to-unit variability for prognostics in accelerated bearings through exponential degradation models with random and Brownian motion error terms. Measurement variability is considered in Tang *et al.* (2014) for RUL prediction of Li-ion batteries through Wiener processes.

In the realm of statistical approaches, stochastic processes play an essential role. A common strategy is to make random increments through time in the models' coefficients, this is known as random coefficient models (GEBRAEEL; PAN, 2008; JIN *et al.*, 2016; SONG; LIU, 2018). These increments are generally Gaussian, which allows the model to describe non-monotonic trends. However, the technique can have its use restricted since it is not capable of describing temporal variability (PANDEY; YUAN; NOORTWIJK, 2009). To address this problem, some authors use Wiener processes because they follow the Markov property in which the future state only depends on the current state (FANG; ZHOU; GEBRAEEL, 2015; LI *et al.*, 2017). A Wiener process was used in Li *et al.* (2019) to predict the RUL of a turbofan engine considering variability between units.

In prognostics, it is common to have monotonic HIs; in such cases, a gamma process would be a better choice because its increments follow a gamma distribution (TSUI *et al.*, 2015). In Son, Fouladirad & Barros (2016), a non-homogenous gamma process was used to model degradation and RUL prediction. Wiener and gamma processes are prominent approaches in prognostics and survey works dedicated to each of them are given by Zhang *et al.* (2018) and Noortwijk (2009), respectively. Another process used to model degradation is the inverse Gaussian process, in which the increments follow an inverse Gaussian distribution. Both gamma and inverse Gaussian processes are restricted to the Markov property and can only model monotonic trends, however, in comparison with gamma processes, inverse Gaussian processes are more flexible and can incorporate different kinds of random effects (PAN; LIU; CAO, 2016; PENG; ZHU; SHEN, 2019).

Simpler linear methods for predicting the RUL based on statistics are the autoregressive models (QIAN; YAN, 2015) and some of its variations, such as, autoregressive with exogenous inputs, Autoregressive Moving Average (ARMA) (CAESARENDRA *et al.*, 2011), autoregressive integrated moving average (ZHOU; HUANG, 2016; LIU *et al.*, 2019), and autoregressive moving average with exogenous inputs (YIU; WANG, 2007) can also be used. The ARMA models, for instance, are well-known and widely employed in the industry for prognostics applications (JARDINE; LIN; BANJEVIC, 2006). It is used in Qian, Yan & Hu (2014) to estimate the dynamic model of rolling bearings, allowing the prediction of faults with advance, while Barraza-Barraza *et al.* (2017) uses the extension of ARMA models that consider exogenous inputs to make prognostics in aluminum plates based on its crack growth. These linear regression models are easy to implement and results are easy to explain, however, they are not effective for long-term predictions (LIAO; KÖTTIG, 2014). Regarding the nonlinear versions of these approaches, such as nonlinear autoregressive moving average with exogenous inputs and nonlinear autoregressive with exogenous inputs, prognostics applications were not found in the literature, although they are commonly found in systems identification problems.

Furthermore, statistical approaches also include Markov models for RUL prediction,

as shown in Liu *et al.* (2015), where a continuous-time Markov model is used to derive the RUL distribution in both homogeneous and non-homogeneous cases. When the degradation state is not directly observable or is partially observable, it is possible to use hidden markov model and gaussian mixture model hidden markov model, as in Zaidi *et al.* (2011) and Ramasso & Denoeux (2014). However, these approaches rely on Markovian properties, which may not be consistent in real-world applications. Therefore, there are still some challenges related to data-driven statistical methods, e.g., the ability to estimate the RUL with very few data, how to fuse multi-dimensional information of condition monitoring data, and models that can deal with multiple degradation stages (SI *et al.*, 2011).

2.2.2.2 Artificial Intelligence approaches

ANNs are strategies based on AI commonly used in fault diagnostics and prognostics problems. Its purpose is to learn the input-output relationship of a complex system, possibly nonlinear (AN; KIM; CHOI, 2015). Parameters from the functions used between the layers along with weights and biases are estimated during the learning process (MENG; LI, 2019). ANN architectures consist of interconnected elements arranged according to the task to be done (AHMADZADEH; LUNDBERG, 2013).

A feed-forward neural network with one layer was used in Yang *et al.* (2016) to map features to HIs and predict the RUL of electric motors. In Xiao *et al.* (2017), the RUL of rolling bearings is predicted using a feed-forward neural network trained with the back-propagation algorithm. Razavi-Far *et al.* (2019) use the extreme learning machine as a learning algorithm to train the feed-forward neural network and predict the RUL of Li-ion batteries considering missing observations. To explicitly deal with time-series and use delayed outputs of the system in the model, recurrent neural networks were used in Liu *et al.* (2015) and Guo *et al.* (2017) to predict the RUL in Li-ion batteries and rolling bearings, respectively. Deep learning architectures, i.e., with multiple middle layers between input and output, can also be used for prognostics tasks. In Zhang *et al.* (2017), an ensemble of deep belief networks was trained using a multiobjective evolutionary algorithm, balancing the accuracy and diversity as conflicting objectives and predicting the RUL of aero-engines from NASA. The same database was used to show the effectiveness of using deep convolutional neural networks for fault prognostics in Li, Ding & Sun (2018). Furthermore, ANN approaches can deal with complex nonlinear systems disregarding physical models of the assessed system; however, it is necessary to retrain the neural network when operating conditions change and the result is a black box system (PENG; DONG; ZUO, 2010).

ANNs can also be combined with an inference system that mimics the human process of prediction through IF-THEN rules (WANG; GOLNARAGHI; ISMAIL, 2004). An example of this combination is the ANFIS since it merges ANNs with a fuzzy inference

system, improving its ability to predict time-series in comparison to autoregressive and feed-forward neural network models (JANG; SUN; MIZUTANI, 1996). Liu *et al.* (2019) used an ANFIS model based on the fuzzy c-means algorithm for fault prognostics in membrane fuel cells.

In addition to the variations of ANNs, a commonly used technique for prognostics is the support vector regression, based on the suport vector machine classifier proposed by Vapnik (2000). The hypothesis is that given a training data set, there exists a function in which its greatest deviation from training data is ε and, at the same time, it is as flat as possible (SMOLA; SCHÖLKOPF, 2004). To predict the RUL in Li-ion batteries, Wang & Mamo (2018) uses the differential evolution algorithm to obtain the kernel parameters of the support vector regression model. In Ma *et al.* (2019a), an ensemble of support vector regression models is proposed to predict the current health state of a Li-ion battery.

To deal with small size and high-dimensional data sets, the Gaussian process regression model can be used to achieve accurate prediction results, besides offering the uncertainty quantification using the state prediction variance (KAN; TAN; MATHEW, 2015). In Aye & Heyns (2017), an integrate Gaussian process regression model is used to predict the RUL of slow speed bearings through acoustic emission data. In Zhou *et al.* (2018), an ANN is used as covariance function of a Gaussian process regression to predict the RUL in Li-ion batteries.

2.2.3 Hybrid approaches

Each of the approaches previously treated has their own limitation; hybrid approaches tries to merge their qualities. In prognostics, it is common to associate physical models to filtering techniques, such as PF, allowing the application of these models in different environments and operational conditions since the model's parameters are estimated using known data through Bayes theorem (LIU *et al.*, 2017; NGUYEN; LIU; ZIO, 2019; LU; CHRISTOU, 2019; WANG *et al.*, 2020; WALKER; RAYMAN; WHITE, 2015).

A commonly used model for cutting tools is the Taylor tool life and its extensions (YEN *et al.*, 2004); Hanachi *et al.* (2019) merges the predictions from an empirical wear-time model with ANFIS predictions using a PF extension to manage the prediction uncertainties of a computer numerical control machine wear states. In Li & Liu (2019), an adaptive hidden markov model that uses Taylor's wear formula to predict the RUL of computer numerical control machines is proposed. A dynamic multi-stage support vector regression model is proposed in Tao, Zio & Zhao (2018) to recursively estimate the health state of computer numerical control machines and turbofan engines from NASA using recent observations through the sequential Monte Carlo paradigm. Li, Wang & Yan (2019) use Gaussian process regression to analyze the capacity incremental curve for RUL prediction and state of health estimation in Li-ion batteries. In Cosme *et al.* (2019),

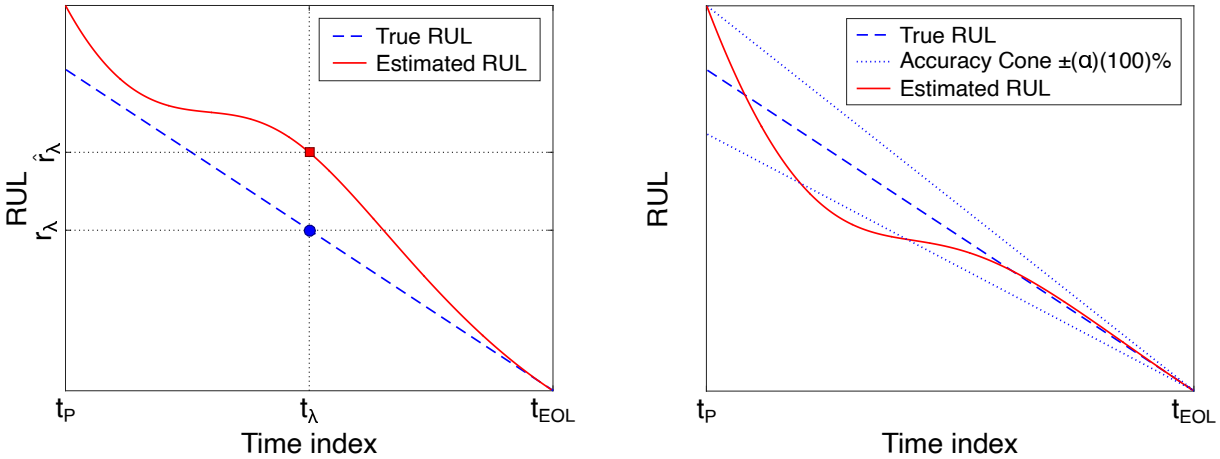
the limitation of having a single degradation model is addressed by using the Interacting Multiple Model filtering theory associated with fuzzy TS systems in a framework called IMMF in order to compute the RUL of rotating bearings.

2.3 Prognostics metrics

In critical systems health management, the prognostics methods need to be certified through rigorous metrics. Saxena *et al.* (2008a) present one of the most comprehensive work on this subject, where a set of metrics for RUL prediction are defined. Relative Accuracy (RA) is a metric to infer the prognostics quality at a given instant t_λ defined as:

$$\text{RA}_\lambda = 1 - \frac{|r_\lambda - \hat{r}_\lambda|}{r_\lambda}, \tag{2.6}$$

where r_λ and $\hat{r}_\lambda$ are, respectively, the actual and estimated RUL at t_λ, as shown in Figure 2.4a. The time instants t_P and t_{EOL} are the first predicting time and End of Life (EOL), respectively. Furthermore, a metric to evaluate the prognostics' results throughout the entire asset's life is the $\alpha - \lambda$ plot; it is usually used to evaluate prognostics strategies since it shows whether the predicted RUL falls within a goal region around the true RUL given by $\pm(\alpha)(100)\%$ (LALL; LOWE; GOEBEL, 2012). Figure 2.4b shows an example of this metric.



(a) True and estimated RUL at time index t_λ.

(b) True and estimated RUL along with $\pm(\alpha)(100)\%$ accuracy cone.

Figure 2.4 – Prognostics metrics.

In addition to metrics related to the prognostics' results, i.e., RUL prediction, it can be useful to quantify how similar the multi-step ahead state prediction is to the actual state. A metric for this task is the Mean Absolute Percentage Error (MAPE), defined as:

$$\text{MAPE}_\lambda = \frac{100}{N} \sum_{k=t_\lambda}^{t_\lambda+N} \left| \frac{x_k - \hat{x}_k}{x_k} \right|, \tag{2.7}$$

where N is the prediction horizon, x_k is the true system's degradation state, and $\hat{x}_k$ represents the predicted system's degradation state.

2.4 Evolving Fuzzy Models

As discussed in Chapter 2.2, regarding RUL prediction methods, all have their limitations. Physics-based methods have high accuracy, but the physics-of-failure models are hard to obtain, and its replication, in general, is limited to equipment of the same type operating under the same conditions. The statistical approaches are depicted with a single degradation model, limiting its applications to multi-step ahead predictions among the data-driven approaches. Also, most of these models assume a degradation rate that depends on the system's age without considering its health state. The AI-based models also have limitations: ANNs are not interpretable; their parameters are randomly or manually initialized, reducing its capacity of replication; and can use lots of computational resources in the training stage. In almost all data-driven approaches, including neuro-fuzzy systems, a large amount of high-quality data is required for training.

Given the discussed limitations of data-driven methods for RUL prediction, it is essential the proposition of new methods that can deal with nonlinear and time variant degradation models and can incorporate multiple degradation stages. Like all data-driven methods under time variant environmental and operational conditions, they need to be provided with adaptive learning capabilities (LIU; ZIO, 2016). In this context, the fuzzy inference systems are competitive candidates, since they can represent different degradation stages through a rule-based model whose combination can embody the system's nonlinearity. Moreover, several models, such as evolving TS and evolving neuro-fuzzy, are proposed to endow these systems of evolving learning mechanisms in which not only the model's parameters are adapted, but its structure also changes in time, i.e., there is no need to specify beforehand the number of degradation models (ŠKRJANC *et al.*, 2019). Moreover, EFMs can extract information through information granules, making the models linguistically interpretable and flexible enough to be incremented with expert knowledge, if needed (ANGELOV, 2010).

The practical applications of EFMs are various. Their recursive nature allows both real-time fault detection and diagnosis (LEMOS; CAMINHAS; GOMIDE, 2013; LEITE *et al.*, 2009; CORDOVIL *et al.*, 2019; INACIO; LEMOS; CAMINHAS, 2015) and systems identification and time-series prediction (LEITE *et al.*, 2012; LEITE *et al.*, 2015; ŠKRJANC *et al.*, 2019). This work's focus lies in the system identification and time-series prediction problems in which the majority of proposed models use variations of the TS fuzzy inference system, i.e., models whose rules consist of functional outputs. In their evolving formulations, these models have a flexible structure that adapts their number of rules, antecedent parameters and consequent parameters through data-streams. The whole process consists of a recursive incremental learning mechanism that decides about rule creation, exclusion and merging in the knowledge-base.

Throughout the years, different kinds of EFMs were proposed to explore various

learning mechanism nuances. The multiple-input multiple-output model described in Angelov & Zhou (2006), namely exTS, consists of a TS model with a functional consequent. It consists of partitioning the input/output data space through an extension of the concepts of clustering from the subtractive clustering approach (CHIU, 1994). The learning process of creating and excluding rules in the knowledge-base is related to recursively computed quality metrics such as the zone of influence of each cluster, their age and support size. The rules are represented by univariate Gaussian Membership Functions (MFs) for each premise variable.

The application of these algorithms to data-streams is one of the reasons for the efforts towards approaches with less computational costs. The Evolving Neo-Fuzzy Neural Network (eNFN) is a version of the neo-fuzzy neuron (YAMAKAWA *et al.*, 1992) to deal with data-streams, providing the ability to change the models' structure in time, i.e., the number of MFs and neurons can increase or decrease as its learning mechanism sees fit (SILVA *et al.*, 2014). The technique consists of tracking the fuzzy model's local and global prediction errors such that, when the local error is greater than the global one, the region will be refined by adding more rules. The exclusion procedure is done by analysing its inactivity in time. The standard neo-fuzzy neuron consists of zero-order TS functional rules where the domain of each input is granulated separately into complementary triangle MFs. The simplified structure of eNFN decreases its computational costs allowing its use in high-frequency real-time dynamic systems.

Defining the right resolution of EFMs whose knowledge is represented by fuzzy information granules is an aspect explored by Leite *et al.* (2012) in an effort to a multi-resolution analysis of the data, enabling the evolving system to use coarser granularities. Leite *et al.* (2012) proposes a fuzzy granular approach that adopts hyperboxes and fuzzy object of trapezoid nature to analyze input-output data streams under different resolutions and decide when to adopt coarser or more detailed granularities. The framework proposed in Leite *et al.* (2012) extends the standard TS model to include a linguistic component, in addition to the functional component, in the rule's consequent to facilitate model interpretation and to enclose the possible outputs of the model.

The application of evolving models for time-series prediction and systems identification has been producing efforts towards the definition of models that can account for the complex relationships between the input variables. Lemos, Caminhas & Gomide (2011) proposes an evolving fuzzy model, namely Evolving Multivariable Gaussian (eMG), with first-order functional and multivariate Gaussians as MF to represent the premise variables. This kind of MF can model the relation between input variables through a recursive computation of a dispersion matrix. The model uses a learning mechanism based on the participatory learning principle proposed in Yager (1990) that endows the algorithm with the capacity to classify whether a sample is an outlier or the first representative of a

new cluster.

The prognostics tasks addressed in this thesis are deeply related to regression and long-term forecasting of a degradation time-series. For this purpose, we hypothesize that two features are highly desirable: accounting for the complex relationships between input variables (or regressors) and explicitly using the estimation error to adjust the model's structure in time. A method with these features will be explored in the next chapter, where a new learning mechanism for EFMs that uses TS representation is proposed. It combines the idea of accounting for the complex relationships between the variables through multivariate Gaussian modeling with the knowledge-base control based on the estimation error tracking.

2.5 Chapter summary

This chapter has introduced the task of fault prognostics with details to its main steps, i.e., data acquisition, HI construction, health stage division and RUL prediction. The latter is one of the most prominent study objects in the literature of PHM; therefore, an overview of different classes of prognostics techniques is made covering the main advantages and limitations of each one. Moreover, a brief overview on evolving fuzzy inference systems is given.

3 ERROR BASED EVOLVING TAKAGI-SUGENO MODEL

In this chapter, a specific type of rule-based fuzzy inference system, the TS model, will be detailed. The proposed model, like many others in the literature, will be endowed with an evolving learning mechanism. This mechanism explicitly considers the prediction error as a trigger to create new rules. The hypothesis, tested in Chapter 5, is that controlling the one-step ahead prediction error will improve the system's multi-step ahead predictions. In addition, to enable its use for prognostics problems, a scheme for uncertainty quantification associated with the iterative propagation of such systems is proposed.

3.1 Problem formulation

A TS fuzzy model is one of the leading mathematical tools to represent a system using fuzzy concepts. Instead of using fuzzy sets to represent the consequent part, such as in Mamdani (1974), a TS fuzzy model employs a functional consequent, usually linear (NGUYEN *et al.*, 2019). In such systems, the i-th fuzzy rule, in a set of C **if-then** rules can be written as

$$\begin{aligned} \text{Rule } i\text{: IF } x_k^1 \text{ IS } \Phi_{i,k-1}^1 \text{ AND } \cdots \text{ AND } x_k^{n_x} \text{ IS } \Phi_{i,k-1}^{n_x} \\ \text{THEN } \hat{y}_{i,k} = \begin{bmatrix} 1 & \mathbf{x}_k^\top \end{bmatrix} \hat{\boldsymbol{\theta}}_{i,k-1} \end{aligned} \tag{3.1}$$

where $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is the vector of premise variables, $\hat{\boldsymbol{\theta}}_{i,k-1} \in \mathbb{R}^{n_x+1}$ is the estimated vector of consequent parameters. Moreover, $\Phi_{i,k-1}^j$ is a fuzzy set for $i \in \mathbb{N}_{\leq C}$ and $j \in \mathbb{N}_{\leq n_x}$, such that $\varphi_{i,k-1}^j \colon \mathbb{R} \to [0,1]$ is its MF. Throughout the text, $\mathbb{N}_{\leq k}$ will be used to denote the set of natural numbers up to k, such that $\mathbb{N}_{\leq k} = \{1, 2, \ldots, k\}$.

The aggregation of fuzzy sets in the antecedent represents the degree of activation $w_i \colon \mathbb{R}^{n_x} \to \mathbb{R}$ of a given rule and can be computed through the product t-norm[1]

$$w_{i,k-1}(\mathbf{x}_k) = \prod_{j=1}^{n_x} \varphi_{i,k-1}^j(x_k^j). \tag{3.2}$$

The output of the TS fuzzy model is given by a convex combination between the linear models in the consequent of each rule weighted by the activation degrees transformed to comply with the convex sum property, i.e., they need to be non-negative and sum one. The output of (3.1) can be inferred as follows:

$$\hat{y}_k = \sum_{i=1}^{C} h_{i,k-1}(\mathbf{x}_k)\, \hat{y}_{i,k}, \tag{3.3}$$

[1] Although the product t-norm is widely used in engineering applications of fuzzy logic (KARNIK; MENDEL, 1998), there are other t-norms, such as the ones described in Gupta & Qi (1991).

$$h_{i,k-1}(\mathbf{x}_k) = \frac{w_{i,k-1}(\mathbf{x}_k)}{\sum_{m=1}^{C} w_{m,k-1}(\mathbf{x}_k)}. \tag{3.4}$$

Given a fixed number of rules (C), the TS fuzzy model creates a coarse fuzzy partitioning in the antecedent domain and adjusts the parameters of the consequent polynomial models of degree one that can locally approximate the system. One of the challenges of working with data streams is not knowing beforehand the model structure, i.e., the number of rules (KASABOV; SONG, 2002). It is particularly relevant in the fault prognostics context since the degradation dynamics are typically nonlinear, time variant, and different for each UUT. In this work, the proposed EBeTS model is updated with its parameters and structure that are adaptively estimated based on the observations obtained at each time instant. Such property enables EBeTS to deal with nonlinear and time-varying dynamics through uncertain measurements and, consequently, with fault prognostics problems. With this scenario in mind, a new procedure to create rules based on a control chart of the estimation error is proposed. Furthermore, a policy based on the Hellinger distance, detailed in Section 3.2.2, is proposed to merge existing clusters.

3.2 Structure adaptation process

In order to consider the complex relations between input dimensions, a multivariate MF can be used instead of univariate MFs aggregated by a t-norm as in (3.2). According to Angelov (2010), the antecedent MF of TS fuzzy models are usually described by Gaussians. In this work, a Gaussian multivariate MF is used to compute the degree of activation, such that

$$w_{i,k-1}(\mathbf{x}_k) = \exp\left(-\frac{1}{2}\left(\mathbf{x}_k - \hat{\boldsymbol{\mu}}_{i,k-1}\right)^{\top} \hat{\boldsymbol{\Sigma}}_{i,k-1}^{-1}\left(\mathbf{x} - \hat{\boldsymbol{\mu}}_{i,k-1}\right)\right), \tag{3.5}$$

is used. The mean $\hat{\boldsymbol{\mu}}_{i,k-1}$ and the dispersion matrix $\hat{\boldsymbol{\Sigma}}_{i,k-1}$ of the antecedent MFs at time $k-1$ are recursively estimated for each rule according to (3.6a)–(3.6c) and the Sherman-Morrison formula (HAGER, 1989), as shown in Appendix A:

$$\boldsymbol{\Delta}_{i,k} = \mathbf{x}_k - \hat{\boldsymbol{\mu}}_{i,k-1}, \tag{3.6a}$$

$$\hat{\boldsymbol{\mu}}_{i,k} = \hat{\boldsymbol{\mu}}_{i,k-1} + \rho_{i,k}\,\boldsymbol{\Delta}_{i,k}, \tag{3.6b}$$

$$\hat{\boldsymbol{\Sigma}}_{i,k}^{-1} = \frac{1}{1-\rho_{i,k}}\left(\hat{\boldsymbol{\Sigma}}_{i,k-1}^{-1} + \rho_{i,k}\,\frac{\hat{\boldsymbol{\Sigma}}_{i,k-1}^{-1}\,\boldsymbol{\Delta}_{i,k}\,\boldsymbol{\Delta}_{i,k}^{\top}\,\hat{\boldsymbol{\Sigma}}_{i,k-1}^{-1}}{1+\rho_{i,k}\,\boldsymbol{\Delta}_{i,k}^{\top}\,\hat{\boldsymbol{\Sigma}}_{i,k-1}^{-1}\,\boldsymbol{\Delta}_{i,k}}\right), \tag{3.6c}$$

where $\rho_{i,k} = 1/n_{i,k}$ weights equally all observations regardless the time of occurrence, $n_{i,k}$ is the number of data points associated to the i-th rule at time instant k. The dispersion is initialized as an identity matrix and the mean is the first data point assigned to the

i-th rule, i.e., (3.6b) and (3.6c) are initialized as $\hat{\boldsymbol{\Sigma}}_{i,1} = \mathbf{I}_{n_x}$ and $\hat{\boldsymbol{\mu}}_{i,1} = \mathbf{x}_{i,1}$, respectively. The Recursive Least Squares (RLS) algorithm, described by (3.7a)–(3.7d) (ÅSTRÖM; WITTENMARK, 1996, p. 518), is employed to estimate the consequent parameters $\hat{\boldsymbol{\theta}}_{i,k}$:

$$\alpha_{i,k} = y_k - \tilde{\mathbf{x}}_k^\top \hat{\boldsymbol{\theta}}_{i,k-1}, \tag{3.7a}$$

$$\mathbf{g}_{i,k} = \mathbf{F}_{i,k-1}\,\tilde{\mathbf{x}}_k \left(1 + \tilde{\mathbf{x}}_k^\top \mathbf{F}_{i,k-1}\,\tilde{\mathbf{x}}_k\right)^{-1}, \tag{3.7b}$$

$$\mathbf{F}_{i,k} = \mathbf{F}_{i,k-1} - \mathbf{g}_{i,k}\,\tilde{\mathbf{x}}_k^\top \mathbf{F}_{i,k-1}, \tag{3.7c}$$

$$\hat{\boldsymbol{\theta}}_{i,k} = \hat{\boldsymbol{\theta}}_{i,k-1} + \alpha_{i,k}\,\mathbf{g}_{i,k}, \tag{3.7d}$$

where $\alpha_{i,k} \in \mathbb{R}$ is known as the *a priori* estimation error, $\mathbf{g}_{i,k} \in \mathbb{R}^{n_x+1}$ is the gain vector, $\mathbf{F}_{i,0} = 10^3 \times \mathbf{I}_{n_x+1}$ can be interpreted as the inverse sample autocorrelation matrix (LIU *et al.*, 2009) multiplied by a large constant that represents the parameters' uncertainty, and $\hat{\boldsymbol{\theta}}_{i,0} = \mathbf{0}_{n_x+1\times 1}$ are the initial values for the coefficients, i.e., a vector of zeros. The augmented vector $\tilde{\mathbf{x}}_k$ is defined as $\tilde{\mathbf{x}}_k = \begin{bmatrix}1 & \mathbf{x}_k^\top\end{bmatrix}^\top$. The system output (3.3) at instant k can be rewritten as

$$\hat{y}_k = \left(\mathbf{h}_{k-1}(\mathbf{x}_k)\right)^\top \hat{\boldsymbol{\Theta}}_{k-1}^\top \tilde{\mathbf{x}}_k, \tag{3.8}$$

where $\mathbf{h}_{k-1}(\mathbf{x}_k) = \begin{bmatrix}h_{1,k-1}(\mathbf{x}_k) & \cdots & h_{C,k-1}(\mathbf{x}_k)\end{bmatrix}^\top \in \mathbb{R}^C$ is the vector of normalized degrees of activation and $\hat{\boldsymbol{\Theta}}_{k-1} = \begin{bmatrix}\hat{\boldsymbol{\theta}}_{1,k-1} & \cdots & \hat{\boldsymbol{\theta}}_{C,k-1}\end{bmatrix}^\top \in \mathbb{R}^{n_x+1\times C}$ is a matrix with the consequent coefficients estimated with RLS in the previous iteration. The uncertainties of each local system are represented by the recursively estimated inverse sample autocorrelation matrix (3.7c) of the parameters.

3.2.1 Rule creation

The thresholds used for rule creating in evolving TS models are defined using, in general, three strategies: (i) estimation error; (ii) similarity measure between cluster prototypes; (iii) ϵ-completeness condition (ŠKRJANC *et al.*, 2019). Most models in the literature that create rules based on estimation error use predefined thresholds, in which the output error can not be above. The evolving fuzzy neural network (KASABOV, 2001) method uses a predefined error threshold along with a similarity measure between new observations and the allocated radial basis neurons to create rules; this is practically the same procedure used in both Dynamic Fuzzy Neural Networks (WU; ER, 2000) and Generalized Fuzzy Neural Networks (WU; ER; GAO, 2001), while Wu, Er & Gao (2001) introduces the concept of ϵ-completeness. Predefined error thresholds were also used in Rong *et al.* (2006), Tzafestas & Zikidis (2001), Pratama *et al.* (2014). The proposed EBeTS model uses a control chart for the residuals estimation as a trigger to create new

rules. The objective is not to force the system to reach a predefined estimation error through the excessive creation of new rules, overfitting the parameters.

The rule creation procedure is related to the local approximation error variation, i.e., it only takes into account what happens in the local affine model expressed by the consequent of the Last Created Rule (LCR) of the TS fuzzy model whose structure initiates with only one rule. The local approximation error is assumed to be a Gaussian random variable, therefore, it is fully described by its mean $\mu_{\mathbf{R}}$ and variance $\sigma^2_{\mathbf{R}}$, with $\mathbf{R}$ being a vector of stored residuals. When a new data point $(\mathbf{x}_k, y_k)$ becomes available at time step k, it is stored in matrix $\mathbf{X}$ and vector $\mathbf{Y}$ to be used in a possible merging procedure. The residual of the data point assignment to the LCR is computed as

$$\varepsilon_k = \left| \tilde{\mathbf{x}}_k^\top \hat{\boldsymbol{\theta}}_{\text{LCR},k} - y_k \right|, \tag{3.9}$$

where $\hat{\boldsymbol{\theta}}_{\text{LCR},k}$ are the consequent parameters of the LCR. In order to classify the data point as anomalies or not, the residual computed by (3.9) is classified through the following control chart

$$\frac{(\varepsilon_k - \mu_{\mathbf{R}})^2}{\sigma^2_{\mathbf{R}}} > (\chi^2)^{-1}_p(\omega), \tag{3.10}$$

where $(\chi^2)^{-1}_p(\omega)$ represents the inverse of the chi-squared statistic with p-degrees of freedom, resulting in an ellipse that covers at least $100\omega\%$ of the data under the Gaussian assumption; the number of degrees of freedom is chosen as the residual's dimension, therefore $p = 1$ (MOSHTAGHI; LECKIE; BEZDEK, 2016). Since inequality (3.10) depends on the residual variance, the first two data points assigned to a new rule (or to the first rule) will be automatically considered non-anomalies and stored in the residual vector $\mathbf{R}$; after this initialization, its mean and variance can be computed, allowing the control chart application. When the number of continuous anomalies exceeds the persistence index τ, arbitrarily defined to control the model complexity, **a new rule will be created**. Since some knowledge is available, the consequent parameters of the new rule will be initialized as

$$\hat{\boldsymbol{\theta}}_{\text{LCR}+1,0} = \frac{1}{C} \sum_{i=1}^{C} \hat{\boldsymbol{\theta}}_{i,k}. \tag{3.11}$$

The rule's parameters will then be updated with the last τ sequential anomalies detected, i.e., the last τ observations stored in $\mathbf{X}$ and $\mathbf{Y}$. Finally, the vector of residuals $\mathbf{R}$ along with data matrices $\mathbf{X}$ and $\mathbf{Y}$ will be cleaned, i.e., $|\mathbf{R}| = |\mathbf{X}| = |\mathbf{Y}| = 0$, where $|\cdot|$ represents the amount of elements of a matrix or vector. The other parameters are initialized in the same way as before.

3.2.2 Rule merge

Whenever a new rule needs to be created, i.e., the number of continuous anomalies exceeds the persistence index τ, the LCR will be compared with the other rules to decide

whether it needs to be merged or not. The algorithm only stores the data associated to the LCR. The stored data will be used to re-estimate the antecedent and consequent parameters of the Most Similar Rule (MSR) in a merge.

The metric to account for similarity needs to be able to compare the fuzzy sets in the antecedent, and in this work, it is described as multivariate Gaussians. A function to quantify the distance between two probability distributions is called f-divergence (CSISZÁR; SHIELDS, 2004). Since a threshold should be defined to merge rules or not, it is useful that the f-divergence has bounded output. In addition, it must also be symmetric, non-negative and respect the triangle inequality to induce a distance metric. The Hellinger distance is an f-divergence with all these properties (GIBBS; SU, 2002; CIESLAK *et al.*, 2012).

For two multivariate fuzzy sets Φ_1 and Φ_2 with Gaussian membership functions described by their means and dispersion matrices $(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, respectively, the squared Hellinger distance is defined as:

$$H^2(\Phi_1, \Phi_2) = 1 - \frac{\sqrt[4]{\det(\boldsymbol{\Sigma}_1)\det(\boldsymbol{\Sigma}_2)}}{\sqrt{\det(\boldsymbol{\Sigma})}} \exp(M_{12}), \tag{3.12a}$$

$$M_{12} = -\frac{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)}{8}, \tag{3.12b}$$

where $\boldsymbol{\Sigma} = \frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}$. Let $\mathcal{S}$, given in (3.13), be the set of rules whose Hellinger distance from the LCR is smaller than a given threshold γ.

$$\mathcal{S} = \{H(\Phi_i, \Phi_{\text{LCR}}) \,|\, H(\Phi_i, \Phi_{\text{LCR}}) < \gamma,\, i \in \mathbb{N}_{\leq C} \setminus \text{LCR}\}\,. \tag{3.13}$$

Whenever the cardinality of $\mathcal{S}$ is positive, namely $\text{card}\,(\mathcal{S}) > 0$, the LCR will be discarded while the MSR, given by $\arg\min_i \mathcal{S}$, will have its parameters updated by the data previously associated with the discarded rule, i.e., matrix $\mathbf{X}$ and vector $\mathbf{Y}$, using (3.6a)–(3.6c) and (3.7a)–(3.7d). The main steps of the learning algorithm EBeTS are summarized in Algorithm 1.

3.3 EBeTS parameters

EBeTS model introduces three parameters to control the model's complexity: ω, τ, and γ. The first parameter (ω) controls the size of the statistical bounds around the residual mean of the LCR as shown in (3.10). It is assumed that the residuals are Gaussian random variables; for this reason, knowing their distribution parameters, i.e., mean $\mu_{\mathbf{R}}$ and variance $\sigma^2_{\mathbf{R}}$, allows the definition of a region where $100\omega\%$ of non-anomalous data are within through the inverse chi-squared distribution. When a given data sample does not belong to this region, it is classified as an anomaly.

Classifying a data point as anomaly or non-anomaly is useful to decide whether to create a new rule or not. However, the rule creation process must be robust regarding

Algorithm 1 Learning algorithm of EBeTS.

procedure EBeTS(γ, τ, ω)
  CA $\leftarrow 0$ ▷ Number of continuous anomalies
  LCR $\leftarrow 1$ ▷ Last created rule
  C $\leftarrow 1$ ▷ Number of rules
  while $\mathbf{x}_k$ is available **do**
    Read $\mathbf{x}_k$ and estimate output $\hat{y}_k$ using (3.8)
    Read y_k
    Append $\mathbf{x}_k$ to $\mathbf{X}$ and y_k to $\mathbf{Y}$
    Adjust **LCR** consequent parameters via (3.7a)–(3.7d)
    Adjust **LCR** antecedent parameters via (3.6a)–(3.6c)
    Compute estimation residual using (3.9)
    if $|\mathbf{R}| > 2$ AND condition (3.10) holds **then** ▷ Anomaly classification
      CA $\leftarrow$ CA $+ 1$ ▷ Count continuous anomalies
    else
      CA $\leftarrow 0$
      Append estimation residual to $\mathbf{R}$
    end if
    if CA $> \tau$ **then**
      Compute $\mathcal{S}$ using (3.13)
      if card($\mathcal{S}$) > 0 **then**
        MSR $\leftarrow \arg\min_i \mathcal{S}$ ▷ Most similar rule
        for all $(\mathbf{x}_k, y_k) \in (\mathbf{X}, \mathbf{Y})$ **do**
          Adjust **MSR** consequent parameters using (3.7a)–(3.7d)
          Adjust **MSR** antecedent parameters using (3.6a)–(3.6c)
        end for
        C $\leftarrow$ C $- 1$
      end if
      Initialize consequent and antecedent parameters for a new rule
      Update new rule's parameters with last τ sequential anomalies
      CA $\leftarrow 0$
      C $\leftarrow C + 1$
      LCR $\leftarrow C$
      Clear vectors $\mathbf{R}$ and $\mathbf{Y}$ and matrix $\mathbf{X}$
    end if
    $k \leftarrow k + 1$
  end while
end procedure

outliers. The second parameter, the persistence index (τ), is introduced to control the rule creation by defining how many anomalies in sequence are allowed to happen before creating another rule. Parameters ω and τ are illustrated in Figure 3.1 where the region defined by $\mu_{\mathbf{R}}$, $\sigma^2_{\mathbf{R}}$ and ω is depicted in blue (solid and dashed lines) and anomalies are depicted as red dots; at the end of the time series a group of τ sequential anomalies is shown.

The last parameter (γ) defines a minimum Hellinger distance threshold for which merging can be considered for two clusters. The distances are computed for each pair of clusters where the closest pair is merged, provided the Hellinger distance between them is less than the threshold γ. Figure 3.2 depicts an example of the Hellinger distance between two bi-variate Gaussian distributions where four degrees of overlapping are shown.

The choice of each of the three EBeTS parameters described in this section will

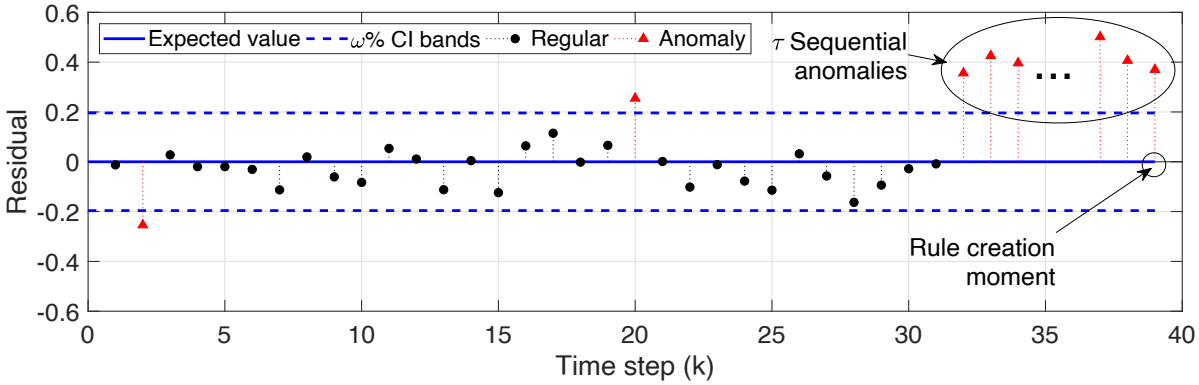


Figure 3.1 – Illustration of EBeTS parameters ω and τ for rule creation.

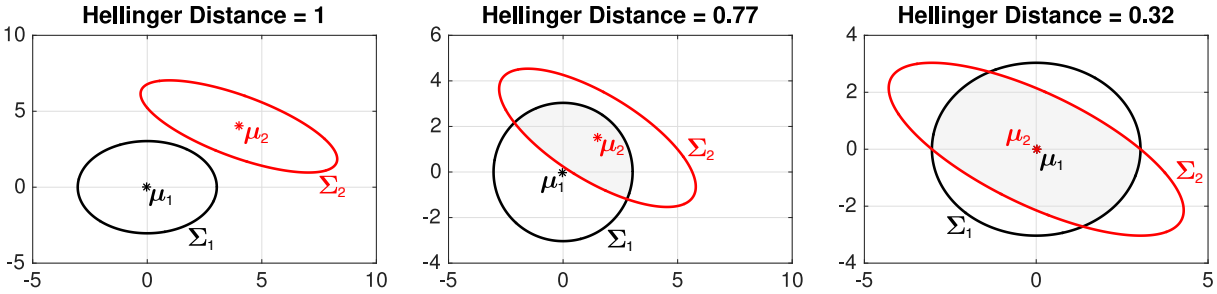


Figure 3.2 – Illustration of EBeTS parameter γ for rule merging.

reflect on the model complexity, i.e., the final number of rules in the model. As the parameter ω decreases, more data points will be considered anomalies, increasing the chance of creating new rules. When the parameter τ increases, more sequential anomalies will be required to create a new rule, decreasing the chance of creating new rules. In the same way, an increase in γ leads to a decrease in the number of rules as more rules become eligible to be merged. Therefore, a typical trade-off between the complexity and the expected model's performance arises from the parameter choices.

Although EBeTS parameters can be defined based on historical data, if available, their values can also be defined in the absence of any knowledge extracted from data. Considering anomalous behavior to be events unlikely to happen, statistically, these would be events in the lower and upper tails of the Gaussian distribution of the residuals. Therefore, a recommendation for ω under unavailable knowledge is a conservative value close to 100%, i.e., the definition of 2σ to 4σ bands around the mean prediction residual. The persistence index is closely related to the estimation of the n_x-dimensional inverse dispersion matrices in both antecedent and consequent parts of the rule, because these sequential anomalies are used to update a recently created rule; therefore, to decrease the chance of rank deficiency on these matrices, it is suggested to use $n_x + 1$ data points to estimate them, i.e., $\tau = n_x + 1$. The last parameter γ does not have a recommendation based on any statistical principle; empirically, in the absence of problem-related knowledge, a good choice would be $\gamma = 0.5$, since it is the midpoint of the interval $[0, 1]$, where the

Hellinger distance lies within.

3.4 Chapter summary

In this chapter, an evolving TS method for time-series prediction multiple steps ahead is developed. First, the problem is formulated using the TS fuzzy model, where the time-series behavior is modeled with a convex combination of local affine functions of the input variables. Then, the recursive structure adaptation process is detailed; the mean and dispersion matrix of the fuzzy antecedent MFs are updated as new data is available and the coefficients of the consequent affine functions are estimated with RLS. The proposed approach combines the idea of accounting for the complex relationships between the variables through multivariate Gaussian modeling with the knowledge-base control based on the estimation error tracking. The novel rule managing mechanisms based on a dynamic control chart and the Hellinger distance allows the suggestion of problem-agnostic parameters, increasing the EBeTS robustness.

4 RUL PREDICTION WITH EFS

A common concept on prognostics is the prediction of how much time remains before a failure occurs, also known as the RUL, given the current machine age, condition and the past operation profile (JARDINE; LIN; BANJEVIC, 2006). In this sense, the RUL can be defined as the amount of time until the system's degradation state reaches a predefined threshold, as shown in (2.3). In particular, data-driven prognostics rely on the assumption that there exist available data from run-to-failure experiments to train degradation models in an offline phase and that data from the UUT will also be available to aid the prognostics up to the time where prediction starts (t_p). It is not always the case, and the possibility of not having available historical data but only a few UUT data is explored in Section 5.2. For this reason, the prognostics problem may benefit from an evolving learning mechanism that can adapt its parameters and change its structure according to new available data while the fuzzy TS representation allows the modeling of nonlinear degradation behavior.

A general data-driven framework for using Evolving Fuzzy Models (EFMs) for prognostics tasks is shown in Figure 4.1. This is a two-phase framework composed by an offline stage and an online stage. The offline phase consists of training a TS fuzzy system using any available technique with historical data; this phase is nonessential considering EFMs can start make predictions from scratch while adapting both its parameters and structure in real-time. In the online phase, the EFM will be updated whenever a new observation becomes available, allowing the start of the prognostics task. The following subsections details how the degradation state can be propagated along with its uncertainty for prognostics purposes using EFMs.

4.1 Problem formulation

The RUL prediction problem can be stated as a multi-step ahead prediction of a system's degradation state until it reaches a threshold η defined by some physical knowledge of the assessed element. The state propagation can be done using a nonlinear state transition auto-regressive model, such that

$$\hat{x}_{k+N|k} = f_k\left(\mathbf{v}_{k+N,L}\right), \tag{4.1}$$

where $\mathbf{v}_{k+N,L}$, given by

$$\mathbf{v}_{k+N,L} = \begin{cases} \left[x_k \quad x_{k-1} \quad \cdots \quad x_{k-L+1}\right]^\top, & \text{if } N = 1 \\ \left[\hat{x}_{k+N-1} \quad \cdots \quad \hat{x}_{k+1} \quad x_k \quad \cdots \quad x_{k+N-L}\right]^\top, & \text{if } 2 \leq N \leq L \\ \left[\hat{x}_{k+N-1} \quad \cdots \quad \hat{x}_{k+N-L}\right]^\top, & \text{if } N > L \end{cases} \tag{4.2}$$

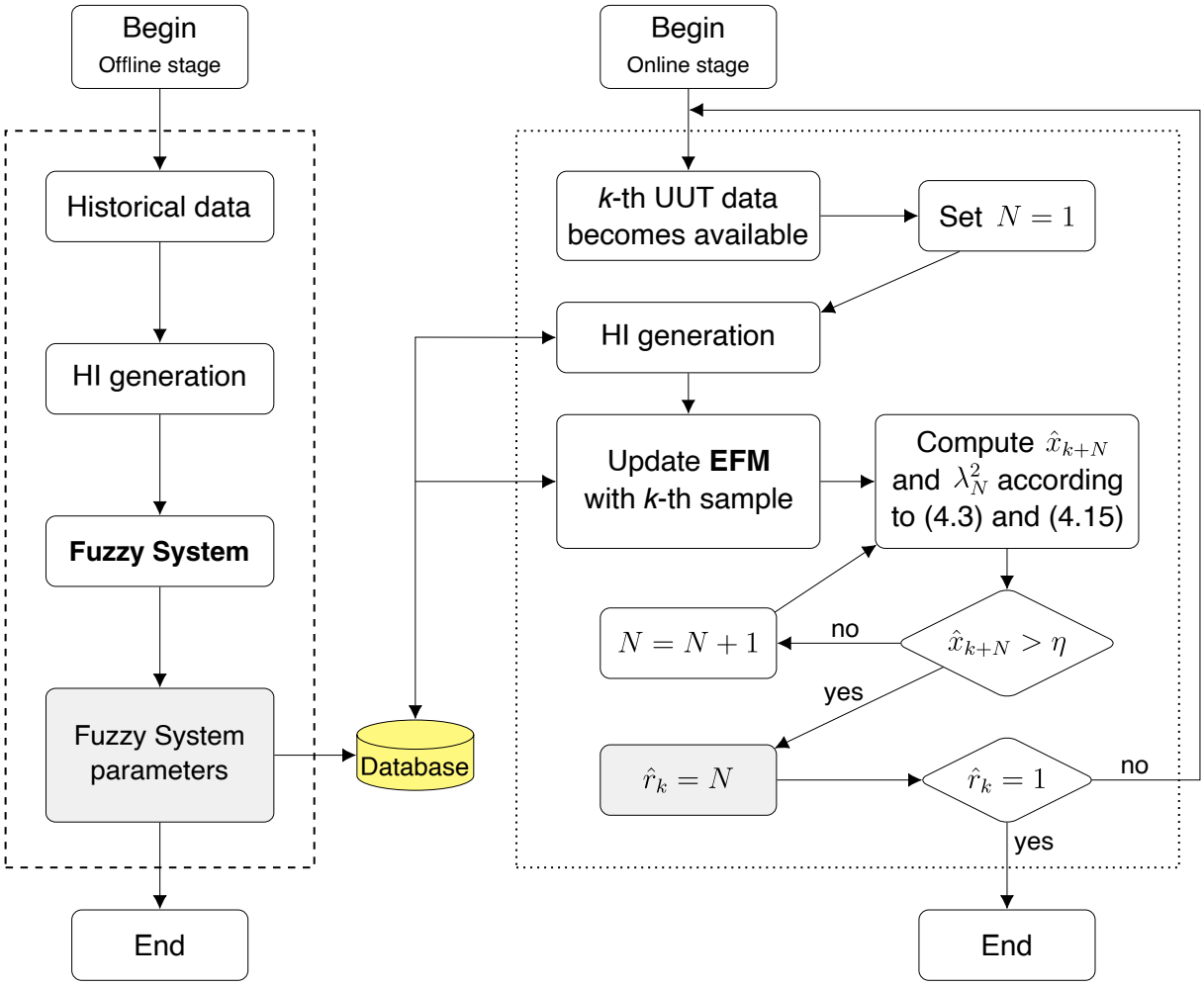


Figure 4.1 – Flowchart of the proposed data-driven framework for prognostics using EFMs.

is a lag vector with the estimates. Furthermore, x_n and $\hat{x}_n$ are, respectively, the observed and estimated degradation state at time instant n, L is the order of the auto-regression polynomial, N is the number of steps ahead for which the degradation state is predicted, and $f_k(\cdot)$ is the state transition function recursively obtained until instant k using EBeTS. The long-term prediction based on systems such as the one proposed in this work can be made in different ways. According to (GOURIVEAU; ZERHOUNI, 2012), there are five main approaches divided into single-output (Iterative, Direct, and DirRec) and multiple-output (Parallel and MISMO). The iterative approach consists of a single prediction tool to make one-step-ahead predictions and uses the last predicted value as a regressor to estimate the next ones; it has the most straightforward implementation and does not require the user to set the prediction horizon (N) beforehand. However, it suffers from error propagation and does not account for temporal behavior, i.e., once f_k is obtained, it will not be changed in the multi-step prediction phase. In the direct approach, H models are used to predict a single different horizon using the same observed data; since each model is independent of each other, they do not consider the complex relations between the variables and may hinder its implementation. The DirRec approach, on the other

hand, combines aspects of the previous two: it uses current predictions as regressors to the next steps, but in each step, a different model is used; however, it can also suffer from error propagation. In the multiple-output approaches, the parallel approach yields the best results in complexity and implementation simplicity; it uses a single model to predict multiple steps in less computing time, but it can raise rounding errors. The MISMO approach consists of several multiple-output models; it can be seen as an approach between the direct and the parallel approaches according to the definition of its parameters. In this work, the iterative approach is used as the mechanism of long-term prediction due to its simple implementation, low complexity, and unnecessary definition of a prediction horizon.

4.2 Uncertainty estimation

Considering a state transition function given by a TS model with rules given in (3.1) and multivariate Gaussian MFs given by (3.5), the degradation state propagation given in (4.1) can be rewritten as:

$$\hat{x}_{k+N} = \left(\mathbf{h}_k\left(\mathbf{v}_{k+N,L}\right)\right)^\top \hat{\mathbf{\Theta}}_k^\top \, \tilde{\mathbf{v}}_{k+N,L} + \epsilon_{k+N}, \quad \forall N > 0 \tag{4.3}$$

where $\mathbf{h}_k(\cdot)$ and $\hat{\mathbf{\Theta}}_k$ are the normalized degrees of activation and consequent parameters for each rule with structure updated until time instant k, respectively; N is the prediction horizon and $\tilde{\mathbf{v}}_{k+N,L}$ is the augmented vector $\tilde{\mathbf{v}}_{k+N,L} \triangleq \begin{bmatrix} 1 & \mathbf{v}_{k+N,L}^\top \end{bmatrix}^\top$. To account for prediction uncertainties, the following white Gaussian noise is added in model (4.3).

$$\epsilon_k \sim \mathcal{N}\left(0, \sigma_\epsilon^2\right). \tag{4.4}$$

where σ_ϵ^2 is considered to be constant. This noise variance will be estimated through Monte Carlo simulations using the consequent parameters' covariance matrix estimated via RLS until time instant k. Another way to compute it is to recursively track the covariance of estimation errors through the online learning operation. Both methods, described in the following subsections, are compared in Section 5.2.

4.2.1 Monte Carlo simulations

A possible way to estimate the variance σ_ϵ^2 in (4.4) is by Monte Carlo simulations using the known model's parameters up until time instant k. The method consists in generating M observations of the parameters and estimate the variance as

$$\hat{\sigma}_\epsilon^2 \approx \frac{1}{M}\sum_{j=1}^{M}\left(x_k - \sum_{i=1}^{C} h_{i,k}(\mathbf{z}_k)\, \tilde{\mathbf{v}}_{k,L}^\top \boldsymbol{\vartheta}_{i,k}\right)^2, \tag{4.5}$$

where $\boldsymbol{\vartheta}_{i,k} \sim \mathcal{N}\left(\hat{\boldsymbol{\theta}}_{i,k}, \mathbf{F}_{i,k}\right)$ and as $M \to \infty$, $\hat{\sigma}_\epsilon^2 \to \sigma_\epsilon^2$. According to (LJUNG, 1999, p. 368), under the assumption of white Gaussian noise, the posterior distribution of the consequent parameters at an instant k is Gaussian with mean $\hat{\boldsymbol{\theta}}_{i,k}$ and covariance matrix $\mathbf{F}_{i,k}$, whose recursion relations are given in (3.7c) and (3.7d), respectively.

4.2.2 Error covariance online tracking

Another possibility is to keep track of the model's error covariance matrix for time instances $n \in \mathbb{N}_{\leq k}$. The error mean can be tracked in a similar way of (3.6a) and (3.6b):

$$\boldsymbol{\Delta}_{\epsilon,n} = \boldsymbol{\epsilon}_n - \hat{\boldsymbol{\mu}}_{\epsilon,n-1}, \tag{4.6}$$

$$\hat{\boldsymbol{\mu}}_{\epsilon,n} = \hat{\boldsymbol{\mu}}_{\epsilon,n-1} + \frac{1}{n}\boldsymbol{\Delta}_{\epsilon,n}, \tag{4.7}$$

where n is the global amount of inputs processed by EBeTS. The error mean starting point is $\hat{\boldsymbol{\mu}}_{\epsilon,0} = \mathbf{0}_{n_y \times 1}$. Once the error mean is estimated, the recursive sum of squares from it can also be computed:

$$\mathbf{M}_{\epsilon,n} = \mathbf{M}_{\epsilon,n-1} + (\boldsymbol{\epsilon}_n - \hat{\boldsymbol{\mu}}_{\epsilon,n-1})(\boldsymbol{\epsilon}_n - \hat{\boldsymbol{\mu}}_{\epsilon,n})^\top, \tag{4.8}$$

where $\mathbf{M}_{\epsilon,0} = \mathbf{0}_{n_y \times n_y}$. Then, at time instant n, the error covariance matrix is given as

$$\boldsymbol{\Sigma}_{\epsilon,n} = \frac{\mathbf{M}_{\epsilon,n}}{n-1}. \tag{4.9}$$

The constant variance σ_ϵ^2 in (4.4) used for the long-term prediction can be estimated as

$$\sigma_\epsilon^2 \approx \boldsymbol{\Sigma}_{\epsilon,k}. \tag{4.10}$$

4.3 Uncertainty propagation

After computing the initial uncertainty in the estimation of one step ahead predictions, the long term propagation of that uncertainty is made by considering the input vector (4.2) to be a vector composed of random variables, such that

$$\mathbf{v}^+_{k+N,L} = \begin{cases} \begin{bmatrix} x_k & x_{k-1} & \cdots & x_{k-L+1} \end{bmatrix}^\top, & \text{if } N = 1 \\ \begin{bmatrix} \hat{x}^+_{k+N-1} & \cdots & \hat{x}^+_{k+1} & x_k & \cdots & x_{k+N-L} \end{bmatrix}^\top, & \text{if } 2 \leq N \leq L \\ \begin{bmatrix} \hat{x}^+_{k+N-1} & \cdots & \hat{x}^+_{k+N-L} \end{bmatrix}^\top, & \text{if } N > L \end{cases} \tag{4.11}$$

It is worth noting that, when $N = 1$, the previous degradation states are known and, therefore, are not random variables. Accordingly, the output $\hat{x}_{k+N}$ of the state transition function in (4.3) is also a random variable:

$$\hat{x}^+_{k+N} = (\mathbf{h}_k(\mathbf{z}_{k+N}))^\top \hat{\boldsymbol{\Theta}}_k^\top \tilde{\mathbf{v}}^+_{k+N,L} + \epsilon_{k+N}, \quad \forall N > 0 \tag{4.12}$$

where $\mathbf{z}_{k+N} \triangleq \mathbb{E}[\tilde{\mathbf{v}}_{k+N,L}]$ are premise variables defined as the expected value of the input vector. For the uncertainty propagation, it is necessary to compute its variance in the multi-step prediction framework. For the first step, the variance can be computed as

$$\begin{aligned} \text{Var}\left(\hat{x}^+_{k+1}\right) &= \text{Cov}\left((\mathbf{h}_k(\mathbf{z}_{k+1}))^\top \hat{\boldsymbol{\Theta}}_k^\top \tilde{\mathbf{v}}^+_{k+1,L} + \epsilon_{k+1}\right) \\ &= \text{Cov}\left((\mathbf{h}_k(\mathbf{z}_{k+1}))^\top \hat{\boldsymbol{\Theta}}_k^\top \tilde{\mathbf{v}}^+_{k+1,L}\right) + \sigma_\epsilon^2. \end{aligned} \tag{4.13}$$

Since the degrees of activation $\mathbf{h}_k(\cdot)$ are computed with the expected value of the random variable $\tilde{\mathbf{v}}^+_{k+1,L}$, it can be considered, along with the parameters vector, as a constant defined as: $\boldsymbol{\Xi}_N \triangleq (\mathbf{h}_k(\mathbf{z}_{k+N}))^\top \hat{\boldsymbol{\Theta}}_k^\top$. Therefore, using the results from Appendix B, the variance (4.13) becomes:

$$\begin{aligned}\text{Var}\left(\hat{x}^+_{k+1}\right) &= \boldsymbol{\Xi}_1 \ \text{Cov}\left(\tilde{\mathbf{v}}^+_{k+1,L}\right) \ \boldsymbol{\Xi}_1^\top + \sigma_\epsilon^2 \\ &= \boldsymbol{\Xi}_1 \ \boldsymbol{\Lambda}_1^L \ \boldsymbol{\Xi}_1^\top + \sigma_\epsilon^2 \\ &= \sigma_\epsilon^2 \\ &= \lambda_1^2, \end{aligned} \tag{4.14}$$

where $\boldsymbol{\Lambda}_N^L \triangleq \text{Cov}\left(\tilde{\mathbf{v}}^+_{k+N,L}\right)$ and $\lambda_N^2 \triangleq \text{Var}\left(\hat{x}^+_{k+N}\right)$. It is worth noting that $\boldsymbol{\Lambda}_1^L = 0$ because, when $N = 1$, the previous degradation states are known. Simirlarly, the variance N steps ahead is computed recursively as

$$\begin{aligned}\text{Var}\left(\hat{x}^+_{k+N}\right) &= \text{Cov}\left((\mathbf{h}_k(\mathbf{z}_{k+N}))^\top \hat{\boldsymbol{\Theta}}_k^\top \ \tilde{\mathbf{v}}^+_{k+N,L} + \epsilon_{k+N}\right) \\ &= \boldsymbol{\Xi}_N \ \text{Cov}\left(\tilde{\mathbf{v}}^+_{k+N,L}\right) \ \boldsymbol{\Xi}_N^\top + \sigma_\epsilon^2 \\ &= \boldsymbol{\Xi}_N \ \boldsymbol{\Lambda}_N^L \ \boldsymbol{\Xi}_N^\top + \sigma_\epsilon^2. \end{aligned} \tag{4.15}$$

The covariance matrix of the random vector $\tilde{\mathbf{v}}^+_{k+1,L}$ is given as

$$\boldsymbol{\Lambda}_N^L = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & \lambda_{N-1}^2 & \cdots & \lambda_{N-L}\lambda_{N-1}\hat{\rho}_{L,1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \lambda_{N-1}\lambda_{N-L}\hat{\rho}_{1,L} & \cdots & \lambda_{N-L}^2 \end{bmatrix}, \tag{4.16}$$

where the first row and column represent the covariance between the random variables with the augmented one. Moreover, $\lambda_i^2 = 0$ when $i < 0$ meaning that the value of x_{k+N} is known. The covariance matrix shown in (4.16) is weighted by Pearson correlation coefficients $\hat{\rho}$, estimated through available UUT data or recursively estimated, as shown in Appendix A.2.

Considering the degradation to be a random variable with Gaussian distribution whose expected value is propagated by successive applications of (4.3), it is possible to suggest lower and upper bounds for the RUL at a $(\alpha)(100)\%$ significance level as

$$\hat{r}_{\text{lower},k} = \inf\{N \in \mathbb{N} : \hat{x}_{k+N} + z_{1-\alpha/2}\,\lambda_N \geq \eta\}, \tag{4.17a}$$

$$\hat{r}_{\text{upper},k} = \inf\{N \in \mathbb{N} : \hat{x}_{k+N} + z_{\alpha/2}\,\lambda_N \geq \eta\}, \tag{4.17b}$$

To provide useful information for decision-making in practical engineering applications, it is important to represent, quantify, propagate and manage uncertainty (SANKARARAMAN,

2015). However, there is still a lack of effective uncertainty quantification for multi-step ahead prediction based on EFMs. In this sense, despite the simplicity, the uncertainty quantification methodology described in this section is a novelty in the context of EFMs and enabled the fault prognostics for this kind of model.

4.4 Chapter summary

This chapter presents the uncertainty quantification and propagation procedures used for evolving TS models, such as EBeTS. The quantified uncertainty are related to model uncertainties, which account for the proposed model's estimation error from Guo, Li & Li (2019). The proposed quantification technique can use the posterior distribution of the consequent parameters of each rule noise or the global error covariance matrix to estimate the model initial uncertainty that will be propagated in the long-term predictions. The uncertainty quantification and propagation procedures enables the definition of a data-driven framework for prognostics using EFMs, as shown in Figure 4.1.

5 RESULTS AND DISCUSSION

5.1 Case study 1: temperature prediction

In order to illustrate the ability of the proposed method for system identification, three time series composed by monthly recorded mean temperatures from Death Valley (Furnace Creek), Ottawa, and Lisbon weather stations are used. The data set consists of records from January of 1901, 1895 and 1910, respectively, up to December of 2009. Systems sensitive to the state of the atmosphere may benefit from weather predictions (LEITE *et al.*, 2012). These predictions can help the decision-making process in many sectors, e.g., agriculture, aviation, and transportation. In all experiments, the data were scaled to the interval $[0, 1]$ using their maximum and minimum temperatures in degrees Celsius.

The proposed EBeTS model is compared with two fuzzy evolving models, the Fuzzy Set Based Evolving Modeling (FBeM) (LEITE *et al.*, 2012) and the exTS (ANGELOV; ZHOU, 2006); these models can also represent time variant nonlinear systems. As discussed in Chapter 2, both learning mechanisms are similar to the proposed technique in terms of the main model structure, which is based on a TS fuzzy inference system. Apart from the learning mechanisms, another difference between these two models and EBeTS is that they do not account for the correlation between the input variables, which is achieved in EBeTS through multivariate MFs in the antecedent part of its rules.

The experiment consists of predicting the temperature one, three, six, and nine months ahead using the last five observations. The objective of this design is to check the prediction capabilities of EBeTS along with its robustness under situations where the outputs are less correlated with the input, e.g., in a monthly recording temperature data set, x_{k+1} is more correlated to x_k than x_{k+6} is to x_k. The input and output provided to the three algorithms are, respectively,

$$\mathbf{x}_k = \begin{bmatrix} x_k & x_{k-1} & \dots & x_{k-4} \end{bmatrix}^\top, \tag{5.1a}$$

$$\begin{bmatrix} x_{k+1} \\ x_{k+3} \\ x_{k+6} \\ x_{k+9} \end{bmatrix}^\top = \begin{bmatrix} f_1(\mathbf{x}_k) \\ f_2(\mathbf{x}_k) \\ f_3(\mathbf{x}_k) \\ f_4(\mathbf{x}_k) \end{bmatrix}^\top. \tag{5.1b}$$

The tuning parameters of EBeTS were set as $\gamma = 0.5$, $\tau = 6$ and $\omega = 95.45\%$, using the recommendations detailed in Section 3.3; along with the recommendation, a sensitivity analysis is made in Subsection 5.1.1 to support the choice. The parameters chosen for FBeM were the ones reported in Leite *et al.* (2012), namely $\rho_{\text{FBeM}} = 0.7$, $\psi_{\text{FBeM}} = 48$

and $\eta_{\text{FBeM}} = 2$. The only parameter set for exTS is the rule covariance initialization constant Ω_{exTS}. It is analogous to the parameter $\mathbf{F}_{i,0}$ of EBeTS, therefore it is also set as $\Omega_{\text{exTS}} = 10^3$.

In the specific case of one-step ahead temperature prediction in the Death Valley region, Figure 5.1a depict the true and predicted values from Nov/1963 to Jan/1968. The prediction error, number of rules and membership degree for each rule from Jan/1901 to Dec/2009 are shown in Figure 5.1b, Figure 5.1c, and Figure 5.1d, respectively.

The results are shown in Table 5.1, where a comparison is made in terms of the Non Dimensional Error Index (NDEI), Theil's U statistic and computing time. In order to illustrate the knowledge-base construction, the average number of rules is also depicted. The NDEI metric consists of the known Root Mean Squared Error (RMSE) divided by the standard deviation of the target time series, defined as:

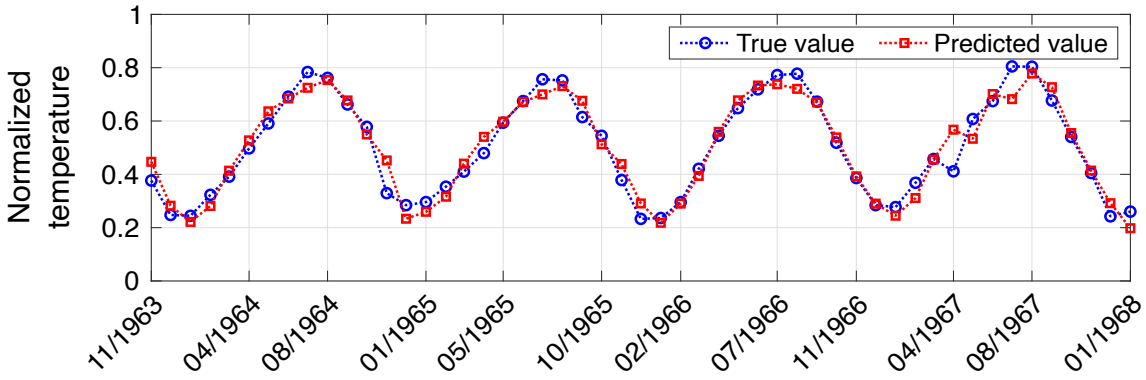
$$\text{NDEI} = \frac{\sqrt{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}}{\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \frac{1}{N}\sum_{j=1}^{N} x_j\right)^2}}. \tag{5.2}$$

The Theil's U statistic, on the other hand, is a metric to check the predictions in relation to a naïve predictor where $\hat{y}_{i+1} = y_i$. According to Makridakis, Wheelwright & Hyndman (1998, p. 50) if the value of U in (5.3) is less than 1, the forecast is better than the naïve method. It will be 1 when they are the same and greater than 1 when the forecast is worse than naïve predictions:

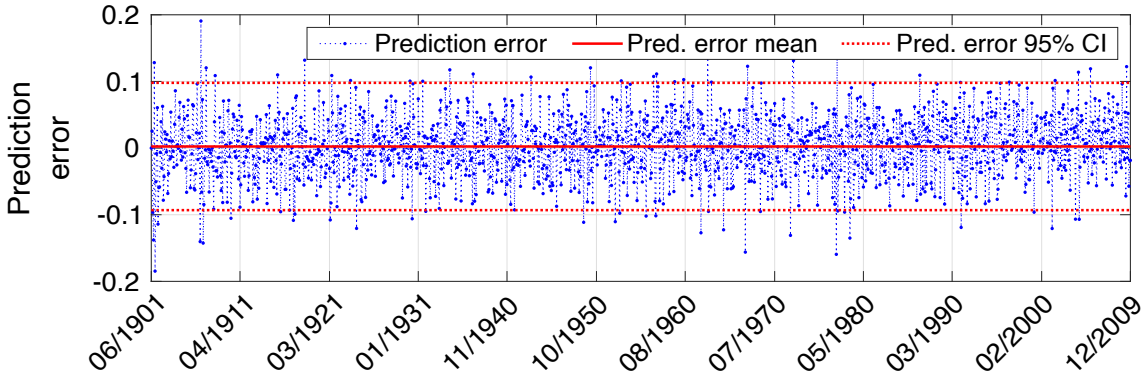
$$\text{U} = \sqrt{\frac{\sum_{i=1}^{N-1}\left(\frac{\hat{x}_{i+1} - x_{i+1}}{x_i}\right)^2}{\sum_{i=1}^{N-1}\left(\frac{x_{i+1} - x_i}{x_i}\right)^2}}. \tag{5.3}$$

The proposed EBeTS can maintain its accuracy with predictions better than naïve forecasts, by controlling the model complexity with a criterion designed to reduce the estimation error. This complexity is reflected by the average number of rules which is a widely used metric to account for complexity-based interpretability. This is one of the two types of interpretability defined in Gacto, Alcalá & Herrera (2011); the other type is called semantics-based interpretability and is not addressed in this work. The low prediction errors of both EBeTS and exTS in this case study indicate their suitability for prognostics, assuming that the long-term predictions required for RUL estimation can be done with these predictors.

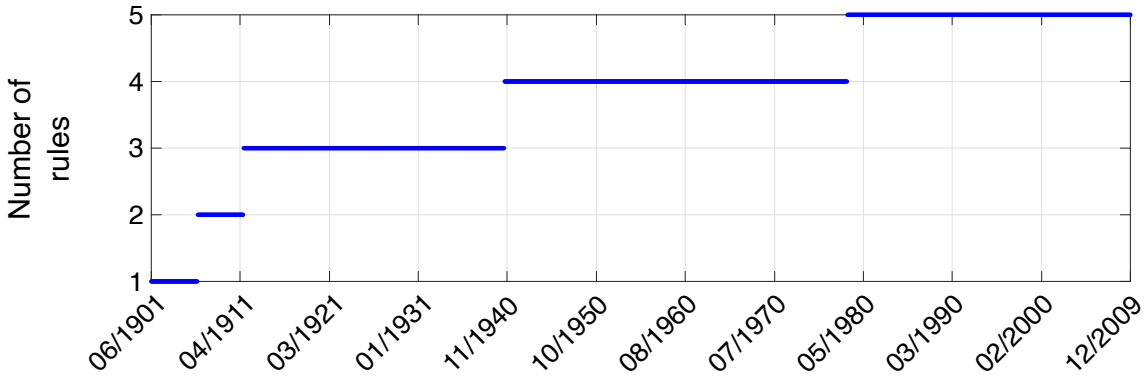
An illustration of EBeTS process of predicting the temperature of the three weather stations is shown in Figure 5.2. Its first row shows the true and estimated output as well as the second row for a smaller time span, indicated in the first row as a grayish stripe. The third row shows both NDEI and the number of rules over time, where it is possible to see the number of rules changing as NDEI changes. It is important to notice that the redundant information of axis labels and title is omitted to improve readability.



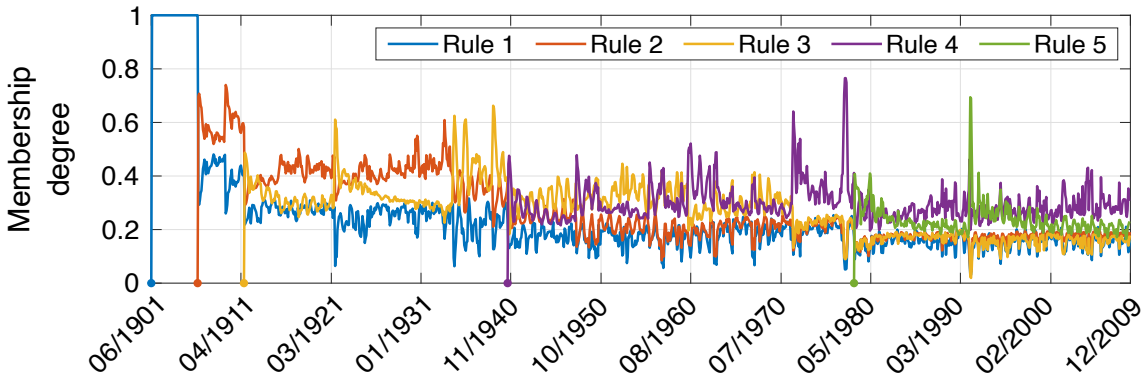
(a) Death Valley's true and predicted temperature from Nov/1963 to Jan/1968.



(b) Death Valley's temperature prediction error from Jan/1901 to Dec/2009.



(c) Death Valley's temperature total number of rules from Jan/1901 to Dec/2009.



(d) Death Valley's temperature membership degree from Jan/1901 to Dec/2009.

Figure 5.1 – Death Valley's temperature one-step ahead prediction results using EBeTS.

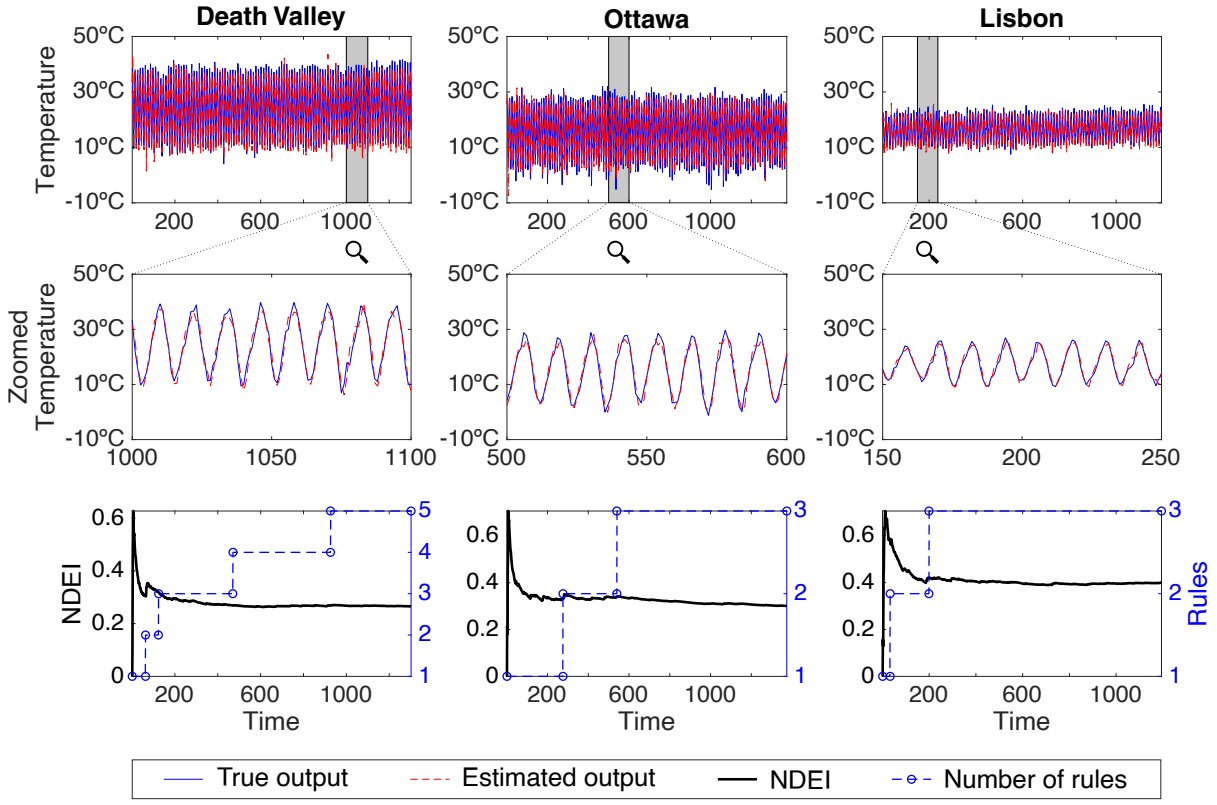


Figure 5.2 – EBeTS prediction of the Death Valley, Ottawa, and Lisbon temperature time series with NDEI and number of rules over time.

5.1.1 Sensitivity analysis

In this case study, the parameters used for EBeTS are chosen based on recommendations detailed in Section 3.3, where a discussion on how the parameters affect the model complexity is also provided. A sensitivity analysis is done to provide a more detailed insight into the parameters' influence in both complexity and accuracy. Both NDEI and the total number of rules are computed in an experiment with the following parameters range:

$$(\omega, \gamma, \tau) \in \mathcal{O}_s \times \mathcal{G}_s \times \mathcal{T}_s, \tag{5.4}$$

where

$$\mathcal{O}_s = \{38.29\%, 68.27\%, 86.64\%, 95.45\%, 98.76\%, 99.73\%, 99.95\%\}, \tag{5.5a}$$

$$\mathcal{G}_s = \{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8\}, \tag{5.5b}$$

$$\mathcal{T}_s = \{3, 4, 5, 6, 7, 8, 9\}. \tag{5.5c}$$

The analysis is performed in a version of the temperature data set prepared for one step ahead predictions, where samples from the three stations are concatenated as depicted in Figure 5.3. The results for each set of parameters from (5.4) are summarized

Table 5.1 – Performance of temperature prediction on Death Valley, Lisbon and Ottawa weather stations using EBeTS, exTS (ANGELOV; ZHOU, 2006) and FBeM (LEITE *et al.*, 2012), with best values in bold.

Metric	Location	Method	Prediction N steps ahead			
			$N = 1$	$N = 3$	$N = 6$	$N = 9$
NDEI	Death Valley	EBeTS	0.2859	0.3268	0.3435	**0.2220**
		FBeM	0.3780	0.3777	0.3369	0.3346
		exTS	**0.2814**	**0.2851**	**0.3040**	0.2315
	Lisbon	EBeTS	0.4116	0.4702	0.5270	**0.3422**
		FBeM	0.5124	0.5016	0.4952	0.4508
		exTS	**0.3933**	**0.4627**	**0.4289**	0.3689
	Ottawa	EBeTS	**0.3198**	**0.3294**	0.3176	**0.2801**
		FBeM	0.3547	0.3603	0.3589	0.3566
		exTS	0.3479	0.3479	**0.3084**	0.3021
Avg. Rules	Death Valley	EBeTS	**3.7857**	**3.9562**	5.0617	**3.5927**
		FBeM	11.2842	11.4762	11.2375	10.9784
		exTS	4.7926	5.6346	**2.9800**	3.8694
	Lisbon	EBeTS	**2.8065**	**5.3951**	**3.6140**	**3.6830**
		FBeM	10.4238	9.9320	9.8276	11.0868
		exTS	4.6809	6.9639	3.7511	6.6324
	Ottawa	EBeTS	**2.4068**	**2.9978**	**2.8795**	**5.7950**
		FBeM	10.1477	9.0882	9.7034	9.1069
		exTS	9.1689	6.8178	9.7093	7.0842
Theil's U	Death Valley	EBeTS	0.5267	0.6138	0.6748	**0.4408**
		FBeM	0.6336	0.6840	0.6086	0.6358
		exTS	**0.4558**	**0.4696**	**0.5881**	0.4548
	Lisbon	EBeTS	0.7175	0.7632	0.8774	**0.5965**
		FBeM	0.8336	0.7852	0.7929	0.7525
		exTS	**0.6710**	**0.7336**	**0.6714**	0.6389
	Ottawa	EBeTS	**0.5604**	**0.5431**	0.5501	**0.5044**
		FBeM	0.5974	0.6690	0.6105	0.6634
		exTS	0.6146	0.6214	**0.5368**	0.5339
Time (s)	Death Valley	EBeTS	**0.2876**	0.2462	0.2438	0.2432
		FBeM	0.3505	0.2916	0.3378	0.3174
		exTS	0.3330	**0.1596**	**0.1572**	**0.1507**
	Lisbon	EBeTS	0.1826	0.2592	0.2062	**0.2053**
		FBeM	0.2402	**0.2364**	0.2606	0.2331
		exTS	**0.1513**	0.2398	**0.1390**	0.2463
	Ottawa	EBeTS	**0.2015**	0.2698	**0.2193**	0.3273
		FBeM	0.4284	0.4506	0.3004	0.3263
		exTS	0.3517	**0.2602**	0.4203	**0.2528**

in Figure 5.4, where pairwise heat maps colored by the average number of rules and the average NDEI are shown. In the first column of Figure 5.4, the discussion from Section 3.3 becomes evident: the lower ω is, or the higher τ and γ are, the lower the average number of rules will be. The second column shows how the NDEI behaves as each pair of parameters changes. According to Figure 5.4, the choice of parameters $(\tau, \omega, \gamma) = (6, 95.45\%, 0.5)$, made without any knowledge about the data, is shown to be within regions where both the number of rules and the value of NDEI are low. A similar analysis was done considering $\omega = 95.45\%$, as shown in Table 5.2. When τ increases, EBeTS tends to create less rules and even large changes in γ will not affect its results in terms of NDEI and the final number of rules.

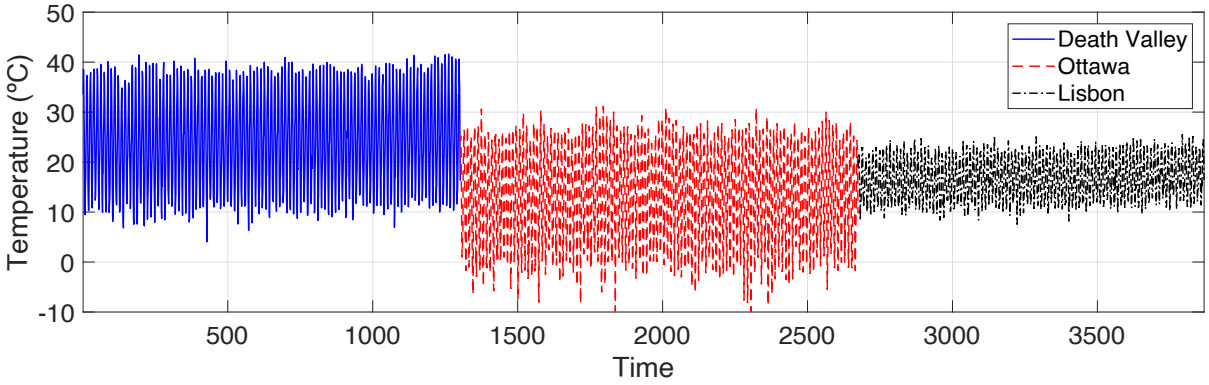


Figure 5.3 – Temperature data set from three different weather stations concatenated.

Table 5.2 – Sensitivity results NDEI/Final number of rules for parameters γ and τ with $\omega = 95.45\%$.

γ	τ						
	3	4	5	6	7	8	9
0.2	0.3467/61	0.3245/47	0.3187/14	0.3240/6	0.2855/5	0.2902/2	0.2892/2
0.3	0.3474/57	0.3288/38	0.3228/13	0.3240/6	0.2829/4	0.2902/2	0.2892/2
0.4	0.3308/46	0.3258/34	0.3185/11	0.3240/6	0.2829/4	0.2902/2	0.2892/2
0.5	0.3374/39	0.3343/19	0.3143/8	0.2868/4	0.2829/4	0.2902/2	0.2892/2
0.6	0.3352/32	0.3366/17	0.3109/7	0.2868/4	0.2822/3	0.2902/2	0.2892/2
0.7	0.3357/25	0.3362/11	0.3021/4	0.2865/3	0.2805/3	0.2902/2	0.2892/2
0.8	0.3550/14	0.3495/8	0.2961/3	0.2872/2	0.2819/2	0.2902/2	0.2892/2

5.2 Case study 2: accelerated ball bearing prognostics

In order to assess RUL prediction using models generated by EBeTS, two experiments with the PRONOSTIA platform data are proposed. The first one aims at comparing the results of the EBeTS based prognostics with a method based on multiple models reported in Cosme *et al.* (2019). The second experiment consists in the comparison of three distinct scenarios for RUL prediction using EBeTS, as discussed in Subsection 5.2.5.

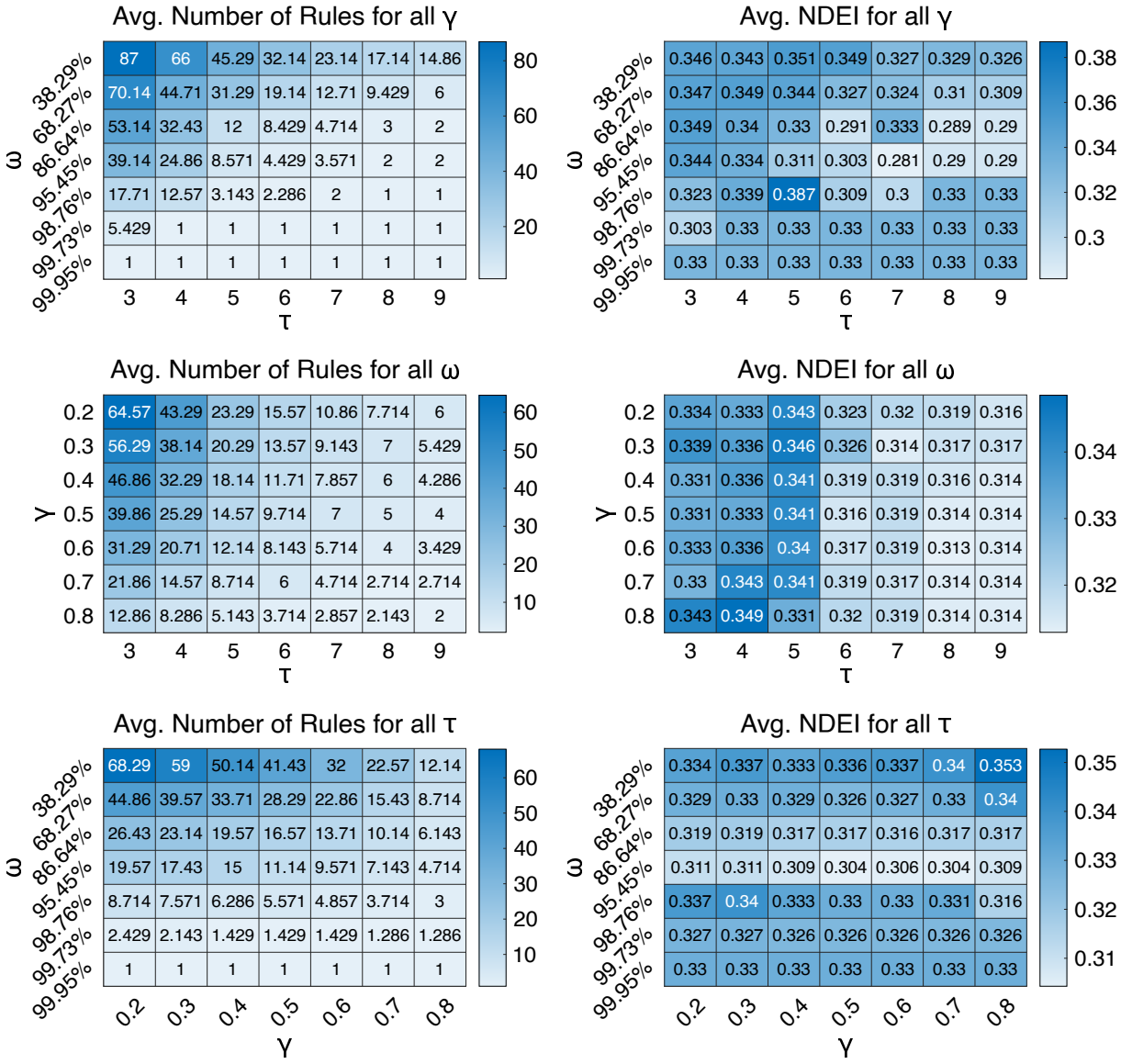


Figure 5.4 – Heat maps colored by the average number of rules and the average NDEI for each pair of EBeTS parameters for the sensitivity analysis of the first case study.

In order to prevent numerical problems due to RLS initialization, only rules with more than one data point associated to it were use for multi-step ahead predictions.

5.2.1 The PRONOSTIA data set

The PRONOSTIA data set consists of accelerated degradation of ball bearings under different operation conditions during their whole operational life (NECTOUX *et al.*, 2012). The testbed, shown in Figure 5.5, has three parts: (i) rotating part, composed by an asynchronous motor that allows the bearing to rotate through a system of gearing and different couplings; (ii) degradation generation part, composed by a pneumatic jack applying a radial force that reduces the bearing's life duration; (iii) measurements part to obtain instantaneous measurements from the radial force applied on the bearing, the

rotation speed of the shaft handling the bearing and the torque inflicted to the bearing. There are three different operation conditions in the data set; the differences between them are related to the horizontal force applied on the bearings and their rotating speeds, as shown in Table 5.3. The influence of these variables' variations are not considered for RUL prediction since they were kept almost stable during the whole operation. The available signals are the vibration in horizontal and vertical direction as well as the temperature of the bearings with a sampling frequency of 25.6 kHz for the vibration signals and of 10 Hz for the temperatures ignals. For each 10 seconds of operation a batch of 2560 samples, or 0.1 seconds, becomes available and is used to generate the HI used for RUL prediction.

The PRONOSTIA data set is widely used to investigate various aspects of prognostics algorithms. The HI construction problem include extraction of time-frequency domain features (SINGLETON; STRANGAS; AVIYENTE, 2015; ZHAO; TANG; TAN, 2016) and correlation-based features (LEI *et al.*, 2016b; MEDJAHER; ZERHOUNI; BAKLOUTI, 2013). Moreover, the RUL is predicted using different techniques such as Wiener processes (HUANG *et al.*, 2017; WANG *et al.*, 2016), neuro-fuzzy networks (PAN *et al.*, 2014), and support vector regression (BENKEDJOUH *et al.*, 2013), to name a few.

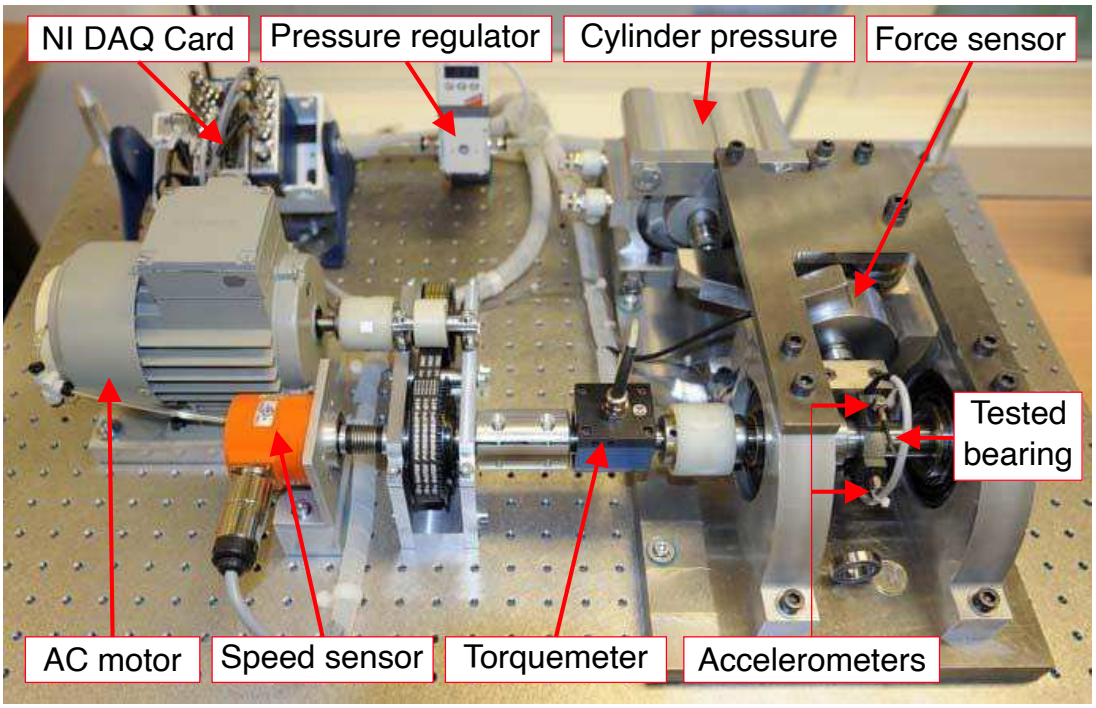


Figure 5.5 – Overview of PRONOSTIA testbed.

Source: Lei, Li & Lin (2016)

According to Nectoux *et al.* (2012), there are two main reasons for using data-driven techniques to tackle PRONOSTIA: the first one is that nothing is known about the degradation nature and origin; the other reason is that there is a mismatch between the experiments and the theoretical framework such as L_{10} life, Ball Pass Frequency of Inner Ring, and Ball Pass Frequency of Outer Ring.

Table 5.3 – PRONOSTIA training and test data sets under different operation conditions.

Condition #	Operation conditions	Training/validation sets	Test sets
1	1800rpm / 4000N	Bear_{1-1} / Bear_{1-2}	Bear_{1-3} to Bear_{1-7}
2	1650rpm / 4200N	Bear_{2-1} / Bear_{2-2}	Bear_{2-3} to Bear_{2-7}
3	1500rpm / 5000N	Bear_{3-1} / Bear_{3-2}	Bear_{3-3}

The first step towards a data-driven approach in prognostics is to extract candidate features for a useful HI. In this work, these features are derived from vibration sensors in the same way described in Javed *et al.* (2015) where the horizontal vibration signals were chosen; since the force is applied in this direction, they convey more information on the degradation than the others. To construct the HI proposed in Javed *et al.* (2015), the following steps must be taken:

Step 1. wavelet transform using the fourth level of decomposition of fourth order Daubechies in each 0.1 s packet;

Step 2. application of the arc tangent trigonometric function in the packet;

Step 3. statistical reduction of the packet by the standard deviation function;

Step 4. application of a smoothing task, in this case an exponential moving average with a window of size 12;

Step 5. computation of cumulative sum for improving trendability and monotonicity.

The steps are also depicted in Figure 5.6, where Step 0 represents the raw data acquisition.

5.2.2 Metrics used for comparison

Two accuracy based metrics are used to evaluate the proposed method performance: the MAPE and RA. They are computed as follows (SAXENA *et al.*, 2008a):

$$\text{MAPE}_k = \frac{100}{N} \sum_{i=k+1}^{k+N} \left| \frac{x_i - \hat{x}_i}{x_i} \right|, \tag{5.6}$$

$$\text{RA}_k = 1 - \frac{|r_k - \hat{r}_k|}{r_k}, \tag{5.7}$$

where N is the number of future predictions until the UUT state reaches the threshold; r_k and $\hat{r}_k$ are, respectively, the actual and estimated RUL at k. For simplicity purposes, polynomials of degree one were used, which means that the number of autoregressive components is one in (4.3).

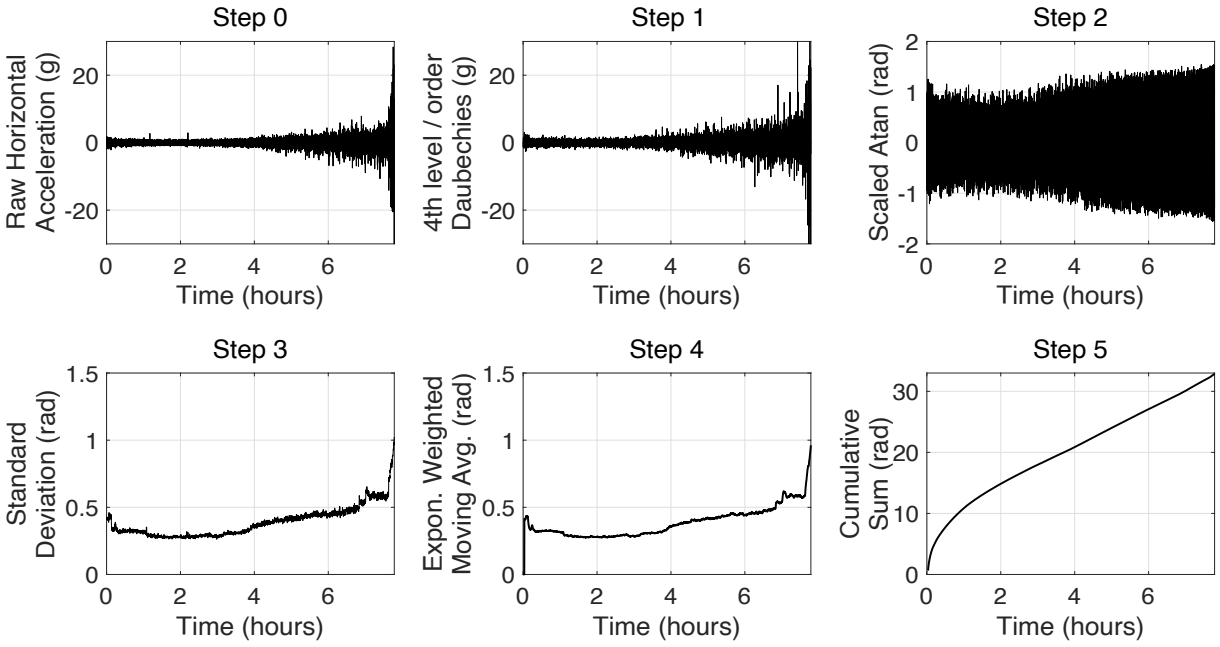


Figure 5.6 – Steps to generate the HI proposed by Javed *et al.* (2015).

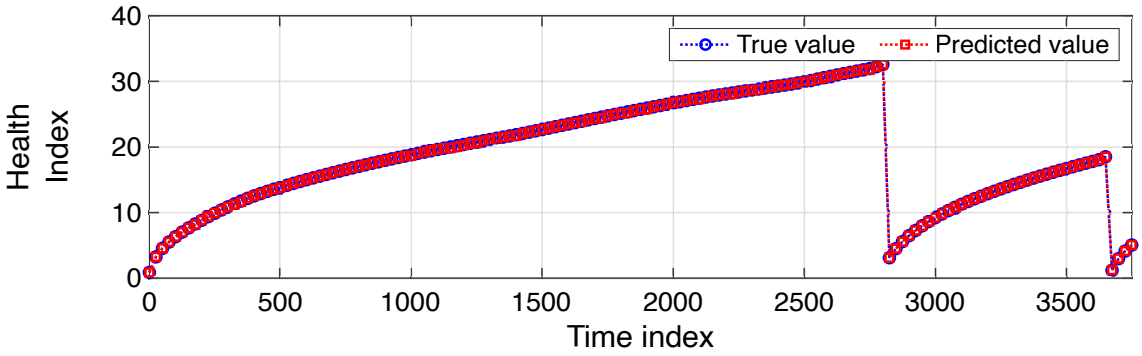
5.2.3 EBeTS parameters tuning

Since there are available historical data to train and validate the algorithms, the sets Bear_{i-1} and Bear_{i-2} of the i-th operation condition were used to define the EBeTS parameters. A grid search was used to find the best set of parameters with the objective of maximizing the mean prognostics RA in the training data set. For parameter γ, the grid interval was set to $\{0.1, 0.2, \ldots, 0.9\}$; for parameter τ, the grid interval was set to $\{1, 2, \ldots, 5\}$; and for parameter ω, the grid interval was set to $\{68.27\%, 86.64\%, \ldots, 99.99\%\}$. The values for ω are related to the band size around the residuals mean, i.e., $\{1\sigma, 1.5\sigma, 2\sigma, \ldots, 4\sigma\}$. The best set of parameters after the search was $\gamma = 0.8$, $\tau = 2$ and $\omega = 68.27\%$.

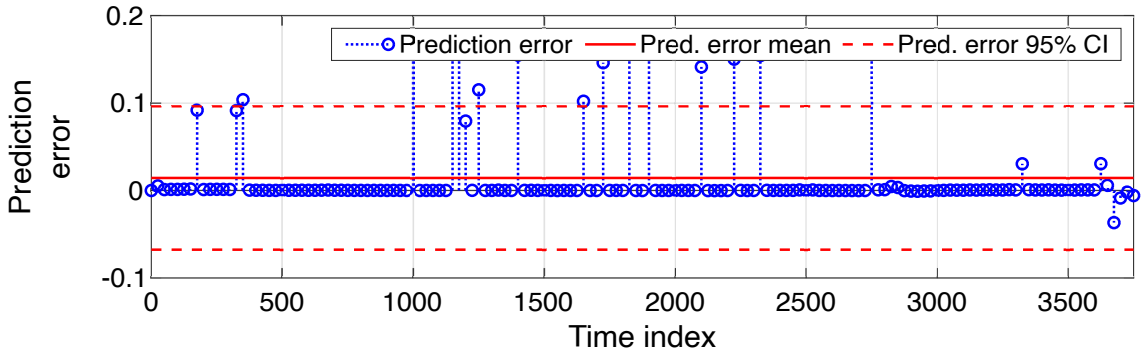
In the specific case of one-step ahead prediction of bearing health condition for condition 1, Figure 5.7a depict the true and predicted values for training and validation datasets Bear_{1-1} and Bear_{1-2} with the beginning of the testing dataset Bear_{1-3} up until its time instant $t_P = 100$. The prediction error, number of rules and membership degree for each rule in the same interval is shown in Figure 5.7b, Figure 5.7c, and Figure 5.7d, respectively. Although the process of merging rules is not depicted in Figure 5.7, due to the choice of parameters, i.e., low values for τ and ω and a high value for γ, redundant rules were merged every 8 processed samples on average. This choice of parameters resulted in a large creation/merge rate throughout the learning phase.

5.2.4 First experiment

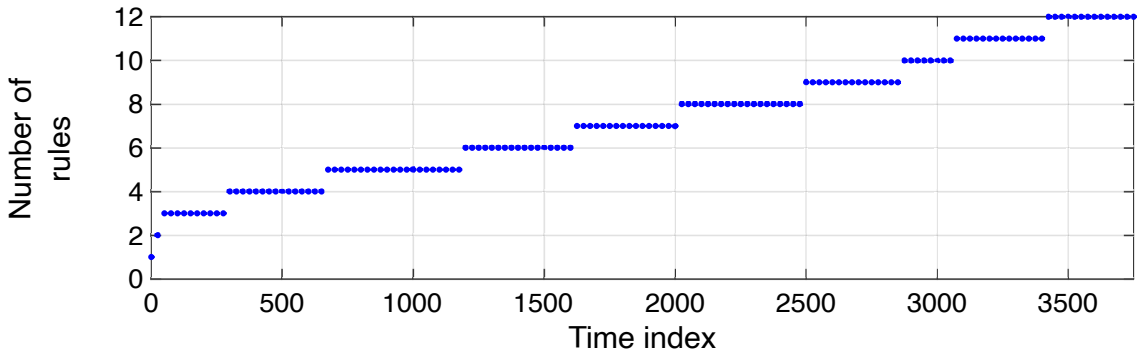
In this experiment, EBeTS is compared with four other prognostics methods: a fuzzy version of the interacting multiple models algorithm, namely IMMF (COSME *et al.*,



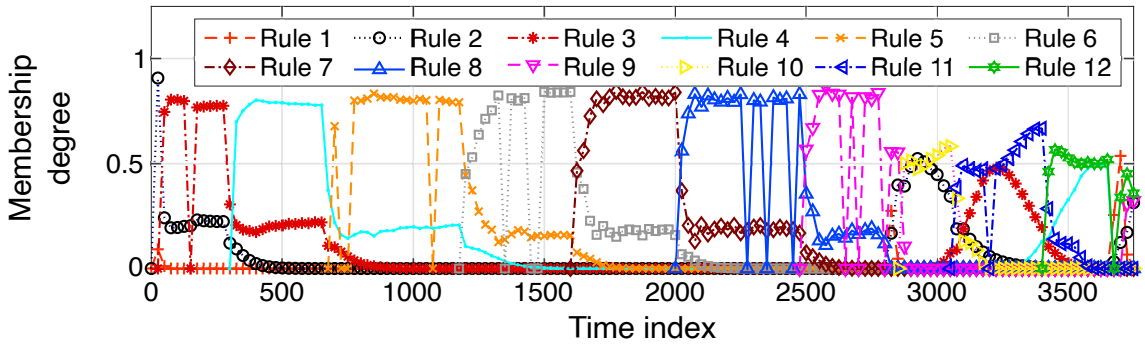
(a) PRONOSTIA's true and predicted HIs for training and initial test sets of condition 1.



(b) PRONOSTIA's HIs prediction error for training and initial test sets of condition 1.



(c) PRONOSTIA's total number of rules for training and initial test sets of condition 1.



(d) PRONOSTIA's fuzzy membership degree for training and initial test sets of condition 1.

Figure 5.7 – PRONOSTIA's one-step ahead prediction results for training and initial test sets of condition 1 using EBeTS.

2019); an EFM based on concepts from the well-known subtractive clustering algorithm (CHIU, 1994), namely exTS (ANGELOV; ZHOU, 2006), the popular ARMA models, and a hybrid method proposed by Camargos *et al.* (2020b) that uses a generic fractional polynomial model propagated through a PF. Both IMMF and exTS are state-of-the-art structures used for fault prognostics based on fuzzy models (GOURIVEAU; ZERHOUNI, 2012); the fact that they are in the same class of fault prognostics techniques justifies the comparison with EBeTS. It is also important to notice that three of them, namely, IMMF, ARMA, and the hybrid approach, are fixed structures where learning happens in an offline stage, and the UUT data are not used to improve their parameter estimates. A brief overview of the techniques used for comparison is given below.

The IMMF described in Cosme *et al.* (2019) consists of a two-stage prognostics approach. The first stage is an offline training phase where known historical data are used to obtain fuzzy rules in a TS fuzzy inference system. The i-th rule's consequent part are affine functions of the past degradation states, such that:

$$\text{Rule } i\text{: IF } x_{k-1} \text{ IS } \Phi_i \text{ THEN } \hat{x}_{i,k} = p_i x_{k-1} + q_i, \tag{5.8}$$

where Φ_i are pi-shaped MFs; each one has four parameters to control their shape. The fuzzy rules are estimated through ANFIS algorithm. The second phase is a modified version of the interactive multiple model (BLOM; BLOEM, 2007) that takes into account the membership degrees to provide the combined estimation of parallel PFs.

The ARMA models are defined as

$$x_t = c + \varepsilon_t + \sum_{i=1}^{p} a_i x_{t-i} + \sum_{i=1}^{q} b_i \varpi_{t-i}, \tag{5.9}$$

where $\varpi_t \sim \mathcal{N}(0, \sigma_\varpi^2)$ are independent identically distributed random variables, c is a constant term, $\mathbf{a} = [a_1 \quad \cdots \quad a_p]^\top \in \mathbb{R}^p$ and $\mathbf{b} = [b_1 \quad \cdots \quad b_q]^\top \in \mathbb{R}^q$ are the model's parameters, and p and q represent the model's structure. Given a pair (p, q) and a data set, the task is to estimate $\mathbf{a}$, $\mathbf{b}$, c and σ_ϖ^2.

The hybrid approach described in Camargos *et al.* (2020b) also consists of a two-phase process where historical data are first used to train the proposed degradation model's parameters. The Levenberg-Marquardt algorithm is used in this step to fit the data into the following generic fractional polynomial model

$$\hat{x}_k = \theta_1 + \theta_2 k^{\theta_3} + \theta_4 x_{k-1}^{\theta_5}. \tag{5.10}$$

Once the parameters are computed, the RUL will be estimated using a PF to propagate the degradation state.

The set of parameters for EBeTS used in this experiment is the one found by the grid search, i.e., $\gamma = 0.8$, $\tau = 2$ and $\omega = 68.27\%$. For the exTS, it is set $\Omega_{\text{exTS}} = 10^3$ as

it is analogous to the rule covariance initialization constant in EBeTS. For all fuzzy-based prognostics approaches, the number of previous states used for predicting the RUL, i.e., the number of regressors, is set to **one**. For the ARMA models, a more flexible number of past states used for prognostics is allowed to fully take advantage of the approach. A grid search is performed with both p and q ranging from 1 to 10, i.e., 100 different combinations of (p, q) for each condition. For each pair (p, q), the first bearing data is used to train the ARMA model's coefficients; the trained models are used for prognostics using the second bearing data enabling the computation of both RA and MAPE metrics. An index that combines both metrics was constructed to find the pair (p, q) that maximizes the RA while minimizing the MAPE and the number of parameters in the model, such that

$$I_k(p,q) = \mathrm{RA}_k(p,q) + \left(1 - \frac{\mathrm{MAPE}_k(p,q)}{100}\right) + \left(1 - \frac{p+q}{20}\right), \tag{5.11}$$

where $\mathrm{RA}_k(p,q)$ and $\mathrm{MAPE}_k(p,q)$ are the RA and MAPE metrics at time k using parameters (p, q), respectively. These metrics were computed with the validation data set, i.e., the second bearing of each condition, and, the chosen value of k for each condition is the time in which prognostics starts in the test data set. The values of (p, q) for each condition that maximizes the index in (5.11) are: $(9, 4)$ for condition 1, $(5, 2)$ for condition 2, and $(4, 1)$ for condition 3. The IMMF requires the number of degradation models to be set beforehand and does not use UUT data to adjust these fixed models; the parameters used are the ones reported in Cosme *et al.* (2019), i.e., for each bearing condition, a fuzzy inference system with three rules, as described in (5.8), is estimated offline. Likewise, the parameters reported in Camargos *et al.* (2020b) are used for the hybrid approach, where each bearing condition has a different set of parameters.

Since there are differences in the bearings' life duration for each operation condition, different starting times and threshold were set for the prognostics task. In bearing conditions 1, 2 and 3, prognostics started at $t_P^1 = 100$ and $t_P^2 = t_P^3 = 20$, respectively, and the threshold was set to $\eta_1 = 20$, $\eta_2 = 8$ and $\eta_3 = 10$, respectively. This choice allows comparisons with the IMMF, for which the proposed method achieves better MAPE results for all bearing conditions, except 1–3 as shown in Table 5.4 and Table 5.5, where MAPE and RA are computed at $k = t_P$. An illustration for the PRONOSTIA's HI prediction for all bearings in the test dataset using these initial conditions is shown in Figure 5.8.

The results in Table 5.4 and Table 5.5 shows a significant improvement in RUL prediction using the proposed method instead of IMMF or ARMA for condition 2, because, in condition 1, the test HI is more similar to the train HI than in condition 2. Figure 5.9 depicts this difference, where the train HI is shown as a continuous black line, the most different test HI, within the same condition, as a red line with squares and the most similar test HI, also within the same condition, as a blue line with circles. This difference indicates that evolving approaches, such as EBeTS and exTS, can yield better results in cases where training and test data have different behavior.

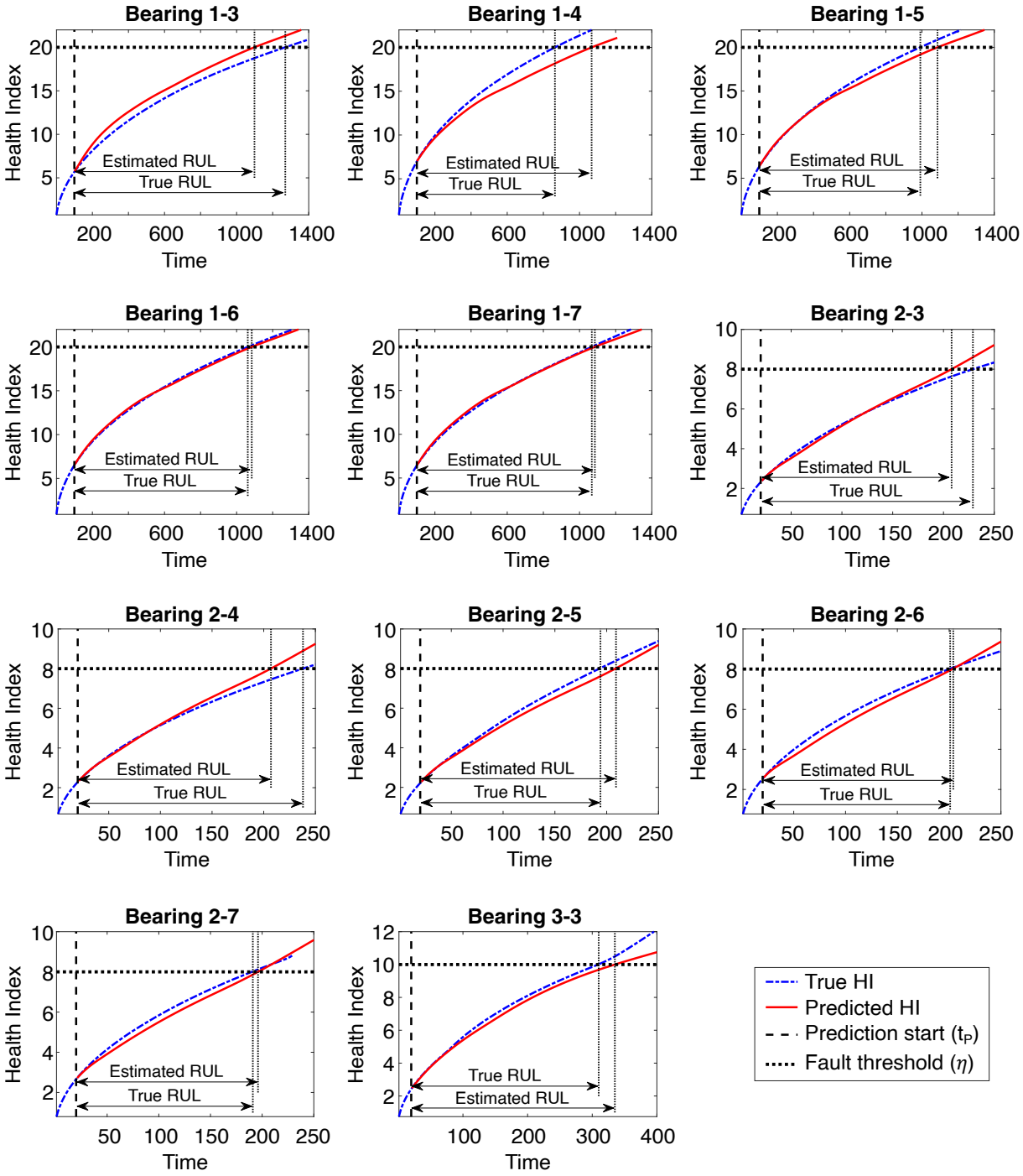


Figure 5.8 – EBeTS long-term prediction of the PRONOSTIA's HIs for bearings 1–3, 2–3 and 3–3 with prediction starting from $t_P^1 = 100$, $t_P^2 = 20$, and $t_P^2 = 20$, respectively, and fault thresholds $\eta_1 = 20$, $\eta_2 = 8$, and $\eta_3 = 10$, respectively.

5.2.5 Second experiment

This experiment evaluates three distinct scenarios for RUL prediction using EBeTS. The first scenario consists of using the available historical data to obtain the TS model in an offline phase, while UUT data are used to adapt the models online, as depicted in Figure 4.1. The second scenario consists of using historical data to train the TS model without any adjustments regarding UUT data, i.e., for condition 1, only data from bearings 1 and 2

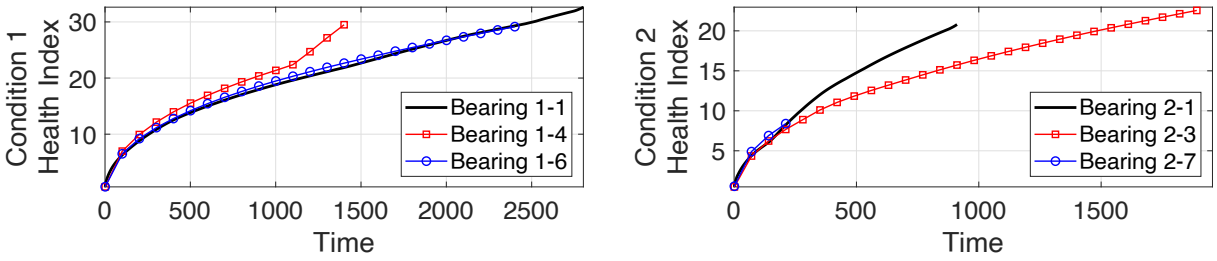


Figure 5.9 – Bearing's HIs of condition 1 and 2 for comparison.

Table 5.4 – Accuracy based metric MAPE computed at $t_P^1 = 100$, $t_P^2 = 20$ and $t_P^3 = 20$ for operation conditions 1, 2 and 3, respectively, with best values in bold.

Bearing	t_P	EBeTS	IMMF	exTS	ARMA	Hybrid
1–3	100	7.01	3.85	9.21	3.58	**2.88**
1–4	100	6.17	9.73	3.71	**2.72**	8.63
1–5	100	2.39	6.41	**1.02**	2.66	1.6
1–6	100	**1.13**	3.01	3.48	7.37	5.45
1–7	100	**1.07**	2.77	3.64	4.94	2.69
Average for condition 1		**3.55**	5.15	4.21	4.25	4.25
2–3	20	**2.48**	20.17	6.52	3.96	10.2
2–4	20	3.55	12.56	7.85	4.98	**1.74**
2–5	20	3.77	19.0	**2.02**	6.75	21.32
2–6	20	5.06	24.91	3.92	**2.96**	4.65
2–7	20	4.56	6.05	3.75	2.65	**1.87**
Average for condition 2		**3.88**	16.54	4.81	4.26	7.96
3–3	20	**2.88**	7.66	4.68	7.85	6.37
Total Average		**3.64**	10.56	4.53	4.58	6.13

are used. In the last scenario, the RUL is predicted when only UUT data are available. Such evaluations are essential to illustrate an important property of EBeTS that is an advantage concerning many of prognostics methods: the ability to take advantage of both historical and UUT data for providing improved RUL predictions. The RUL in these three scenarios is computed at different starting times for each bearing. In some experiments of scenario 3, the failure threshold was not reached within the prediction horizon defined by the failure time; this can happen when knowledge is not yet sufficient to create transition models that can predict the RUL, i.e., the prediction converges prematurely to a value below the failure threshold. The set of parameters used in this experiment for scenarios 1 and 2 is the one found by the grid search, i.e., $\gamma = 0.8$, $\tau = 2$ and $\omega = 68.27\%$. For scenario 3, however, this set can not be used since no training data is available; therefore,

Table 5.5 – Accuracy based metric RA computed at $t_P^1 = 100$, $t_P^2 = 20$ and $t_P^3 = 20$ for operation conditions 1, 2 and 3, respectively, with best values in bold.

Bearing	t_P	EBeTS	IMMF	exTS	ARMA	Hybrid
1–3	100	0.85	**0.97**	0.75	0.87	0.96
1–4	100	0.73	0.91	**0.97**	0.9	0.82
1–5	100	0.89	0.85	**0.94**	0.9	0.91
1–6	100	**0.98**	0.94	0.84	0.79	0.96
1–7	100	0.98	0.97	0.86	0.85	**0.99**
Average for condition 1		0.89	**0.93**	0.87	0.86	0.93
2–3	20	**0.9**	0.84	0.79	0.71	0.75
2–4	20	0.86	0.91	0.76	0.47	**1**
2–5	20	0.91	0.53	**0.97**	0.2	0.42
2–6	20	**0.98**	0.75	0.87	0.74	0.82
2–7	20	**0.97**	0.85	0.9	0.82	0.95
Average for condition 2		**0.92**	0.78	0.86	0.59	0.79
3–3	20	0.91	0.96	0.97	0.76	**1**
Total Average		**0.91**	0.86	0.87	0.73	0.87

the recommendation if Section 3.3 is used, i.e., $\gamma = 0.5$, $\tau = 2$ and $\omega = 95.45\%$.

The results are shown in Table 5.6, where each bearing has a different failure time. All three scenarios were tested with prediction starting from different instants before failure. In these scenarios, it is expected that MAPE decreases and RA increases as the starting prediction time becomes closer to the failure time while maintaining low values of MAPE and high values of RA.

In scenario 1, the proposed scheme, for conditions 1 and 3, values of MAPE decrease in all cases while RA values increase or fluctuate around a value whereas new data are not affecting the models in a perceptive way; this fluctuation is related to the continuous process of creation and combination of rules. For condition 2, the high variability of life duration along all data sets may be responsible for more significant oscillations in MAPE and RA. The results of scenario 3 follow what is expected, but initial values of MAPE are too high since only data from UUT are known. In general, the comparison between scenarios 1 and 3 shows that starting with some historical knowledge gives better RUL estimations, especially when starting prediction time is far from failure time. Additionally, the comparison between scenarios 1 and 2 shows that updating the degradation model when new data are available also gives better RUL estimations.

Table 5.6 – Three distinct scenarios for RUL prediction using EBeTS for different times in which predictions started (t_P).

Bearing	Fails at	t_P	Scenario 1		Scenario 2		Scenario 3	
			MAPE	RA	MAPE	RA	MAPE	RA
1–3	1271	100	**7.01**	**0.85**	7.96	0.78	39.80	–
		700	1.99	**0.84**	3.71	0.69	**0.85**	**0.84**
		1200	**0.05**	**0.96**	0.92	0.56	0.12	0.83
1–4	865	100	**6.17**	0.73	6.57	**0.86**	21.19	–
		500	2.22	0.76	**1.60**	**0.99**	5.54	–
		800	**0.12**	**0.92**	0.54	0.80	0.12	0.88
1–5	992	100	2.39	0.89	**2.23**	**1.00**	30.71	–
		500	1.99	0.87	**1.33**	**0.93**	7.24	–
		800	**0.30**	**0.92**	2.15	0.66	0.86	0.46
1–6	1063	100	1.13	**0.98**	**1.09**	0.92	30.64	–
		500	**1.07**	**0.96**	1.23	0.86	8.42	–
		900	**0.06**	**0.98**	2.34	0.58	0.50	0.66
1–7	1069	100	**1.07**	**0.98**	1.24	0.92	35.32	–
		400	**0.97**	**0.95**	1.02	0.90	14.16	–
		800	**0.13**	**0.96**	2.14	0.71	4.21	–
2–3	230	20	2.48	0.90	**2.30**	**0.91**	44.67	–
		80	3.35	0.84	**3.23**	**0.85**	19.84	–
		200	**0.48**	**0.90**	1.38	0.73	2.13	0
2–4	239	20	3.55	0.86	**3.18**	**0.88**	34.99	–
		80	4.63	0.81	**4.57**	**0.81**	13.31	–
		200	1.88	**0.72**	**1.73**	0.72	1.76	0.08
2–5	195	20	**3.77**	**0.91**	4.44	0.90	44.47	–
		80	**2.26**	**0.92**	2.65	0.90	17.90	–
		160	**0.18**	**1.00**	0.48	0.94	2.73	0
2–6	202	20	**5.06**	**0.98**	5.77	0.97	28.25	–
		100	**0.52**	**0.92**	0.60	0.92	2.73	0.51
		180	**0.27**	**0.91**	0.74	0.77	0.35	0.86
2–7	192	20	**4.56**	**0.97**	7.54	0.92	27.32	–
		80	**0.54**	0.95	0.59	**0.97**	7.35	–

Table 5.6 – Three distinct scenarios for RUL prediction using EBeTS for different times in which predictions started (t_P).

Bearing	Fails at	t_P	Scenario 1		Scenario 2		Scenario 3	
			MAPE	RA	MAPE	RA	MAPE	RA
		140	0.94	0.85	0.54	0.88	**0.22**	**0.96**
3–3	311	20	**2.88**	**0.91**	4.44	0.86	42.53	–
		140	**0.42**	**0.92**	1.29	0.87	1.76	0.64
		280	0.36	**0.84**	0.48	0.77	**0.32**	**0.84**

The $\alpha - \lambda$ plot is usually used to evaluate prognostics strategies since it shows whether the predicted RUL falls within a goal region around the true RUL given by $\pm(\alpha)(100)\%$ (LALL; LOWE; GOEBEL, 2012). Figure 5.10 and Figure 5.11 depicts the $\alpha - \lambda$ plot with goal region of $\alpha = 0.2$ for all bearings on scenario 1, where the RUL is computed for all time instants from the starting prediction point up until the fail time of each bearing, listed in Table 5.6. In almost all situations, the predicted RUL is inside the gol region the whole time and the suggested RUL bounds given by (4.17a) and (4.17b) at a 95% confidence level are shown for some points.

Considering the Monte Carlo method to estimate the Gaussian noise variance, as described in Subsection 4.2.2, it is observed that these bounds increase after a rule is created or combined since the covariance matrix of the model parameters is initialized with a large value; in this case, the new rule has not learned enough to stabilize the covariance matrix used for the posterior distribution. Thus, it leads to an increase of the quantified uncertainty, as depicted in Figure 5.10. When such rules are merged with rules more stable, i.e., rules with knowledge acquired from a more significant data set, the overall uncertainty of the model decreases, as depicted in Figure 5.10. It is also worth pointing out that, for some points of this experiment, the EBeTS models produced confidence lower bounds for the degradation series below the fault threshold. In these cases, the upper bound RUL estimation in (4.17b) cannot be computed; therefore, the RUL lower bound is used reflected as the upper bound, making the suggested bounds symmetric concerning the estimated value.

Although the uncertainty quantification and propagation techniques using Monte Carlo simulations are provided for evolving TS models, its behavior during the rule creation and merging processes still needs to be addressed to make it manageable and offer support for decision-makers. In contrast with the Monte Carlo initialization method, the method described in Subsection 4.2.2 shows more stability concerning the events of creation and combination of rules. The same experiment is depicted in Figure 5.11, showing that not only the uncertainty estimation were reduced but this design allowed the RUL lower bound

to be computed for all tested data points.

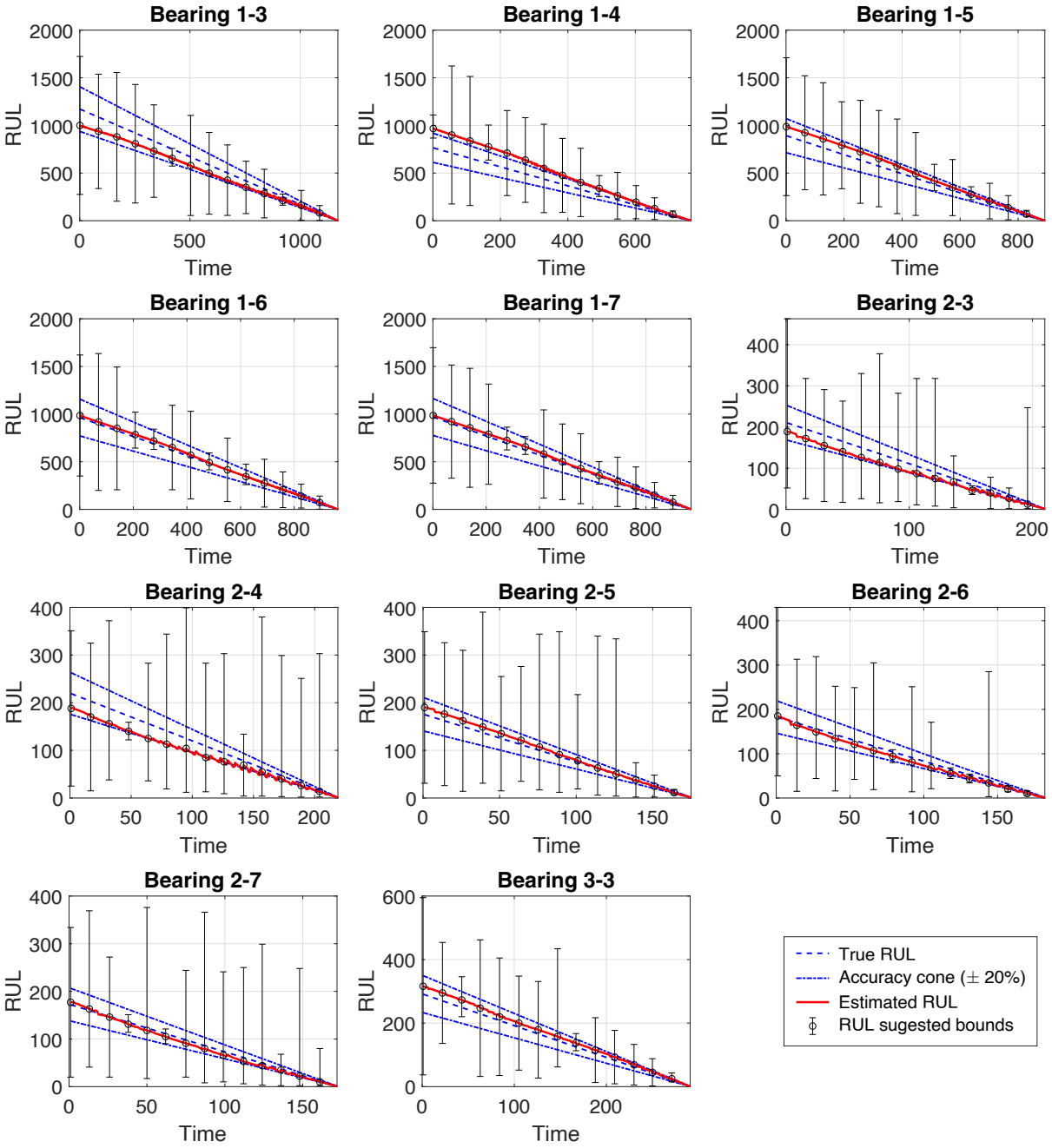


Figure 5.10 – $\alpha - \lambda$ plot of the estimated RUL in scenario 1 with goal region of $\alpha = 0.2$ using the Monte Carlo method to estimate the initial variance.

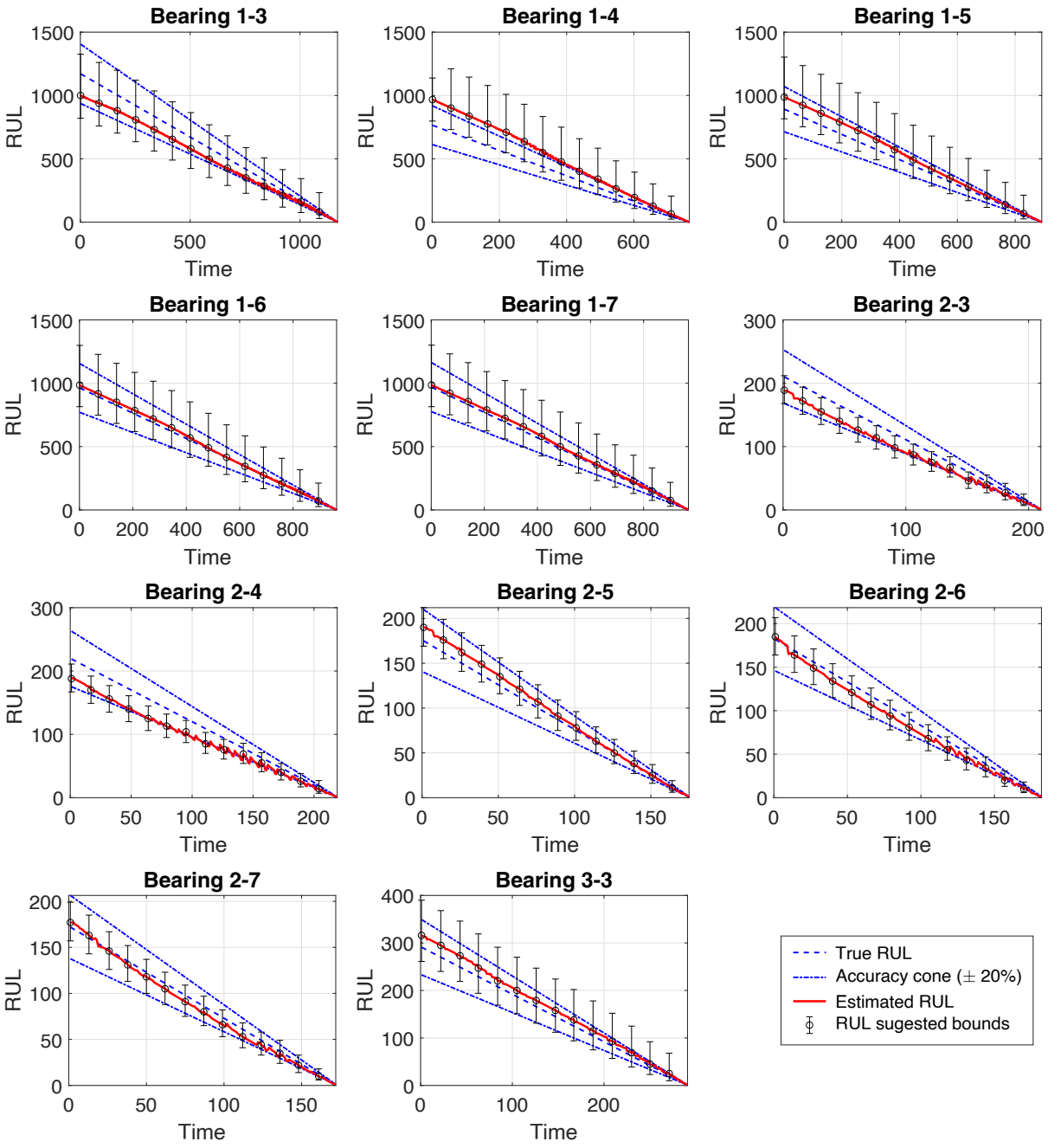
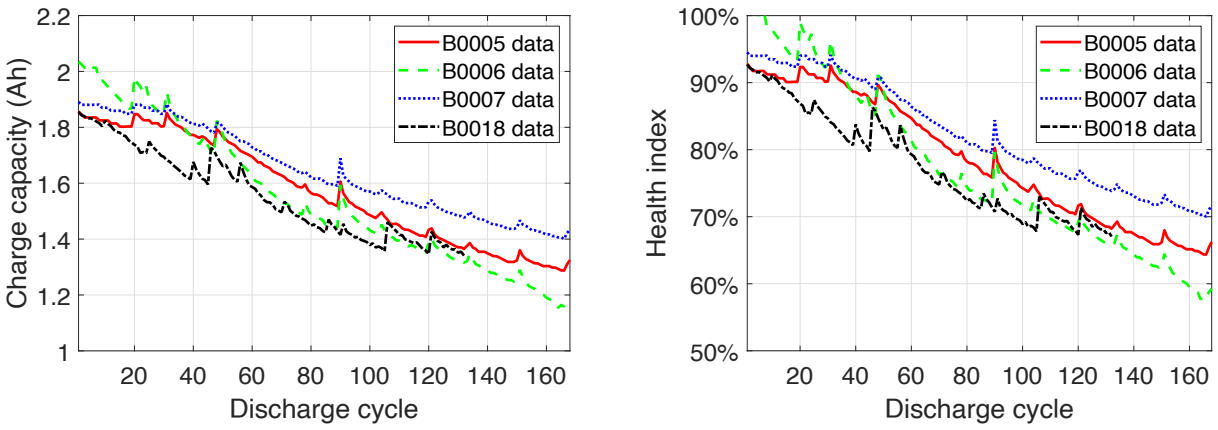


Figure 5.11 – $\alpha - \lambda$ plot of the estimated RUL in scenario 1 with goal region of $\alpha = 0.2$ using the RMSE method to estimate the initial variance.

5.3 Case study 3: battery capacity prediction

The third case deals with the degradation of Li-ion batteries. This type of battery can be found both in the industry and in our day-to-day lives, such as in electric vehicles, micro-grids, and a significant number of consumer electronics (LI; WANG; YAN, 2019; SAHA; GOEBEL, 2009). The cycle aging data sets of four Li-ion batteries are provided by the testbed in NASA Ames Prognostics Center of Excellence (PCoE) and comprises the commercial Li-ion 18650 sized rechargeable batteries from Idaho National Laboratory; programmable 4-channel DC electronic load and power supply; voltmeter, ammeter and thermocouple sensor suite; custom electrochemical impedance spectrometry equipment; and environmental chamber to impose different operational conditions. The batteries were run at room temperature (23º C), where charging was done in a constant mode at 1.5 A until the voltage reached 4.2 V. The discharge process was done at a constant current level of 2 A until the battery voltage reached 2.7 V. More details on the experiments can be found in Saha & Goebel (2009).

In this experiment, the HI used is the percentage charge capacity, shown in Figure 5.12b, instead of its true value, shown in Figure 5.12a. According to Saha & Goebel (2009), when the batteries reached 30% fade in rated capacity (from 2 Ah to 1.4 Ah), the experiments stopped. Therefore, the FT used in this case study is 70%. Figure 5.12 shows the available data set for this experiment, namely, B0005, B0006, B0007, and B0018. The data from B0006 were arbitrarily chosen as the training data set.



(a) True charge capacity of each battery.

(b) Percentage of capacity charge.

Figure 5.12 – True and percentage charge capacity along discharge cycles.

The prognostics results of EBeTS are compared with exTS and ARMA models, both previously used, and three other algorithms, namely, eNFN, eMG, and Long Short-Term Memory (LSTM). To make a more comprehensive comparison, in terms of different classes of prognostics algorithms, the deep learning LSTM model is also used. The recurrent neural networks with LSTM consists of cells that maintain the state over time and using nonlinear gating unites to regulate the information flow into and out of it (HOCHREITER;

SCHMIDHUBER, 1997); these structures are effective and scalable models for learning problems related to sequential data that requires capturing long-term temporal dependence (GREFF *et al.*, 2017).

5.3.1 Parameter tuning

The specific parameters of each algorithm were chosen as standard values reported in the literature. For EBeTS, the choice was made based on the recommendations made in Section 3.3, as $\omega_{\text{EBeTS}} = 95.45\%$, $\tau_{\text{EBeTS}} = \ell + 1$ with ℓ being the number of input lags or autoregressors, $\gamma_{\text{EBeTS}} = 0.5$, $\delta_{\text{EBeTS}} = 10^3$. As in the previous experiments, for exTS the selection is $\Omega_{\text{exTS}} = 10^3$. For the eNFN, the learning rate has been set to $\beta_{\text{eNFN}} = 0.01$, the granularity controller is $\eta_{\text{eNFN}} = 10$, and the threshold to remove MFs based on age is $w_{\text{eNFN}} = 100$. For the eMG algorithm, the learning rate has been set to $\beta_{\text{eMG}} = 0.05$, the unilateral confidence interval to define the compatibility threshold is $\alpha_{\text{eMG}} = 0.01$, the window size for the alert mechanism is $w_{\text{eMG}} = 20$, and the initial dispersion matrix at cluster creation is $\Sigma_{\text{eMG}}^{init} = 10^{-3} \times \mathbf{I}_\ell$. The LSTM algorithm was defined to have 200 hidden units with a solver configured to use the Adam optimizer (KINGMA; BA, 2015) with 250 epochs. The initial learning rate is set to 0.005 and will drop by a factor of 0.2 at every 125 epochs[1].

Tuning the parameters for these techniques is a rather complicated task to do. Although some generic choices are available, they are not robust enough to be chosen in a problem-agnostic manner, as will be shown in the results. The results are susceptible to different parameters, even for exTS with only one adjustable parameter in the present formulation. The EBeTS, on the other hand, shows to be robust to problem independent choices for its three parameters. Using the standard parameters, in this experiment, the free parameter is set to be the number of input lags.

The known data, i.e., the training data from B0006, and 20 samples of each testing data set, are used to validate each approach's lags in each data set. The following modified version of the index in (5.11) is used

$$I_k(\ell, \zeta, \varkappa) = \text{RA}_k(\ell, \zeta, \varkappa) + \left(1 - \frac{\text{MAPE}_k(\ell, \zeta, \varkappa)}{100}\right) + \left(1 - \frac{\ell}{20}\right), \tag{5.12}$$

where $\zeta \in \{\text{B0005}, \text{B0007}, \text{B0018}\}$ is one of the testing batteries, the used algorithm is $\varkappa \in \{\text{EBeTS}, \text{exTS}, \text{ARMA}, \text{eNFN}, \text{eMG}, \text{LSTM}\}$, and ℓ is the number of lags that will range from 1 to 20 in the validation procedure. For the ARMA models the index will be the same as in (5.11) with both p and q ranging from 1 to 10. The number of lags will be

[1] From Mathworks' example on time-series forecasting using deep learning. Available in <https://www.mathworks.com/help/deeplearning/ug/time-series-forecasting-using-deep-learning.html>.

defined as the following maximization problem:

$$\ell(\zeta, \varkappa) = \arg\max_{l} \quad \frac{1}{4} \sum_{j \in \{5,10,15,20\}} I_j(l, \zeta, \varkappa). \tag{5.13}$$

The index in (5.12) depends on the actual RUL of the testing battery to compute the RA, which is not known. Therefore, we propose an approximation for validation purposes that uses the model in (2.5), commonly used to quantify Li-ion battery's charge capacity:

$$C(k; \mathbf{c}) = c_1 \exp(c_2 k) + c_3 \exp(c_4 k), \tag{5.14}$$

where the vector parameters $\mathbf{c} = [c_1 \quad c_2 \quad c_3 \quad c_4]^\top$ can be estimated using the known battery data in the least-squares sense. The function `lsqcurvefit`[2] is used to find the parameters for battery B0006, which will be used as the initial value to find the parameters using the first 20 samples of each testing battery. The model used to approximate the RUL for each test battery will be the average between the model found with training data set B0006 and the model found with the first 20 samples of the testing data, as depicted in Figure 5.13. The model computed with B0006 data is shown in Figure 5.13a; this model is also shown in Figure 5.13b, along with the partial model using B0005 data and the average model. Similarly, these models are also shown for B0007 and B0018 in Figure 5.13c and Figure 5.13d, respectively.

The parameters of the exponential models depicted in Figure 5.13 are listed in Table 5.7. In order to find the parameters of the exponential model that uses the training data of battery B0006, the starting point is $\mathbf{c}_0 = [1 \quad 1 \quad 1 \quad 0]^\top$, while for the other batteries, the coefficients from the B0006 model are used as starting point.

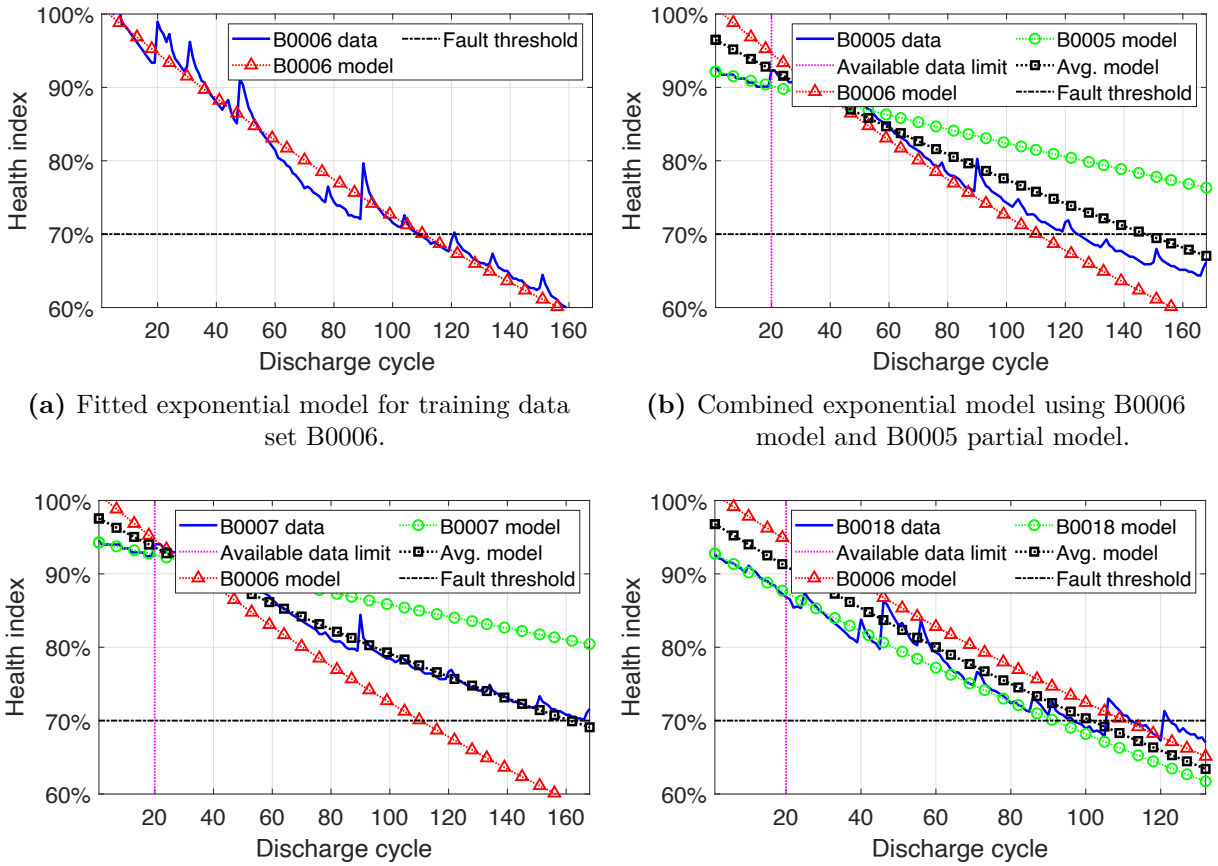
Table 5.7 – Fitting parameters of the exponential models for battery charge capacity available data.

Battery	Exponential model's (5.14) parameters			
	c_1	c_2	c_3	c_4
B0006	-0.4512	13.3905	1.0115	0.0033
B0005	-0.4512	13.3905	0.9226	0.0011
B0007	-0.4512	13.3905	0.9437	0.0010
B0018	-0.4512	13.3905	0.9305	0.0031

Once the parameters of (5.14) are computed for each testing battery data set, a RUL estimate can be done to compute the index (5.12) which will be used in the optimization problem (5.13). Given a battery data ζ and an algorithm $\varkappa$, the following steps are performed to find the best number of lags:

Step 1. Fit an instance of (5.14) using the training data (B0006).

[2] Available in <https://www.mathworks.com/help/optim/ug/lsqcurvefit.html>



(a) Fitted exponential model for training data set B0006.

(b) Combined exponential model using B0006 model and B0005 partial model.

(c) Combined exponential model using B0006 model and B0007 partial model.

(d) Combined exponential model using B0006 model and B0018 partial model.

Figure 5.13 – Fitted exponential models for battery charge capacity data to approximate the true RUL of testing data sets.

Step 2. Fit an instance of (5.14) using the available test data of battery ζ, i.e., the first 20 samples of it.

Step 3. Find a RUL estimate using the average predictions of models from Steps 1 and 2.

Step 4. For each lag between 2 and 20, prepare the training data (B0006) using a Hankel matrix.

Step 5. Use the algorithm $\varkappa$ to train a model for each lag using data from Step 4.

Step 6. Solve the optimization problem in (5.13) to find the best number of lags.

5.3.2 Results and discussion

The optimization problem in (5.13) defines the number of input lags for each algorithm-battery pair. The third column of Table 5.8 shows the number of input lags used, while the subsequent columns show the RA for different starting prognostics points t_P. In this table, the symbol '*' indicates the prognostics task that was not carried out for

that t_P and '–' represents an algorithm's impossibility to compute the RUL for that t_P. This impossibility can happen when the long-term predictions converge to a value greater than the FT or even when they become crescent, never reaching the FT. It is important to note that, for the ARMA models, the Lags column of Table 5.8 corresponds to the p coefficient in (5.9); the q coefficient found in the validation was zero for all batteries. To compare the algorithms between each other considering the same data knowledge, the sample index (s_i) in which the prognostics start is adjusted considering the t_P and the input lags of each battery-algorithm pair. Using the EBeTS' input lags as a basis, the adjustment is $s_i = t_P - (\ell - 3)$. For this reason, the B0018-exTS pair is unable to start the prognostics task at $t_P = 20$ because $s_i = 6$ will be less than the amount of lags $\ell = 17$.

The results in Table 5.8 indicate that multivariate Gaussian models can better represent some problems. An example is the similar results achieved for battery B0005 using both EBeTS and eMG. However, this phenomenon was not observed in batteries B0007 and B0018 in which the non-evolving algorithms ARMA and LSTM performed better in some situations. The reason for that is associated with each technique's chosen parameters as different choices of parameters could lead to better results. In the absence of a fine-tuning procedure for choosing these parameters, the evolving approaches were more prone to create models with wrong long-term tendencies. As for EBeTS, it is essential to emphasize its parameter choice was made based on Section 3.3.

The non-evolving models, such as ARMA and LSTM, produced competitive results in terms of RA. However, when comparing the computation time to train the models and to make successive RUL estimates ($\alpha - \lambda$ plot), the non-evolving LSTM takes a significantly higher time to be trained and used. This disparity is expected since LSTM is a deep learning approach that requires a large amount of parameters to be estimated. The proposed algorithm, was also competitive in terms of computing time during the learning phase. The high computing time for evolving approaches, such as exTS, is related to the creation of more rules to represent the degradation state.

The $\alpha - \lambda$ plot is shown in Figure 5.14, Figure 5.15, and Figure 5.16, for batteries B0005, B0007, and B0018, respectively. The uncertainty is quantified for all evolving techniques using the RMSE method in Section 4.2.2 using a 99% confidence level. The ARMA models' built-in[3] uncertainty propagation shows confidence bounds sufficiently large to encompass the entire goal region, offering poor decision-making support. The LSTM method as no uncertainty quantification mechanism in its standard form. For battery B0005, as shown in Figure 5.14, it is possible to note the similarity between EBeTS and eMG, also shown in Table 5.8. In some experiments, the estimated RUL's absence (red line) in these figures; these are situations where the long-term prediction failed to

[3] From MATLAB's System Identification Toolbox. Available in <https://www.mathworks.com/help/ident/ref/forecast.html>

Table 5.8 – RA for all algorithm-battery pairs with prognostics starting at different t_P, with best values in bold.

Battery	Algorithm	ℓ	t_P					
			20	40	60	80	100	120
B0005 fails at cycle 125	EBeTS	3	**0.9412**	0.7805	0.7581	**0.9762**	**0.9545**	*
	exTS	9	–	–	–	0.9524	0.9091	*
	ARMA	1	0.7647	0.8293	0.8548	0.7381	0.8182	*
	eNFN	3	0.1373	0.3415	0.2097	0.3333	–	*
	eMG	5	0.8922	**0.9756**	**0.9355**	0.9048	**0.9545**	*
	LSTM	3	0.9020	0.9634	0.7903	0.5952	–	*
B0007 fails at cycle 166	EBeTS	3	**0.8182**	**0.8943**	**0.8350**	0.7229	0.7460	0.8140
	exTS	10	0.6853	0.5447	–	–	0.8254	**0.8372**
	ARMA	1	0.5944	0.6179	0.5728	0.5060	0.5238	0.5349
	eNFN	4	–	–	0.1650	0.2530	0.4444	0.2791
	eMG	5	0.6923	0.7561	0.7087	0.6265	–	–
	LSTM	2	0.6643	0.7317	0.7961	**0.8916**	**0.9048**	0.5116
B0018 fails at cycle 97	EBeTS	3	0.9054	**0.9630**	0.7941	**0.7857**	*	*
	exTS	17	*	0.5926	–	–	*	*
	ARMA	1	0.7973	0.7778	**0.9118**	0.5714	*	*
	eNFN	1	–	0.1111	0.2647	0.5714	*	*
	eMG	5	0.8378	0.8889	–	–	*	*
	LSTM	3	**0.9324**	0.7407	0.2647	–	*	*

* prognostics task was not carried out.
– algorithm's impossibility to compute the RULs.

Table 5.9 – Computational time during learning phases of training and testing stages where FNR is the final number of rules.

Algorithm	Training stage		Testing stage ($\alpha - \lambda$ plot)					
	B0006		B0005		B0007		B0018	
	Time (s)	FNR	Time (s)	FNR	Time (s)	FNR	Time (s)	FNR
EBeTS	0.0431	2	2.4630	2	3.7910	2	1.2972	2
exTS	0.2111	11	23.4547	8	49.7472	11	22.2839	12
ARMA	0.2230	–	14.1833	–	14.6436	–	7.0072	–
eNFN	0.0751	2	5.5241	2	10.4547	2	2.9943	2
eMG	0.0453	1	2.0638	1	8.5818	1	10.4159	1
LSTM	11.4016	–	337.8710	–	544.6485	–	326.0498	–

reach the FT, either for converging too soon or assuming a wrong tendency direction.

The long-term prediction for batteries B0005, B0007, and B0018 are shown in Figure 5.17, Figure 5.18, and Figure 5.19 for $t_P = 20$, respectively. As discussed before, the

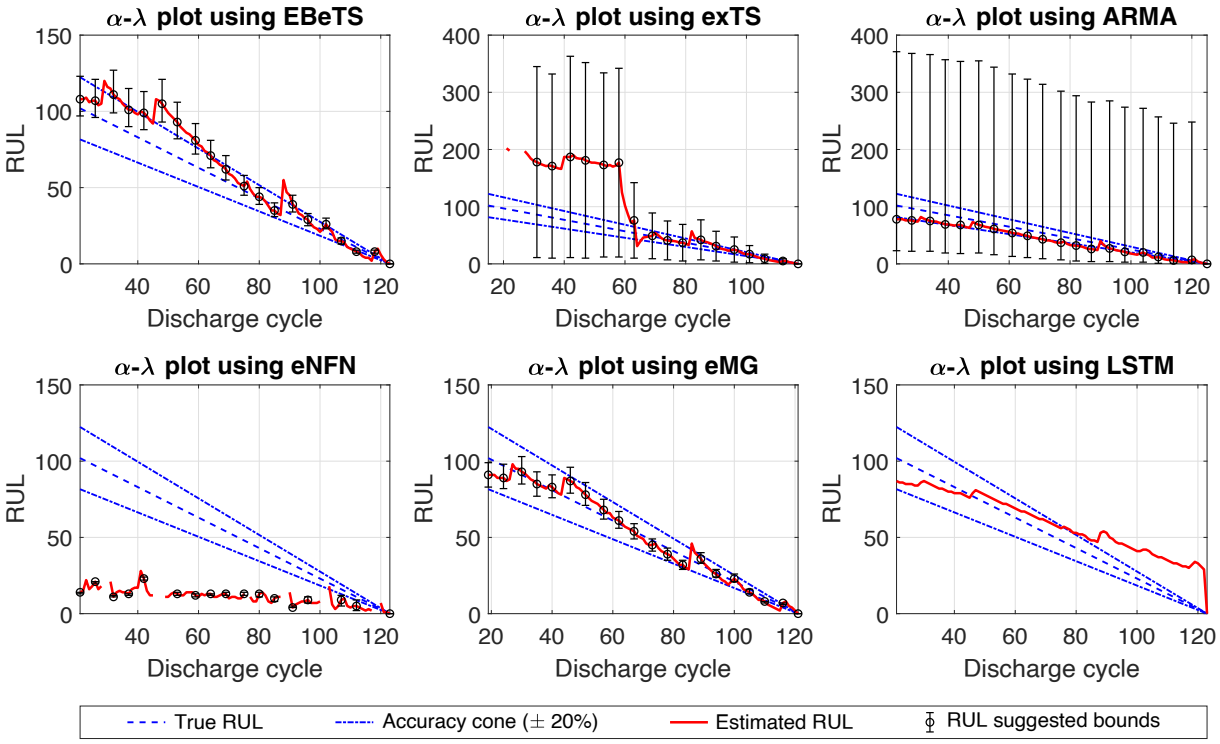


Figure 5.14 – $\alpha-\lambda$ plot of the estimated RUL of battery B0005 with goal region of $\alpha = 0.2$.

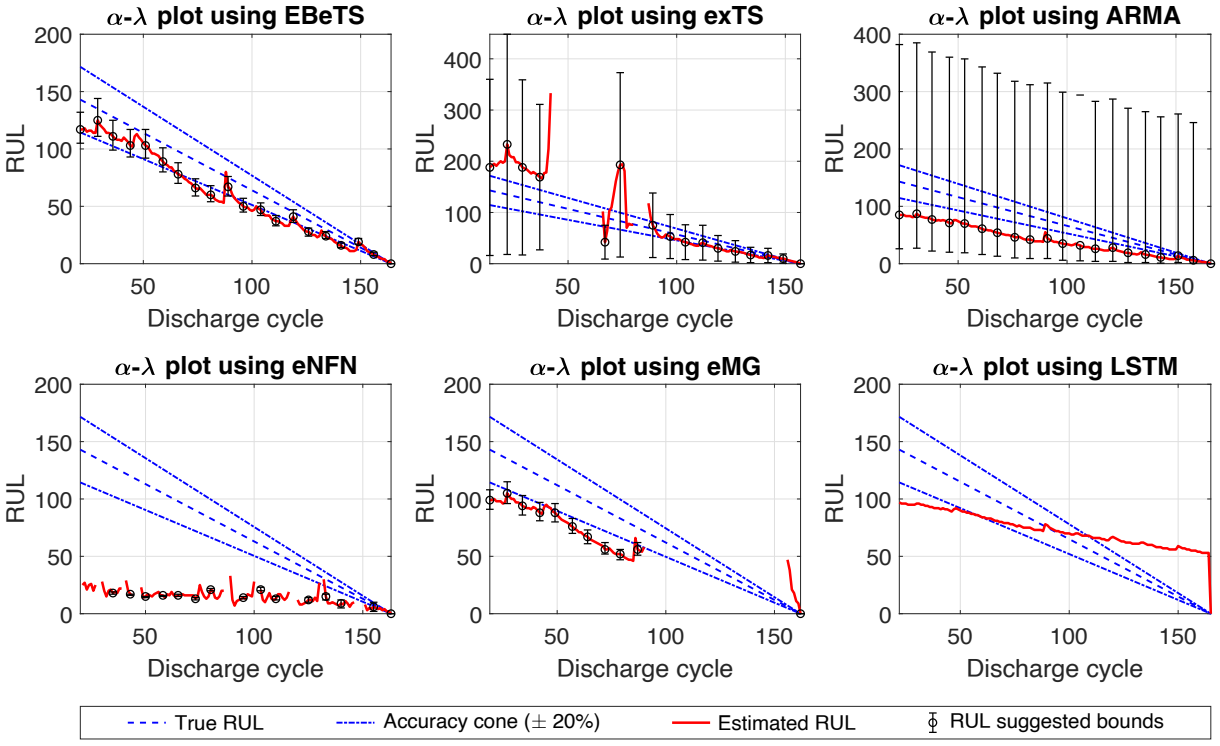


Figure 5.15 – $\alpha-\lambda$ plot of the estimated RUL of battery B0007 with goal region of $\alpha = 0.2$.

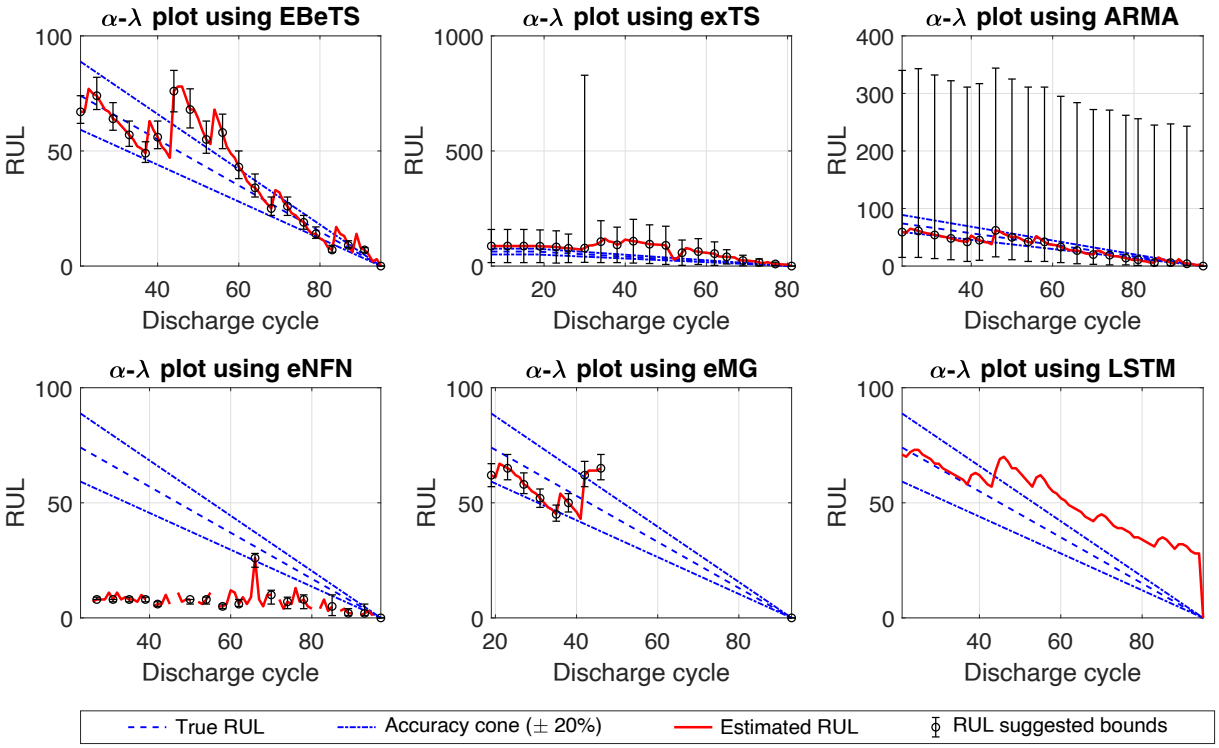


Figure 5.16 – $\alpha-\lambda$ plot of the estimated RUL of battery B0018 with goal region of $\alpha = 0.2$.

sample index is adjusted to match the same sample in the data set for all battery-algorithm pairs. For the pair B0018-exTS, the prediction was made considering $s_i = 17$, which does not allow a numerical comparison as done in Table 5.8 but allows to illustrate the long-term prediction, as shown in Figure 5.19. For the evolving approaches, the fuzzy membership degree is also shown right below each one, except for the non-evolving approaches ARMA and LSTM.

5.4 Chapter summary

In this chapter, the application of fuzzy evolving models to solve real-world problems is explored through three applications. The first case study is a proof-of-concept to evaluate the proposed model's predictive capabilities concerning two well-established techniques in the EFMs literature. The proposed technique is competitive in terms of both accuracy and interpretability metrics. Furthermore, the executed sensitivity study indicates that the proposed method's parametrization can be done through a problem-agnostic methodology, i.e., using knowledge not necessarily related to the problem itself. This work's main focus has been also explored in other two real-world prognostics problems that involve common assets of many industrial applications: rolling bearings and Li-ion batteries. In the bearings experiment, the prognostics' results were compared with different architectures' techniques; the fuzzy evolving models showed more consistent predictions in situations where the training dataset exhibited distinct behaviors than the testing sets. In the battery experiment, the proposed model was tested against other types of EFMs showing similar

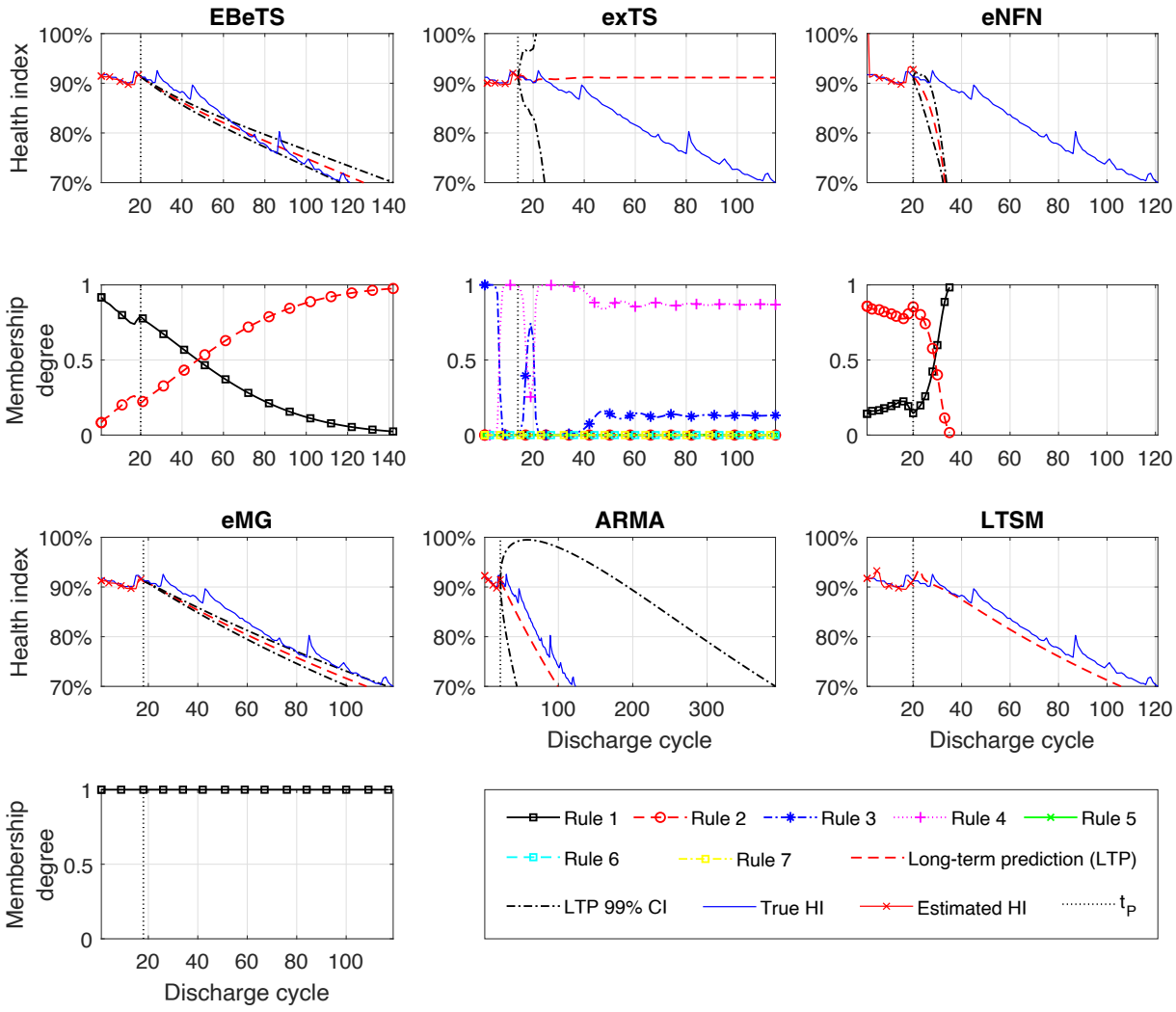


Figure 5.17 – Long-term prediction with 99% confidence intervals of different algorithms in Battery B0005.

results to the eMG technique that also uses multivariate Gaussian membership functions to represent the data clusters. They also exhibited similar results in terms of training and testing time. It is crucial to notice that the chosen techniques' parametrization was based on previous works that did not tackle the long-term prediction problem, which may have considerably affected their performances. In order to carry a fair comparison, the EBeTS parameters were chosen based on the recommendations of Section 3.3, that is, completely independent of the problem.

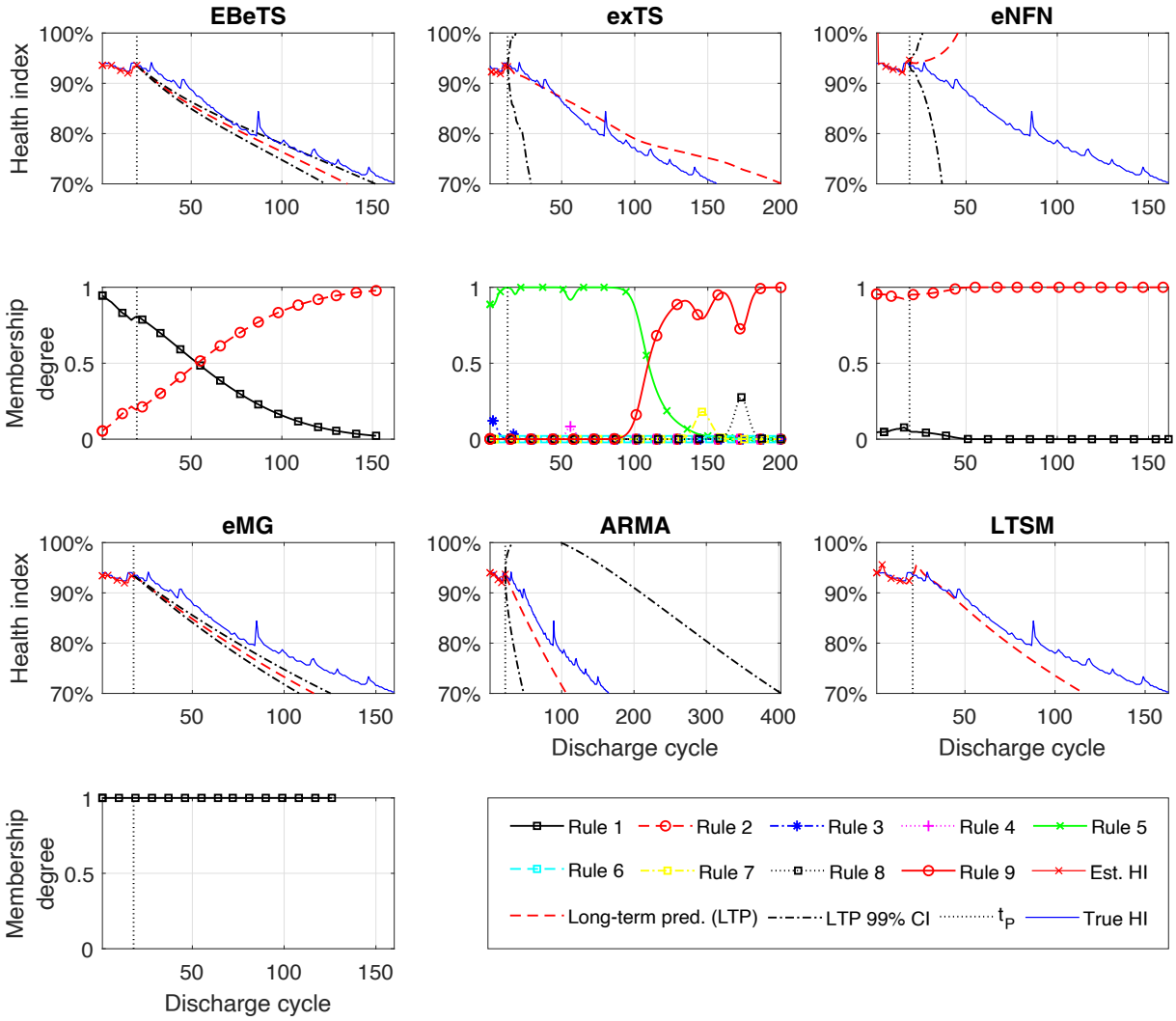


Figure 5.18 – Long-term prediction with 99% confidence intervals of different algorithms in Battery B0007.

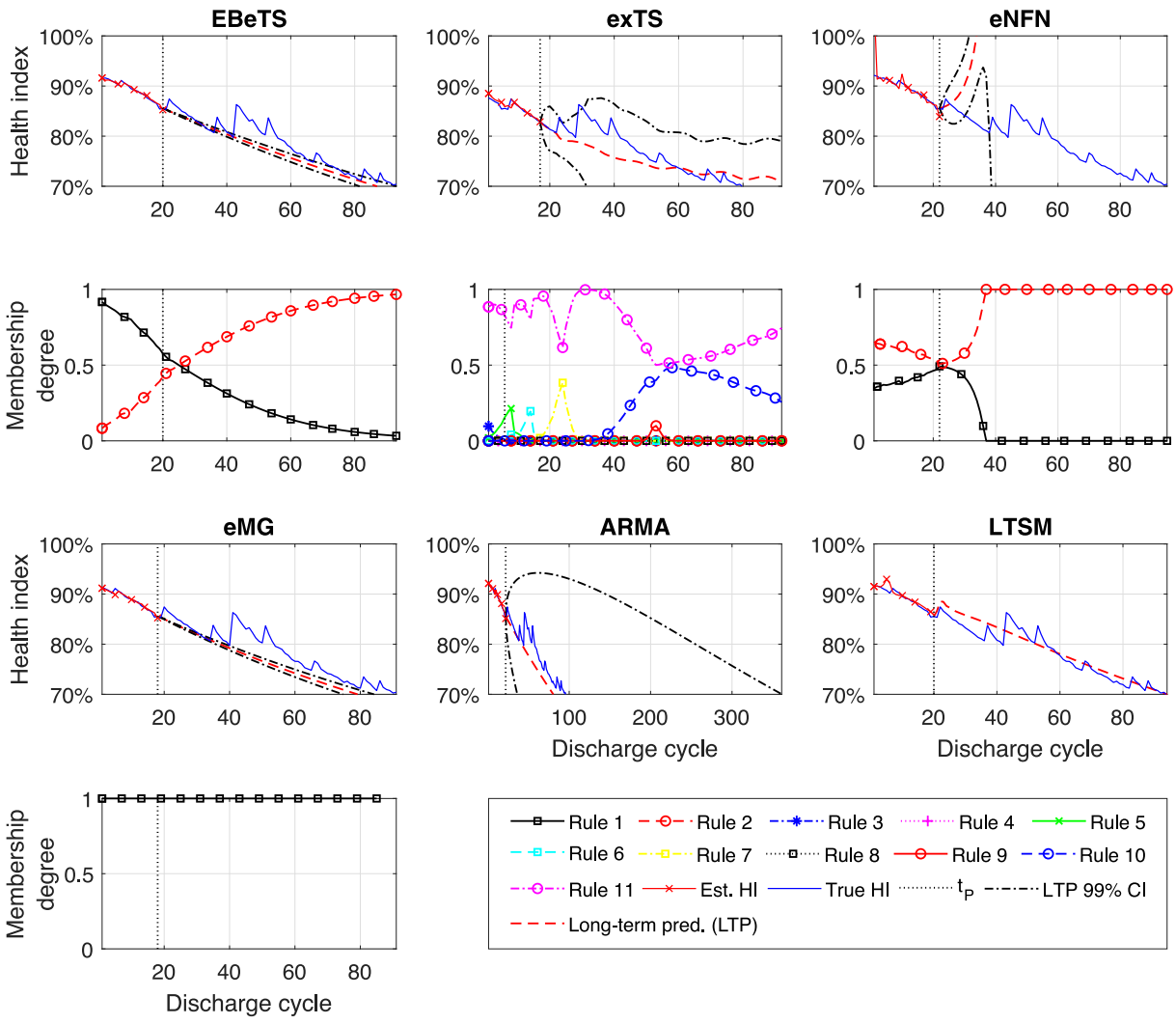


Figure 5.19 – Long-term prediction with 99% confidence intervals of different algorithms in Battery B0018.

6 CONCLUSIONS AND FUTURE WORK

In the last few decades, the industry has been struggling to reduce the life cycle cost by creating smarter maintenance policies to avoid catastrophic failures, guaranteeing safety for its critical systems. In this context, new CBM frameworks are being proposed; these frameworks' main task is the fault prognostics tasks, which include algorithms capable of estimating the RUL of a given asset. In the last years, different techniques for RUL prediction were proposed. Using the asset's physical model can yield very accurate results, but they are challenging and expensive to obtain. Among the data-driven techniques, many depend on a large amount of high-quality training data from different operational conditions to create their models; others have fixed structures that may not work under time variant environments where human intervention to redefine the problem's domain is not always possible.

The Evolving Fuzzy Models are promising alternatives to solve nonlinear problems in time variant environments. Their structure and parameters are flexible and can be adapted recursively according to a data-stream. The low computational complexity of such recursive updates allows its use in real-time systems and enables starting the knowledge-base from scratch. This Thesis has proposed a novel data-driven prognostics approach based on evolving TS fuzzy models denominated EBeTS, whose parameters are estimated through the recursive evaluation of estimation errors. The proposed EBeTS can model nonlinearities and time variant behavior due to its evolving fuzzy structure. The use of multivariate MFs avoids the loss of information related to the complex interactions between input variables.

This capability is illustrated via the EBeTS application in three benchmarks: prediction of temperature time series from three weather stations, the prognostics of rolling bearings in the PRONOSTIA problem, and the prognostics of charge capacity in Li-ion batteries. All the three problems have nonlinear and time-varying dynamics. The proposed algorithm has been compared with different state-of-the-art evolving and non-evolving approaches and has showed very competitive results in terms of different metrics. The prognostics approach based on the EBeTS model provides a way of fusing multivariate condition monitoring information to describe multiple-stage degradation phenomena. Furthermore, a framework to quantify and propagate uncertainties related to the model's estimation error has been provided such that it can be plugged into any EFM, enabling their use in real-world prognostics problems. The proposed quantification technique can use the posterior distribution of the consequent parameters of each rule's noise or the global error covariance matrix to estimate the model's initial uncertainty that will be propagated in the long-term predictions.

This thesis envisions to provide a deepened study concerning the use of EFMs applied to the industrial problem of fault prognostics. This type of model has been chosen due to its suitability for challenging situations, such as (i) modeling systems with nonlinear dynamics; (ii) running in time variant environments; (iii) being able to generate predictions without needing a large amount of training data; (iv) environments that require real-time predictions; (v) to meet the necessity of having interpretable models that can include expert knowledge. The obtained results have demonstrated that there is still room for improvement, such as:

- To improve the uncertainty management framework, enabling a better decision-making process. A bank of adaptive Kalman filters could be considered in an evolving framework whose final prediction is a mixture of all the filters through the Interactive Multiple Model theory. This setting allows the incorporation of different uncertainty sources, such as measurement variability. Furthermore, having such a framework can also enable the exploration of particle filters to model non-Gaussian uncertainties.

- During the tests, some of the long-term predictions made by some of the evolving algorithms faced early convergence to values different from the pre-defined FT, making it impossible to predict the RUL. Once the creating of HIs can consider the trendability and monotonicity degrees, it could be useful to endow the algorithms with a trend-aware mechanism to create more stable long-term predictions.

- Considering exogenous inputs in the evolving models can also help create control laws that can maximize the RUL of the assets. It also allows a cost-effective maintenance policy creation in hostile environments, such as offshore wind turbines.

- The investigation of new HIs with good predictability trough the metrics of monotonicity and trendability that can also allow the detection of the first time of prognostics, i.e., the detection of the degradation beginning.

We hope this thesis can help to consolidate the use of EFMs to solve fault prognostics problems and provides the required foundations for the subsequent developments.

BIBLIOGRAPHY

AHMAD, W.; KHAN, S. A.; ISLAM, M. M. M.; KIM, J. M. A reliable technique for remaining useful life estimation of rolling element bearings using dynamic regression models. **Reliability Engineering and System Safety**, Elsevier Ltd, v. 184, p. 67–76, 2019. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2018.02.003>. pages 21

AHMADZADEH, F.; LUNDBERG, J. Remaining useful life prediction of grinding mill liners using an artificial neural network. **Minerals Engineering**, Elsevier Ltd, v. 53, p. 1–8, 2013. ISSN 08926875. Disponível em: <https://doi.org/10.1016/j.mineng.2013.05.026>. pages 29

AJAMI, A.; DANESHVAR, M. Data driven approach for fault detection and diagnosis of turbine in thermal power plant using Independent Component Analysis (ICA). **International Journal of Electrical Power and Energy Systems**, Elsevier Ltd, v. 43, n. 1, p. 728–735, 2012. ISSN 01420615. Disponível em: <https://doi.org/10.1016/j.ijepes.2012.06.022>. pages 24

AN, D.; KIM, N. H.; CHOI, J. H. Practical options for selecting data-driven or physics-based prognostics algorithms with reviews. **Reliability Engineering and System Safety**, Elsevier, v. 133, p. 223–236, 2015. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2014.09.014>. pages 17, 29

ANGELOV, P. Evolving Takagi-Sugeno Fuzzy Systems from Streaming Data (eTS+). **Evolving Intelligent Systems: Methodology and Applications**, p. 21–50, 2010. Disponível em: <https://doi.org/10.1002/9780470569962.ch2>. pages 32, 36

ANGELOV, P.; ZHOU, X. Evolving fuzzy systems from data streams in real-time. In: **2006 International Symposium on Evolving Fuzzy Systems**. IEEE, 2006. Disponível em: <https://doi.org/10.1109/isefs.2006.251157>. pages 10, 33, 49, 53, 60

ÅSTRÖM, K. J.; WITTENMARK, B. **Computer-Controlled Systems: Theory and Design**. Third edit. [S.l.]: Prentice Hall, 1996. 555 p. pages 37

AYE, S. A.; HEYNS, P. S. An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission. **Mechanical Systems and Signal Processing**, Elsevier, v. 84, p. 485–498, 2017. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2016.07.039>. pages 30

BARRAZA-BARRAZA, D.; TERCERO-GÓMEZ, V. G.; BERUVIDES, M. G.; LIMÓN-ROBLES, J. An adaptive ARX model to estimate the RUL of aluminum plates based on its crack growth. **Mechanical Systems and Signal Processing**, Elsevier, v. 82, p. 519–536, 2017. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2016.05.041>. pages 28

BARTLETT, A.; MARCICKI, J.; ONORI, S.; RIZZONI, G.; YANG, X. G.; MILLER, T. Electrochemical Model-Based State of Charge and Capacity Estimation for a Composite Electrode Lithium-Ion Battery. **IEEE Transactions on Control Systems Technology**, v. 24, n. 2, p. 384–399, 2016. ISSN 10636536. Disponível em: <https://doi.org/10.1109/TCST.2015.2446947>. pages 27

BENKEDJOUH, T.; MEDJAHER, K.; ZERHOUNI, N.; RECHAK, S. Remaining useful life estimation based on nonlinear feature reduction and support vector regression. **Engineering Applications of Artificial Intelligence**, v. 26, n. 7, p. 1751–1760, 2013. Disponível em: <https://doi.org/10.1016/j.engappai.2013.02.006>. pages 56

BLOM, H.; BLOEM, E. Exact bayesian and particle filtering of stochastic hybrid systems. **IEEE Transactions on Aerospace and Electronic Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 43, n. 1, p. 55–70, January 2007. Disponível em: <https://doi.org/10.1109/taes.2007.357154>. pages 60

CADINI, F.; SBARUFATTI, C.; CANCELLIERE, F.; GIGLIO, M. State-of-life prognosis and diagnosis of lithium-ion batteries by data-driven particle filters. **Applied Energy**, Elsevier, v. 235, n. September 2018, p. 661–672, 2018. Disponível em: <https://doi.org/10.1016/j.apenergy.2018.10.095>. pages 17

CAESARENDRA, W.; WIDODO, A.; THOM, P. H.; YANG, B.; SETIAWAN, J. D. Combined probability approach and indirect data-driven method for bearing degradation prognostics. **IEEE Transactions on Reliability**, Institute of Electrical and Electronics Engineers (IEEE), v. 60, n. 1, p. 14–20, March 2011. Disponível em: <https://doi.org/10.1109/tr.2011.2104716>. pages 28

CAMARGOS, M. O.; BESSA, I.; D'ANGELO, M. F. S. V.; COSME, L. B.; PALHARES, R. M. Data-driven prognostics of rolling element bearings using a novel error based evolving takagi–sugeno fuzzy model. **Applied Soft Computing**, Elsevier BV, v. 96, p. 106628, November 2020. Disponível em: <https://doi.org/10.1016/j.asoc.2020.106628>. pages 20

CAMARGOS, M. O.; BESSA, I.; D'ANGELO, M. F. S. V.; PALHARES, R. M. Fault prognostics of rolling bearings using a hybrid approach. **IFAC-PapersOnLine**, Elsevier BV, v. 50, n. 1, July 2020. pages 20, 60, 61

CAMCI, F.; MEDJAHER, K.; ZERHOUNI, N.; NECTOUX, P. Feature evaluation for effective bearing prognostics. **Quality and Reliability Engineering International**, v. 29, n. 4, p. 477–486, 2013. ISSN 07488017. Disponível em: <https://doi.org/10.1002/qre.1396>. pages 23

CARINO, J. A.; ZURITA, D.; DELGADO, M.; ORTEGA, J. A.; ROMERO-TRONCOSO, R. J. Remaining useful life estimation of ball bearings by means of monotonic score calibration. In: **2015 IEEE International Conference on Industrial Technology (ICIT)**. IEEE, 2015. v. 2015-June, n. June, p. 1752–1758. ISBN 978-1-4799-7800-7. Disponível em: <https://doi.org/10.1109/ICIT.2015.7125351>. pages 23

CHATURVEDI, N. A.; KLEIN, R.; CHRISTENSEN, J.; AHMED, J.; KOJIC, A. Algorithms for Advanced Battery-Management Systems. **IEEE Control Systems**, v. 30, n. 3, p. 49–68, jun 2010. ISSN 1066-033X. Disponível em: <https://doi.org/10.1109/MCS.2010.936293>. pages 27

CHEN, C.; VACHTSEVANOS, G.; ORCHARD, M. E. Machine remaining useful life prediction: An integrated adaptive neuro-fuzzy and high-order particle filtering approach. **Mechanical Systems and Signal Processing**, Elsevier, v. 28, p. 597–607, 2012. ISSN 08883270. Disponível em: <https://doi.org/10.1016/j.ymssp.2011.10.009>. pages 21

CHEN, Z. S.; YANG, Y. M.; HU, Z. A technical framework and roadmap of embedded diagnostics/prognostics for complex mechanical systems in PHM systems. **2011 Prognostics and System Health Management Conference, PHM-Shenzhen 2011**, v. 61, n. 2, p. 314–322, 2011. Disponível em: <https://doi.org/10.1109/PHM.2011.5939468>. pages 16

CHENG, S.; AZARIAN, M. H.; PECHT, M. G. Sensor Systems for Prognostics and Health Management. **Sensors**, v. 10, n. 6, p. 5774–5797, jun 2010. ISSN 1424-8220. Disponível em: <https://doi.org/10.3390/s100605774>. pages 21

CHIACHÍO, J.; CHIACHÍO, M.; SANKARARAMAN, S.; SAXENA, A.; GOEBEL, K. Condition-based prediction of time-dependent reliability in composites. **Reliability Engineering and System Safety**, Elsevier, v. 142, p. 134–147, 2015. Disponível em: <https://doi.org/10.1016/j.ress.2015.04.018>. pages 25

CHIACHÍO, J.; JALÓN, M. L.; CHIACHÍO, M.; KOLIOS, A. A Markov chains prognostics framework for complex degradation processes. **Reliability Engineering and System Safety**, v. 195, 2020. Disponível em: <https://doi.org/10.1016/j.ress.2019.106621>. pages 17

CHIU, S. L. Fuzzy model identification based on cluster estimation. **Journal of Intelligent and Fuzzy Systems**, v. 2, n. 3, p. 267–278, 1994. Disponível em: <https://doi.org/10.3233/IFS-1994-2306>. pages 33, 60

CIESLAK, D. A.; HOENS, T. R.; CHAWLA, N. V.; KEGELMEYER, W. P. Hellinger distance decision trees are robust and skew-insensitive. **Data Mining and Knowledge Discovery**, v. 24, n. 1, p. 136–158, 2012. ISSN 13845810. Disponível em: <https://doi.org/10.1007/s10618-011-0222-1>. pages 39

COBLE, J. B. **Merging Data Sources to Predict Remaining Useful Life - An Automated Method to Identify Prognostic Parameters**. PhD dissertation, 2010. Disponível em: <https://trace.tennessee.edu/utk_graddiss/683>. pages 15

CORDOVIL, L. A. Q.; COUTINHO, P. H.; BESSA, I. V.; CAMARGOS, M. O.; PALHARES, R. M. Análise de compromisso entre granularidade e interpretabilidade em sistemas granulares evolutivos. In: **Anais do 14º Simpósio Brasileiro de Automação Inteligente**. Galoa, 2019. Disponível em: <https://doi.org/10.17648/sbai-2019-111316>. pages 32

CORDOVIL, L. A. Q.; COUTINHO, P. H. S.; BESSA, I. V. d.; D'ANGELO, M. F. S. V.; PALHARES, R. M. Uncertain data modeling based on evolving ellipsoidal fuzzy information granules. **IEEE Transactions on Fuzzy Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 28, n. 10, p. 2427–2436, October 2020. Disponível em: <https://doi.org/10.1109/tfuzz.2019.2937052>. pages 18

COSME, L. B.; CAMINHAS, W. M.; D'ANGELO, M. F. S. V.; PALHARES, R. M. A Novel Fault-Prognostic Approach Based on Interacting Multiple Model Filters and Fuzzy Systems. **IEEE Transactions on Industrial Electronics**, v. 66, n. 1, p. 519–528, 2019. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2018.2826449>. pages 18, 30, 54, 60, 61

COSME, L. B.; D'ANGELO, M. F. S. V.; CAMINHAS, W. M.; YIN, S.; PALHARES, R. M. A novel fault prognostic approach based on particle filters and differential evolution.

Applied Intelligence, v. 48, n. 4, p. 834–853, 2018. Disponível em: <https://doi.org/10.1007/s10489-017-1013-1>. pages 17

CSISZÁR, I.; SHIELDS, P. C. Information Theory and Statistics: A Tutorial. **Foundations and Trends™ in Communications and Information Theory**, v. 1, n. 4, p. 417–528, 2004. ISSN 1567-2190. Disponível em: <https://doi.org/10.1561/0100000004>. pages 39

CUBILLO, A.; PERINPANAYAGAM, S.; ESPERON-MIGUEZ, M. A review of physics-based models in prognostics: Application to gears and bearings of rotating machinery. **Advances in Mechanical Engineering**, v. 8, n. 8, p. 1–21, 2016. ISSN 16878140. Disponível em: <https://doi.org/10.1177/1687814016664660>. pages 16, 26

DOWNEY, A.; LUI, Y. H.; HU, C.; LAFLAMME, S.; HU, S. Physics-based prognostics of lithium-ion battery using non-linear least squares with dynamic bounds. **Reliability Engineering and System Safety**, v. 182, n. March 2018, p. 1–12, 2019. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2018.09.018>. pages 27

EL-KOUJOK, M.; GOURIVEAU, R.; ZERHOUNI, N. Reducing arbitrary choices in model building for prognostics: An approach by applying parsimony principle on an evolving neuro-fuzzy system. **Microelectronics Reliability**, v. 51, n. 2, p. 310–320, 2011. Disponível em: <https://doi.org/10.1016/j.microrel.2010.09.014>. pages 18

FANG, X.; ZHOU, R.; GEBRAEEL, N. An adaptive functional regression-based prognostic model for applications with missing data. **Reliability Engineering and System Safety**, Elsevier, v. 133, p. 266–274, 2015. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2014.08.013>. pages 28

GACTO, M. J.; ALCALÁ, R.; HERRERA, F. Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. **Information Sciences**, Elsevier Inc., v. 181, n. 20, p. 4340–4360, 2011. Disponível em: <https://doi.org/10.1016/j.ins.2011.02.021>. pages 50

GEBRAEEL, N.; PAN, J. Prognostic degradation models for computing and updating residual life distributions in a time-varying environment. **IEEE Transactions on Reliability**, v. 57, n. 4, p. 539–550, 2008. ISSN 00189529. Disponível em: <https://doi.org/10.1109/TR.2008.928245>. pages 28

GEBRAEEL, N. Z.; LAWLEY, M. A.; LI, R.; RYAN, J. K. Residual-life distributions from component degradation signals: A Bayesian approach. **IIE Transactions (Institute of Industrial Engineers)**, v. 37, n. 6, p. 543–557, 2005. Disponível em: <https://doi.org/10.1080/07408170590929018>. pages 27

GIBBS, A. L.; SU, F. E. On choosing and bounding probability metrics. **International Statistical Review**, v. 70, n. 3, p. 419–435, 2002. ISSN 03067734. Disponível em: <https://doi.org/10.1111/j.1751-5823.2002.tb00178.x>. pages 39

GOURIVEAU, R.; ZERHOUNI, N. Connexionist-systems-based long term prediction approaches for prognostics. **IEEE Transactions on Reliability**, v. 61, n. 4, p. 909–920, 2012. ISSN 00189529. Disponível em: <https://doi.org/10.1109/TR.2012.2220700>. pages 18, 44, 60

GREFF, K.; SRIVASTAVA, R. K.; KOUTNIK, J.; STEUNEBRINK, B. R.; SCHMIDHUBER, J. LSTM: A search space odyssey. **IEEE Transactions on Neural Networks and Learning Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 28, n. 10, p. 2222–2232, October 2017. Disponível em: <https://doi.org/10.1109/tnnls.2016.2582924>. pages 70

GUBNER, J. **Probability and random processes for electrical and computer engineers**. New York Cambridge: Cambridge University Press, 2006. ISBN 978-0-521-86470-1. pages 102

GUO, J.; LI, Z.; LI, M. A review on prognostics methods for engineering systems. **IEEE Transactions on Reliability**, IEEE, p. 1–20, 2019. Disponível em: <https://doi.org/10.1109/TR.2019.2957965>. pages 16, 17, 26, 48

GUO, L.; GAO, H.; HUANG, H.; HE, X.; LI, S. Multifeatures Fusion and Nonlinear Dimension Reduction for Intelligent Bearing Condition Monitoring. **Shock and Vibration**, v. 2016, 2016. ISSN 10709622. Disponível em: <https://doi.org/10.1155/2016/4632562>. pages 24

GUO, L.; LI, N.; JIA, F.; LEI, Y.; LIN, J. A recurrent neural network based health indicator for remaining useful life prediction of bearings. **Neurocomputing**, Elsevier B.V., v. 240, p. 98–109, 2017. ISSN 18728286. Disponível em: <https://doi.org/10.1016/j.neucom.2017.02.045>. pages 29

GUPTA, M. M.; QI, J. Theory of T-norms and fuzzy inference methods. **Fuzzy Sets and Systems**, v. 40, n. 3, p. 431–450, 1991. Disponível em: <https://doi.org/10.1016/0165-0114(91)90171-L>. pages 35

HAGER, W. W. Updating the Inverse of a Matrix. **SIAM Review**, v. 31, n. 2, p. 221–239, jun 1989. ISSN 0036-1445. Disponível em: <https://doi.org/10.1137/1031049>. pages 36

HANACHI, H.; YU, W.; KIM, I. Y.; LIU, J.; MECHEFSKE, C. K. Hybrid data-driven physics-based model fusion framework for tool wear prediction. **International Journal of Advanced Manufacturing Technology**, The International Journal of Advanced Manufacturing Technology, v. 101, n. 9-12, p. 2861–2872, 2019. ISSN 14333015. Disponível em: <https://doi.org/10.1007/s00170-018-3157-5>. pages 30

HE, W.; WILLIARD, N.; OSTERMAN, M.; PECHT, M. Prognostics of lithium-ion batteries based on Dempster-Shafer theory and the Bayesian Monte Carlo method. **Journal of Power Sources**, Elsevier B.V., v. 196, n. 23, p. 10314–10321, 2011. ISSN 03787753. Disponível em: <https://doi.org/10.1016/j.jpowsour.2011.08.040>. pages 27

HENDERSON, H. V.; SEARLE, S. R. On deriving the inverse of a sum of matrices. **SIAM Review**, Society for Industrial and Applied Mathematics, v. 23, n. 1, p. 53–60, 1981. ISSN 00361445. Disponível em: <http://www.jstor.org/stable/2029838>. pages 103

HENG, A.; ZHANG, S.; TAN, A. C. C.; MATHEW, J. Rotating machinery prognostics: State of the art, challenges and opportunities. **Mechanical Systems and Signal Processing**, v. 23, n. 3, p. 724–739, 2009. ISSN 08883270. Disponível em: <https://doi.org/10.1016/j.ymssp.2008.06.009>. pages 26

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural Computation**, MIT Press - Journals, v. 9, n. 8, p. 1735–1780, November 1997. Disponível em: <https://doi.org/10.1162/neco.1997.9.8.1735>. pages 70

HU, C.; YOUN, B. D.; WANG, P.; Taek Yoon, J. Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life. **Reliability Engineering and System Safety**, Elsevier, v. 103, p. 120–135, 2012. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2012.03.008>. pages 22

HUANG, Z.; XU, Z.; KE, X.; WANG, W.; SUN, Y. Remaining useful life prediction for an adaptive skew-Wiener process model. **Mechanical Systems and Signal Processing**, Elsevier, v. 87, n. October 2016, p. 294–306, 2017. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2016.10.027>. pages 23, 56

INACIO, M.; LEMOS, A.; CAMINHAS, W. Fault diagnosis with evolving fuzzy classifier based on clustering algorithm and drift detection. **Mathematical Problems in Engineering**, Hindawi Limited, v. 2015, p. 1–14, 2015. Disponível em: <https://doi.org/10.1155/2015/368190>. pages 32

JANG, J. R.; SUN, C.; MIZUTANI, E. **Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence**. First edit. [S.l.]: Prentice Hall, 1996. 652 p. pages 30

JARDINE, A. K. S.; LIN, D.; BANJEVIC, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. **Mechanical Systems and Signal Processing**, v. 20, n. 7, p. 1483–1510, oct 2006. ISSN 08883270. Disponível em: <https://doi.org/10.1016/j.ymssp.2005.09.012>. pages 15, 21, 25, 26, 28, 43

JAVED, K.; GOURIVEAU, R.; ZERHOUNI, N. A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering. **IEEE Transactions on Cybernetics**, v. 45, n. 12, p. 2626–2639, 2015. ISSN 21682267. Disponível em: <https://doi.org/10.1109/TCYB.2014.2378056>. pages 24

_____. State of the art and taxonomy of prognostics approaches, trends of prognostics applications and open issues towards maturity at different technology readiness levels. **Mechanical Systems and Signal Processing**, v. 94, p. 214–236, 2017. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2017.01.050>. pages 16, 26

JAVED, K.; GOURIVEAU, R.; ZERHOUNI, N.; NECTOUX, P. Enabling health monitoring approach based on vibration data for accurate prognostics. **IEEE Transactions on Industrial Electronics**, v. 62, n. 1, p. 647–656, 2015. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2014.2327917>. pages 8, 23, 57, 58

JIN, X.; QUE, Z.; SUN, Y.; GUO, Y.; QIAO, W. A Data-Driven Approach for Bearing Fault Prognostics. **IEEE Transactions on Industry Applications**, v. 55, n. 4, p. 3394–3401, 2019. ISSN 00939994. Disponível em: <https://doi.org/10.1109/TIA.2019.2907666>. pages 23

JIN, X.; SUN, Y.; QUE, Z.; WANG, Y.; CHOW, T. W. S. Anomaly detection and fault prognosis for bearings. **IEEE Transactions on Instrumentation and Measurement**, v. 65, n. 9, p. 2046–2054, 2016. ISSN 00189456. Disponível em: <https://doi.org/10.1109/TIM.2016.2570398>. pages 28

JOUIN, M.; GOURIVEAU, R.; HISSEL, D.; PÉRA, M. C.; ZERHOUNI, N. Particle filter-based prognostics: Review, discussion and perspectives. **Mechanical Systems and Signal Processing**, v. 72-73, p. 2–31, 2016. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2015.11.008>. pages 16

KAN, M. S.; TAN, A. C. C.; MATHEW, J. A review on prognostic techniques for non-stationary and non-linear rotating systems. **Mechanical Systems and Signal Processing**, Elsevier, v. 62, p. 1–20, 2015. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2015.02.016>. pages 16, 26, 30

KARNIK, N. N.; MENDEL, J. M. Introduction to type-2 fuzzy logic systems. **1998 IEEE International Conference on Fuzzy Systems Proceedings - IEEE World Congress on Computational Intelligence**, v. 2, p. 915–920, 1998. Disponível em: <https://doi.org/10.1109/FUZZY.1998.686240>. pages 35

KASABOV, N. Evolving fuzzy neural networks for supervised/unsupervised online knowledge-based learning. **IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)**, v. 31, n. 6, p. 902–918, 2001. ISSN 10834419. Disponível em: <https://doi.org/10.1109/3477.969494>. pages 37

KASABOV, N. K.; SONG, Q. DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. **IEEE Transactions on Fuzzy Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 10, n. 2, p. 144–154, April 2002. Disponível em: <https://doi.org/10.1109/91.995117>. pages 36

KHALEGHI, S.; FIROUZ, Y.; Van Mierlo, J.; Van den Bossche, P. Developing a real-time data-driven battery health diagnosis method, using time and frequency domain condition indicators. **Applied Energy**, Elsevier, v. 255, n. August, p. 113813, 2019. ISSN 03062619. Disponível em: <https://doi.org/10.1016/j.apenergy.2019.113813>. pages 21

KHUMPROM, P.; YODO, N. A data-driven predictive prognostic model for lithium-ion batteries based on a deep learning algorithm. **Energies**, v. 12, n. 4, 2019. ISSN 19961073. Disponível em: <https://doi.org/10.3390/en12040660>. pages 21

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. In: **3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings**. [s.n.], 2015. Disponível em: <https://arxiv.org/abs/1412.6980>. pages 70

LALL, P.; LOWE, R.; GOEBEL, K. Prognostics health management of electronic systems under mechanical shock and vibration using Kalman filter models and metrics. **IEEE Transactions on Industrial Electronics**, v. 59, n. 11, p. 4301–4314, 2012. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2012.2183834>. pages 31, 66

LEE, J.; WU, F.; ZHAO, W.; GHAFFARI, M.; LIAO, L.; SIEGEL, D. Prognostics and health management design for rotary machinery systems - Reviews, methodology and applications. **Mechanical Systems and Signal Processing**, Elsevier, v. 42, n. 1-2, p. 314–334, 2014. ISSN 08883270. Disponível em: <https://doi.org/10.1016/j.ymssp.2013.06.004>. pages 26

LEI, Y.; LI, N.; GONTARZ, S.; LIN, J.; RADKOWSKI, S.; DYBALA, J. A Model-Based Method for Remaining Useful Life Prediction of Machinery. **IEEE Transactions on**

Reliability, v. 65, n. 3, p. 1314–1326, 2016. ISSN 00189529. Disponível em: <https://doi.org/10.1109/TR.2016.2570568>. pages 23

_____. A model-based method for remaining useful life prediction of machinery. **IEEE Transactions on Reliability**, v. 65, n. 3, p. 1314–1326, 2016. Disponível em: <https://doi.org/10.1109/tr.2016.2570568>. pages 56

LEI, Y.; LI, N.; GUO, L.; LI, N.; YAN, T.; LIN, J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. **Mechanical Systems and Signal Processing**, Elsevier Ltd, v. 104, n. May, p. 799–834, 2018. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2017.11.016>. pages 16, 21, 22, 23, 24, 25, 26

LEI, Y.; LI, N.; LIN, J. A New Method Based on Stochastic Process Models for Machine Remaining Useful Life Prediction. **IEEE Transactions on Instrumentation and Measurement**, v. 65, n. 12, p. 2671–2684, 2016. ISSN 00189456. Disponível em: <https://doi.org/10.1109/TIM.2016.2601004>. pages 16, 23, 27, 56

LEITE, D.; BALLINI, R.; COSTA, P.; GOMIDE, F. Evolving fuzzy granular modeling from nonstationary fuzzy data streams. **Evolving Systems**, v. 3, n. 2, p. 65–79, 2012. ISSN 18686478. Disponível em: <https://doi.org/10.1007/s12530-012-9050-9>. pages 10, 32, 33, 49, 53

LEITE, D.; PALHARES, R. M.; CAMPOS, V. C. S.; GOMIDE, F. Evolving granular fuzzy model-based control of nonlinear dynamic systems. **IEEE Transactions on Fuzzy Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 23, n. 4, p. 923–938, ago. 2015. Disponível em: <https://doi.org/10.1109/tfuzz.2014.2333774>. pages 32

LEITE, D. F.; HELL, M. B.; COSTA, P.; GOMIDE, F. Real-time fault diagnosis of nonlinear systems. **Nonlinear Analysis: Theory, Methods & Applications**, Elsevier BV, v. 71, n. 12, p. e2665–e2673, dez. 2009. Disponível em: <https://doi.org/10.1016/j.na.2009.06.037>. pages 32

LEITE, G. D. N. P.; ARAÚJO, A. M.; ROSAS, P. A. C. Prognostic techniques applied to maintenance of wind turbines: a concise and specific review. **Renewable and Sustainable Energy Reviews**, v. 81, n. February, p. 1917–1925, 2018. ISSN 18790690. Disponível em: <https://doi.org/10.1016/j.rser.2017.06.002>. pages 21

LEMOS, A.; CAMINHAS, W.; GOMIDE, F. Multivariable gaussian evolving fuzzy modeling system. **IEEE Transactions on Fuzzy Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 19, n. 1, p. 91–104, February 2011. Disponível em: <https://doi.org/10.1109/tfuzz.2010.2087381>. pages 33

LEMOS, A.; CAMINHAS, W.; GOMIDE, F. Adaptive fault detection and diagnosis using an evolving fuzzy classifier. **Information Sciences**, Elsevier BV, v. 220, p. 64–85, jan. 2013. Disponível em: <https://doi.org/10.1016/j.ins.2011.08.030>. pages 32

LI, N.; LEI, Y.; GUO, L.; YAN, T.; LIN, J. Remaining Useful Life Prediction Based on a General Expression of Stochastic Process Models. **IEEE Transactions on Industrial Electronics**, v. 64, n. 7, p. 5709–5718, 2017. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2017.2677334>. pages 28

LI, N.; LEI, Y.; LIN, J.; DING, S. X. An Improved Exponential Model for Predicting Remaining Useful Life of Rolling Element Bearings. **IEEE Transactions on Industrial Electronics**, v. 62, n. 12, p. 7762–7773, 2015. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2015.2455055>. pages 25, 26

LI, N.; LEI, Y.; YAN, T.; LI, N.; HAN, T. A wiener-process-model-based method for remaining useful life prediction considering unit-to-unit variability. **IEEE Transactions on Industrial Electronics**, v. 66, n. 3, p. 2092–2101, 2019. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2018.2838078>. pages 27, 28

LI, R.; HE, D. Rotational machine health monitoring and fault detection using EMD-based acoustic emission feature quantification. **IEEE Transactions on Instrumentation and Measurement**, v. 61, n. 4, p. 990–1001, 2012. ISSN 00189456. Disponível em: <https://doi.org/10.1109/TIM.2011.2179819>. pages 24

LI, W.; LIU, T. Time varying and condition adaptive hidden Markov model for tool wear state estimation and remaining useful life prediction in micro-milling. **Mechanical Systems and Signal Processing**, Elsevier Ltd, v. 131, p. 689–702, 2019. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2019.06.021>. pages 30

LI, X.; DING, Q.; SUN, J. Q. Remaining useful life estimation in prognostics using deep convolution neural networks. **Reliability Engineering and System Safety**, Elsevier Ltd, v. 172, p. 1–11, 2018. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2017.11.021>. pages 29

LI, X.; ELASHA, F.; SHANBR, S.; MBA, D. Remaining Useful Life Prediction of Rolling Element Bearings Using Supervised Machine Learning. **Energies**, v. 12, n. 14, p. 2705, jul 2019. ISSN 1996-1073. Disponível em: <https://doi.org/10.3390/en12142705>. pages 21

LI, X.; WANG, Z.; YAN, J. Prognostic health condition for lithium battery using the partial incremental capacity and Gaussian process regression. **Journal of Power Sources**, Elsevier, v. 421, n. February, p. 56–67, 2019. ISSN 03787753. Disponível em: <https://doi.org/10.1016/j.jpowsour.2019.03.008>. pages 30, 69

LIAO, H.; TIAN, Z. A framework for predicting the remaining useful life of a single unit under time-varying operating conditions. **IIE Transactions (Institute of Industrial Engineers)**, v. 45, n. 9, p. 964–980, 2013. ISSN 0740817X. Disponível em: <https://doi.org/10.1080/0740817X.2012.705451>. pages 23

LIAO, L.; KÖTTIG, F. Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. **IEEE Transactions on Reliability**, v. 63, n. 1, p. 191–207, 2014. ISSN 00189529. Disponível em: <https://doi.org/10.1109/TR.2014.2299152>. pages 16, 21, 27, 28

LIU, D.; XIE, W.; LIAO, H.; PENG, Y. An Integrated Probabilistic Approach to Lithium-Ion. **Ieee Transactions on Instrumentation and Measurement**, v. 64, n. 3, p. 660–670, 2015. ISSN 0018-9456. Disponível em: <https://doi.org/10.1109/TIM.2014.2348613>. pages 29

LIU, H.; CHEN, J.; HISSEL, D.; SU, H. Short-Term Prognostics of PEM Fuel Cells: A Comparative and Improvement Study. **IEEE Transactions on Industrial Electronics**, IEEE, v. 66, n. 8, p. 6077–6086, 2019. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2018.2873105>. pages 30

LIU, J.; ZIO, E. An adaptive online learning approach for Support Vector Regression: Online-SVR-FID. **Mechanical Systems and Signal Processing**, Elsevier, v. 76-77, p. 796–809, 2016. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2016.02.056>. pages 32

LIU, Q.; DONG, M.; LV, W.; GENG, X.; LI, Y. A novel method using adaptive hidden semi-Markov model for multi-sensor monitoring equipment health prognosis. **Mechanical Systems and Signal Processing**, Elsevier, v. 64-65, p. 217–232, 2015. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2015.03.029>. pages 24

LIU, W.; PARK, I.; WANG, Y.; PRINCIPE, J. C. Extended kernel recursive least squares algorithm. **IEEE Transactions on Signal Processing**, v. 57, n. 10, p. 3801–3814, 2009. ISSN 1053587X. Disponível em: <https://doi.org/10.1109/TSP.2009.2022007>. pages 37

LIU, X.; SONG, P.; YANG, C.; HAO, C.; PENG, W. Prognostics and health management of bearings based on logarithmic linear recursive least-squares and recursive maximum likelihood estimation. **IEEE Transactions on Industrial Electronics**, v. 65, n. 2, p. 1549–1558, 2017. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2017.2733469>. pages 26, 30

LIU, Y.; CHANG, Y.; LIU, S.; CHEN, S. Data-driven prognostics of remaining useful life for milling machine cutting tools. **2019 IEEE International Conference on Prognostics and Health Management (ICPHM)**, IEEE, p. 1–5, 2019. Disponível em: <https://doi.org/10.1109/icphm.2019.8819400>. pages 28

LIU, Y.; ZUO, M. J.; LI, Y. F.; HUANG, H. Z. Dynamic Reliability Assessment for Multi-State Systems Utilizing System-Level Inspection Data. **IEEE Transactions on Reliability**, v. 64, n. 4, p. 1287–1299, 2015. ISSN 00189529. Disponível em: <https://doi.org/10.1109/TR.2015.2418294>. pages 29

LJUNG, L. **System Identification: Theory for the User**. 2nd. ed. [S.l.]: Prentice Hall, 1999. 640 p. pages 45

LU, Y.; CHRISTOU, A. Prognostics of IGBT modules based on the approach of particle filtering. **Microelectronics Reliability**, Elsevier, v. 92, n. November 2018, p. 96–105, 2019. ISSN 00262714. Disponível em: <https://doi.org/10.1016/j.microrel.2018.11.012>. pages 27, 30

LUO, J.; NAMBURU, M.; PATTIPATI, K.; QIAO, L.; KAWAMOTO, M.; CHIGUSA, S. Model-based prognostic techniques [maintenance applications]. In: **Proceedings AUTOTESTCON 2003. IEEE Systems Readiness Technology Conference.** IEEE, 2003. p. 330–340. Disponível em: <https://doi.org/10.1109/autest.2003.1243596>. pages 16

LY, C.; TOM, K.; BYINGTON, C. S.; PATRICK, R.; VACHTSEVANOS, G. J. Fault diagnosis and failure prognosis for engineering systems: A global perspective. **2009 IEEE International Conference on Automation Science and Engineering, CASE 2009**, p. 108–115, 2009. Disponível em: <https://doi.org/10.1109/COASE.2009.5234094>. pages 16

MA, C.; ZHAI, X.; WANG, Z.; TIAN, M.; YU, Q.; LIU, L.; LIU, H.; WANG, H.; YANG, X. State of health prediction for lithium-ion batteries using multiple-view feature fusion

and support vector regression ensemble. **International Journal of Machine Learning and Cybernetics**, Springer Berlin Heidelberg, v. 10, n. 9, p. 2269–2282, 2019. ISSN 1868-8071. Disponível em: <https://doi.org/10.1007/s13042-018-0865-y>. pages 30

MA, H. G.; WU, J. P.; LI, X. Y.; KANG, R. Condition-based maintenance optimization for multicomponent systems under imperfect repair-based on RFAD model. **IEEE Transactions on Fuzzy Systems**, v. 27, n. 5, p. 917–927, 2019. ISSN 10636706. Disponível em: <https://doi.org/10.1109/TFUZZ.2018.2878165>. pages 15

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. **Forecasting : methods and applications**. New York: John Wiley & Sons, 1998. ISBN 978-0-471-53233-0. pages 50

MAMDANI, E. H. Application of Fuzzy Algorithms for Control of Simple Dynamic Plant. **Proceedings of the Institution of Electrical Engineers**, v. 121, n. 12, p. 1585–1588, 1974. ISSN 00203270. Disponível em: <https://doi.org/10.1049/piee.1974.0328>. pages 35

MEDJAHER, K.; ZERHOUNI, N.; BAKLOUTI, J. Data-driven prognostics based on health indicator construction: Application to PRONOSTIA's. In: **2013 European Control Conference (ECC)**. [s.n.], 2013. Disponível em: <https://doi.org/10.23919/ecc.2013.6669223>. pages 56

MENG, H.; LI, Y. A review on prognostics and health management (PHM) methods of lithium-ion batteries. **Renewable and Sustainable Energy Reviews**, Elsevier Ltd, v. 116, n. September, p. 109405, 2019. ISSN 13640321. Disponível em: <https://doi.org/10.1016/j.rser.2019.109405>. pages 29

MOSHTAGHI, M.; LECKIE, C.; BEZDEK, J. C. Online clustering of multivariate time-series. **16th SIAM International Conference on Data Mining 2016, SDM 2016**, n. June, p. 360–368, 2016. Disponível em: <https://doi.org/10.1137/1.9781611974348.41>. pages 38

NECTOUX, P.; GOURIVEAU, R.; MEDJAHER, K.; RAMASSO, E.; CHEBEL-MORELLO, B.; ZERHOUNI, N.; VARNIER, C. PRONOSTIA : An experimental platform for bearings accelerated degradation tests. In: **IEEE International Conference on Prognostics and Health Management, PHM'12.** Denver, Colorado, United States: IEEE Catalog Number : CPF12PHM-CDR, 2012. sur CD ROM, p. 1–8. Disponível em: <https://hal.archives-ouvertes.fr/hal-00719503>. pages 20, 55, 56

NGUYEN, A. T.; TANIGUCHI, T.; ECIOLAZA, L.; CAMPOS, V.; PALHARES, R.; SUGENO, M. Fuzzy control systems: Past, present and future. **IEEE Computational Intelligence Magazine**, v. 14, n. 1, p. 56–68, 2019. ISSN 15566048. Disponível em: <https://doi.org/10.1109/MCI.2018.2881644>. pages 35

NGUYEN, H. P.; LIU, J.; ZIO, E. Ensemble of Models for Fatigue Crack Growth Prognostics. **IEEE Access**, IEEE, v. 7, p. 49527–49537, 2019. ISSN 21693536. Disponível em: <https://doi.org/10.1109/ACCESS.2019.2910611>. pages 26, 30

NOORTWIJK, J. M. van. A survey of the application of gamma processes in maintenance. **Reliability Engineering and System Safety**, v. 94, n. 1, p. 2–21, 2009. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2007.03.019>. pages 28

PALHARES, R. M.; YUAN, Y.; WANG, Q. Artificial intelligence in industrial systems. **IEEE Transactions on Industrial Electronics**, v. 66, n. 12, p. 9636–9640, 2019. Disponível em: <https://doi.org/10.1109/TIE.2019.2916709>. pages 16

PAN, D.; LIU, J. B.; CAO, J. Remaining useful life estimation using an inverse Gaussian degradation model. **Neurocomputing**, Elsevier, v. 185, p. 64–72, 2016. ISSN 18728286. Disponível em: <https://doi.org/10.1016/j.neucom.2015.12.041>. pages 28

PAN, Y.; ER, M. J.; LI, X.; YU, H.; GOURIVEAU, R. Machine health condition prediction via online dynamic fuzzy neural networks. **Engineering Applications of Artificial Intelligence**, v. 35, p. 105–113, 2014. Disponível em: <https://doi.org/10.1016/j.engappai.2014.05.015>. pages 56

PANDEY, M. D.; YUAN, X. X.; NOORTWIJK, J. M. van. The influence of temporal uncertainty of deterioration on life-cycle management of structures. **Structure and Infrastructure Engineering**, v. 5, n. 2, p. 145–156, 2009. ISSN 15732479. Disponível em: <https://doi.org/10.1080/15732470601012154>. pages 28

PARIS, P.; ERDOGAN, F. A critical analysis of crack propagation laws. **Journal of Fluids Engineering, Transactions of the ASME**, v. 85, n. 4, p. 528–533, 1963. ISSN 1528901X. Disponível em: <https://doi.org/10.1115/1.3656900>. pages 26

PENG, W.; ZHU, S. P.; SHEN, L. The transformed inverse Gaussian process as an age- and state-dependent degradation model. **Applied Mathematical Modelling**, Elsevier Inc., v. 75, p. 837–852, 2019. ISSN 0307904X. Disponível em: <https://doi.org/10.1016/j.apm.2019.07.004>. pages 28

PENG, Y.; DONG, M.; ZUO, M. J. Current status of machine prognostics in condition-based maintenance: A review. **International Journal of Advanced Manufacturing Technology**, v. 50, n. 1-4, p. 297–313, 2010. ISSN 02683768. Disponível em: <https://doi.org/10.1007/s00170-009-2482-0>. pages 17, 29

PRATAMA, M.; ANAVATTI, S. G.; ANGELOV, P. P.; LUGHOFER, E. PANFIS: A novel incremental learning machine. **IEEE Transactions on Neural Networks and Learning Systems**, v. 25, n. 1, p. 55–68, 2014. ISSN 2162237X. Disponível em: <https://doi.org/10.1109/TNNLS.2013.2271933>. pages 37

QIAN, Y.; YAN, R. Remaining Useful Life Prediction of Rolling Bearings Using an Enhanced Particle Filter. **IEEE Transactions on Instrumentation and Measurement**, v. 64, n. 10, p. 2696–2707, 2015. ISSN 00189456. Disponível em: <https://doi.org/10.1109/TIM.2015.2427891>. pages 28

QIAN, Y.; YAN, R.; HU, S. Bearing degradation evaluation using recurrence quantification analysis and Kalman filter. **IEEE Transactions on Instrumentation and Measurement**, v. 63, n. 11, p. 2599–2610, 2014. Disponível em: <https://doi.org/10.1109/tim.2014.2313034>. pages 28

QIU, H.; LEE, J.; LIN, J.; YU, G. Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. **Advanced Engineering Informatics**, v. 17, n. 3-4, p. 127–140, 2003. ISSN 14740346. Disponível em: <https://doi.org/10.1016/j.aei.2004.08.001>. pages 23

RAGAB, A.; YACOUT, S.; OUALI, M.; OSMAN, H. Prognostics of multiple failure modes in rotating machinery using a pattern-based classifier and cumulative incidence functions. **Journal of Intelligent Manufacturing**, Springer US, v. 30, n. 1, p. 255–274, jan 2019. ISSN 0956-5515. Disponível em: <https://doi.org/10.1007/s10845-016-1244-8>. pages 21

RAMASSO, E.; DENOEUX, T. Making use of partial knowledge about hidden states in HMMs: An approach based on belief functions. **IEEE Transactions on Fuzzy Systems**, v. 22, n. 2, p. 395–405, 2014. ISSN 10636706. Disponível em: <https://doi.org/10.1109/TFUZZ.2013.2259496>. pages 18, 29

RAMASSO, E.; ROMBAUT, M.; ZERHOUNI, N. Joint prediction of continuous and discrete states in time-series based on belief functions. **IEEE Transactions on Cybernetics**, v. 43, n. 1, p. 37–50, 2013. ISSN 21682267. Disponível em: <https://doi.org/10.1109/TSMCB.2012.2198882>. pages 24

RASHID, A.; HITCHCOCK, L.; IBIDO, M.; GIGANT, A. **Condition monitoring and diagnostics of wind turbines – Part 1: General guidelines**. Geneva, CH, 2017. Disponível em: <https://www.iso.org/standard/55634.html>. pages 21

RAZAVI-FAR, R.; CHAKRABARTI, S.; SAIF, M.; ZIO, E. An integrated impu-tation-prediction scheme for prognostics of battery data with missing observations. **Expert Systems with Applications**, Elsevier Ltd, v. 115, p. 709–723, 2019. ISSN 09574174. Disponível em: <https://doi.org/10.1016/j.eswa.2018.08.033>. pages 21, 29

RONG, H. J.; SUNDARARAJAN, N.; HUANG, G. B.; SARATCHANDRAN, P. Sequential Adaptive Fuzzy Inference System (SAFIS) for nonlinear system identification and prediction. **Fuzzy Sets and Systems**, v. 157, n. 9, p. 1260–1275, 2006. ISSN 01650114. Disponível em: <https://doi.org/10.1016/j.fss.2005.12.011>. pages 37

SAHA, B.; GOEBEL, K. Modeling li-ion battery capacity depletion in a particle filtering framework. In: **Proceedings of the annual conference of the prognostics and health management society**. [s.n.], 2009. p. 2909–2924. Disponível em: <https://www.phmsociety.org/node/167>. pages 69

SANKARARAMAN, S. Significance, interpretation, and quantification of uncertainty in prognostics and remaining useful life prediction. **Mechanical Systems and Signal Processing**, Elsevier, v. 52-53, n. 1, p. 228–247, 2015. ISSN 10961216. Disponível em: <https://doi.org/10.1016/j.ymssp.2014.05.029>. pages 48

SAXENA, A.; CELAYA, J.; BALABAN, E.; GOEBEL, K.; SAHA, B.; SAHA, S.; SCHWABACHER, M. Metrics for evaluating performance of prognostic techniques. **2008 International Conference on Prognostics and Health Management, PHM 2008**, 2008. Disponível em: <https://doi.org/10.1109/PHM.2008.4711436>. pages 19, 21, 31, 57

SAXENA, A.; GOEBEL, K.; SIMON, D.; EKLUND, N. Damage propagation modeling for aircraft engine run-to-failure simulation. In: **2008 International Conference on Prognostics and Health Management**. IEEE, 2008. Disponível em: <https://doi.org/10.1109/phm.2008.4711414>. pages 21

SCANLON, P.; KAVANAGH, D. F.; BOLAND, F. M. Residual life prediction of rotating machines using acoustic noise signals. **IEEE Transactions on Instrumentation and Measurement**, v. 62, n. 1, p. 95–108, 2013. ISSN 00189456. Disponível em: <https://doi.org/10.1109/TIM.2012.2212508>. pages 24

SCHWABACHER, M. A Survey of Data-Driven Prognostics. In: **Infotech@ Aerospace**. Reston, Virigina: American Institute of Aeronautics and Astronautics, 2005. p. 1–5. ISBN 978-1-62410-069-7. Disponível em: <https://doi.org/10.2514/6.2005-7002>. pages 21

SI, X. S.; WANG, W.; HU, C. H.; ZHOU, D. H. Remaining useful life estimation - A review on the statistical data driven approaches. **European Journal of Operational Research**, Elsevier B.V., v. 213, n. 1, p. 1–14, 2011. ISSN 03772217. Disponível em: <https://doi.org/10.1016/j.ejor.2010.11.018>. pages 16, 27, 29

SI, X. S.; WANG, W.; HU, C. H.; CHEN, M. Y.; ZHOU, D. H. A Wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation. **Mechanical Systems and Signal Processing**, Elsevier, v. 35, n. 1-2, p. 219–237, 2013. ISSN 08883270. Disponível em: <https://doi.org/10.1016/j.ymssp.2012.08.016>. pages 25

SIKORSKA, J. Z.; HODKIEWICZ, M.; MA, L. Prognostic modelling options for remaining useful life estimation by industry. **Mechanical Systems and Signal Processing**, v. 25, n. 5, p. 1803–1836, 2011. ISSN 08883270. Disponível em: <https://doi.org/10.1016/j.ymssp.2010.11.018>. pages 26

SILVA, A. M.; CAMINHAS, W.; LEMOS, A.; GOMIDE, F. A fast learning algorithm for evolving neo-fuzzy neuron. **Applied Soft Computing Journal**, Elsevier B.V., v. 14, n. PART B, p. 194–209, 2014. ISSN 15684946. Disponível em: <https://doi.org/10.1016/j.asoc.2013.03.022>. pages 33

SINGLETON, R. K.; STRANGAS, E. G.; AVIYENTE, S. Extended Kalman filtering for remaining-useful-life estimation of bearings. **IEEE Transactions on Industrial Electronics**, v. 62, n. 3, p. 1781–1790, 2015. Disponível em: <https://doi.org/10.1109/tie.2014.2336616>. pages 56

ŠKRJANC, I.; IGLESIAS, J.; SANCHIS, A.; LEITE, D.; LUGHOFER, E.; GOMIDE, F. Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: A Survey. **Information Sciences**, v. 490, p. 344–368, 2019. ISSN 00200255. Disponível em: <https://doi.org/10.1016/j.ins.2019.03.060>. pages 32, 37

SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and Computing**, v. 14, n. 3, p. 199–222, aug 2004. ISSN 0960-3174. Disponível em: <https://doi.org/10.1023/B:STCO.0000035301.49549.88>. pages 30

SON, K. L.; FOULADIRAD, M.; BARROS, A. Remaining useful lifetime estimation and noisy gamma deterioration process. **Reliability Engineering & System Safety**, Elsevier BV, v. 149, p. 76–87, May 2016. Disponível em: <https://doi.org/10.1016/j.ress.2015.12.016>. pages 28

SONG, C.; LIU, K. Statistical degradation modeling and prognostics of multiple sensor signals via data fusion: A composite health index approach. **IISE Transactions**, Taylor & Francis, v. 50, n. 10, p. 853–867, 2018. ISSN 24725862. Disponível em: <https://doi.org/10.1080/24725854.2018.1440673>. pages 28

SOUALHI, A.; MEDJAHER, K.; ZERHOUNI, N. Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression. **IEEE Transactions on Instrumentation and Measurement**, v. 64, n. 1, p. 52–62, 2015. ISSN 00189456. Disponível em: <https://doi.org/10.1109/TIM.2014.2330494>. pages 24

SOUALHI, A.; RAZIK, H.; CLERC, G.; DOAN, D. D. Prognosis of bearing failures using hidden markov models and the adaptive neuro-fuzzy inference system. **IEEE Transactions on Industrial Electronics**, v. 61, n. 6, p. 2864–2874, 2014. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2013.2274415>. pages 24

SUTRISNO, E.; OH, H.; VASAN, A. S. S.; PECHT, M. Estimation of remaining useful life of ball bearings using data driven methodologies. **PHM 2012 - 2012 IEEE Int. Conf.on Prognostics and Health Management: Enhancing Safety, Efficiency, Availability, and Effectiveness of Systems Through PHM Technology and Application, Conference Program**, v. 2, p. 1–7, 2012. Disponível em: <https://doi.org/10.1109/ICPHM.2012.6299548>. pages 23

TANG, S. J.; YU, C. Q.; WANG, X.; GUO, X. S.; SI, X. S. Remaining useful life prediction of lithium-ion batteries based on the Wiener process with measurement error. **Energies**, v. 7, n. 2, p. 520–547, 2014. Disponível em: <https://doi.org/10.3390/en7020520>. pages 27

TAO, T.; ZIO, E.; ZHAO, W. A novel support vector regression method for online reliability prediction under multi-state varying operating conditions. **Reliability Engineering and System Safety**, Elsevier Ltd, v. 177, p. 35–49, 2018. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2018.04.027>. pages 30

TOBON-MEJIA, D. A.; MEDJAHER, K.; ZERHOUNI, N.; TRIPOT, G. A Data-Driven Failure Prognostics Method Based on Mixture of Gaussians Hidden Markov Models. **IEEE Transactions on Reliability**, v. 61, n. 2, p. 491–503, jun 2012. ISSN 0018-9529. Disponível em: <https://doi.org/10.1109/TR.2012.2194177>. pages 25

TOMS, L. A.; TOMS, A. M. **Machinery Oil Analysis: Methods, Automation and Benefits - a Guide for Maintenance Managers, Supervisors & Technicians**. [S.l.]: Society of Tribologists and Lubrication Engineers, 2008. ISBN 9780981751207. pages 15

TSUI, K. L.; CHEN, N.; ZHOU, Q.; HAI, Y.; WANG, W. Prognostics and health management: A review on data driven approaches. **Mathematical Problems in Engineering**, v. 2015, 2015. ISSN 15635147. Disponível em: <https://doi.org/10.1155/2015/793161>. pages 28

TZAFESTAS, S. G.; ZIKIDIS, K. C. NeuroFAST: On-line neuro-fuzzy ART-based structure and parameter learning TSK model. **IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics**, v. 31, n. 5, p. 797–802, 2001. ISSN 10834419. Disponível em: <https://doi.org/10.1109/3477.956041>. pages 37

VAPNIK, V. N. **The Nature of Statistical Learning Theory**. New York, NY: Springer New York, 2000. ISBN 978-1-4419-3160-3. Disponível em: <https://doi.org/10.1007/978-1-4757-3264-1>. pages 30

VARGAS, R. E. V.; MUNARO, C. J.; CIARELLI, P. M.; MEDEIROS, A. G.; AMARAL, B. G. do; BARRIONUEVO, D. C.; ARAÚJO, J. C. D. de; RIBEIRO, J. L.; MAGALHÃES, L. P. A realistic and public dataset with rare undesirable real events in oil wells. **Journal of Petroleum Science and Engineering**, Elsevier BV, v. 181, p. 106223, out. 2019. Disponível em: <https://doi.org/10.1016/j.petrol.2019.106223>. pages 15

WALKER, E.; RAYMAN, S.; WHITE, R. E. Comparison of a particle filter and other state estimation methods for prognostics of lithium-ion batteries. **Journal of Power Sources**, Elsevier B.V, v. 287, p. 1–12, 2015. ISSN 03787753. Disponível em: <https://doi.org/10.1016/j.jpowsour.2015.04.020>. pages 27, 30

WANG, F. K.; MAMO, T. A hybrid model based on support vector regression and differential evolution for remaining useful lifetime prediction of lithium-ion batteries. **Journal of Power Sources**, Elsevier, v. 401, n. March, p. 49–54, 2018. ISSN 03787753. Disponível em: <https://doi.org/10.1016/j.jpowsour.2018.08.073>. pages 30

WANG, J.; LIANG, Y.; ZHENG, Y.; GAO, R. X.; ZHANG, F. An integrated fault diagnosis and prognosis approach for predictive maintenance of wind turbine bearing with limited samples. **Renewable Energy**, Elsevier Ltd, v. 145, p. 642–650, 2020. ISSN 18790682. Disponível em: <https://doi.org/10.1016/j.renene.2019.06.103>. pages 27, 30

WANG, T. Bearing life prediction based on vibration signals: A case study and lessons learned. In: **2012 IEEE Conference on Prognostics and Health Management**. IEEE, 2012. p. 1–7. ISBN 978-1-4673-0358-3. Disponível em: <https://doi.org/10.1109/ICPHM.2012.6299547>. pages 23

WANG, W. Q.; GOLNARAGHI, M. F.; ISMAIL, F. Prognosis of machine health condition using neuro-fuzzy systems. **Mechanical Systems and Signal Processing**, v. 18, n. 4, p. 813–831, 2004. ISSN 08883270. Disponível em: <https://doi.org/10.1016/S0888-3270(03)00079-7>. pages 29

WANG, Y.; PENG, Y.; ZI, Y.; JIN, X.; TSUI, K. L. A Two-Stage Data-Driven-Based Prognostic Approach for Bearing Degradation Problem. **IEEE Transactions on Industrial Informatics**, v. 12, n. 3, p. 924–932, 2016. ISSN 15513203. Disponível em: <https://doi.org/10.1109/TII.2016.2535368>. pages 23, 24, 56

WIDODO, A.; YANG, B. S. Application of relevance vector machine and survival probability to machine degradation assessment. **Expert Systems with Applications**, Elsevier Ltd, v. 38, n. 3, p. 2592–2599, 2011. Disponível em: <https://doi.org/10.1016/j.eswa.2010.08.049>. pages 23

WU, J.; SU, Y.; CHENG, Y.; SHAO, X.; DENG, C.; LIU, C. Multi-sensor information fusion for remaining useful life prediction of machining tools by adaptive network based fuzzy inference system. **Applied Soft Computing Journal**, Elsevier B.V., v. 68, p. 13–23, 2018. Disponível em: <https://doi.org/10.1016/j.asoc.2018.03.043>. pages 17

WU, S.; ER, M. J. Dynamic fuzzy neural networks - a novel approach to function approximation. **IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics**, v. 30, n. 2, p. 358–364, 2000. ISSN 10834419. Disponível em: <https://doi.org/10.1109/3477.836384>. pages 37

WU, S.; ER, M. J.; GAO, Y. A fast approach for automatic generation of fuzzy rules by generalized dynamic fuzzy neural networks. **IEEE Transactions on Fuzzy Systems**, v. 9, n. 4, p. 578–594, 2001. ISSN 10636706. Disponível em: <https://doi.org/10.1109/91.940970>. pages 37

XI, F.; SUN, Q.; KRISHNAPPA, G. Bearing diagnostics based on pattern recognition of statistical parameters. **Journal of Vibration and Control**, SAGE Publications, v. 6, n. 3,

p. 375–392, March 2000. Disponível em: <https://doi.org/10.1177/10775463000600303>. pages 23

XI, Z.; JING, R.; WANG, P.; HU, C. A copula-based sampling method for data-driven prognostics. **Reliability Engineering and System Safety**, Elsevier, v. 132, p. 72–82, 2014. ISSN 09518320. Disponível em: <https://doi.org/10.1016/j.ress.2014.06.014>. pages 23

XIAO, L.; CHEN, X.; ZHANG, X.; LIU, M. A novel approach for bearing remaining useful life estimation under neither failure nor suspension histories condition. **Journal of Intelligent Manufacturing**, Springer US, v. 28, n. 8, p. 1893–1914, 2017. ISSN 15728145. Disponível em: <https://doi.org/10.1007/s10845-015-1077-x>. pages 29

XU, J.; WANG, Y.; XU, L. PHM-oriented integrated fusion prognostics for aircraft engines based on sensor data. **IEEE Sensors Journal**, v. 14, n. 4, p. 1124–1132, 2014. ISSN 1530437X. Disponível em: <https://doi.org/10.1109/JSEN.2013.2293517>. pages 16

YAGER, R. R. A model of participatory learning. **IEEE Transactions on Systems, Man, and Cybernetics**, Institute of Electrical and Electronics Engineers (IEEE), v. 20, n. 5, p. 1229–1234, 1990. Disponível em: <https://doi.org/10.1109/21.59986>. pages 33

YAMAKAWA, T.; UCHINO, E.; MIKI, T.; KUSANAGI, H. A neo fuzzy neuron and its applications to system identification and prediction of the system behavior. In: **Proc. of the 2nd Int. Conf. on Fuzzy Logic & Neural Networks**. [S.l.]: Fuzzy Logic Systems Institute, 1992. p. 477–483. pages 33

YANG, F.; HABIBULLAH, M. S.; ZHANG, T.; XU, Z.; LIM, P.; NADARAJAN, S. Health index-based prognostics for remaining useful life predictions in electrical machines. **IEEE Transactions on Industrial Electronics**, v. 63, n. 4, p. 2633–2644, 2016. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2016.2515054>. pages 23, 29

YEN, Y. C.; SÖHNER, J.; LILLY, B.; ALTAN, T. Estimation of tool wear in orthogonal cutting using the finite element analysis. **Journal of Materials Processing Technology**, v. 146, n. 1, p. 82–91, 2004. Disponível em: <https://doi.org/10.1016/S0924-0136(03)00847-1>. pages 30

YIU, J. C. M.; WANG, S. Multiple ARMAX modeling scheme for forecasting air conditioning system performance. **Energy Conversion and Management**, v. 48, n. 8, p. 2276–2285, 2007. ISSN 01968904. Disponível em: <https://doi.org/10.1016/j.enconman.2007.03.018>. pages 28

ZAIDI, S. S. H.; AVIYENTE, S.; SALMAN, M.; SHIN, K. K.; STRANGAS, E. G. Prognosis of gear failures in dc starter motors using hidden Markov models. **IEEE Transactions on Industrial Electronics**, v. 58, n. 5, p. 1695–1706, 2011. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2010.2052540>. pages 29

ZHANG, B.; SCONYERS, C.; BYINGTON, C.; PATRICK, R.; ORCHARD, M. E.; VACHTSEVANOS, G. A probabilistic fault detection approach: Application to bearing fault detection. **IEEE Transactions on Industrial Electronics**, v. 58, n. 5, p. 2011–2018, 2011. ISSN 02780046. Disponível em: <https://doi.org/10.1109/TIE.2010.2058072>. pages 24

ZHANG, B.; ZHANG, L.; XU, J. Degradation Feature Selection for Remaining Useful Life Prediction of Rolling Element Bearings. **Quality and Reliability Engineering International**, v. 32, n. 2, p. 547–554, 2016. ISSN 10991638. Disponível em: <https://doi.org/10.1002/qre.1771>. pages 23

ZHANG, C.; LIM, P.; QIN, A. K.; TAN, K. C. Multiobjective Deep Belief Networks Ensemble for Remaining Useful Life Estimation in Prognostics. **IEEE Transactions on Neural Networks and Learning Systems**, v. 28, n. 10, p. 2306–2318, 2017. ISSN 21622388. Disponível em: <https://doi.org/10.1109/TNNLS.2016.2582798>. pages 29

ZHANG, Z.; SI, X.; HU, C.; LEI, Y. Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods. **European Journal of Operational Research**, Elsevier B.V., v. 271, n. 3, p. 775–796, 2018. ISSN 03772217. Disponível em: <https://doi.org/10.1016/j.ejor.2018.02.033>. pages 28

ZHANG, Z. X.; SI, X. S.; HU, C. H. An Age- and State-Dependent Nonlinear Prognostic Model for Degrading Systems. **IEEE Transactions on Reliability**, v. 64, n. 4, p. 1214–1228, 2015. ISSN 00189529. Disponível em: <https://doi.org/10.1109/TR.2015.2419220>. pages 23

ZHAO, M.; TANG, B.; TAN, Q. Bearing remaining useful life estimation based on time–frequency representation and supervised dimensionality reduction. **Measurement**, v. 86, p. 41–55, 2016. Disponível em: <https://doi.org/10.1016/j.measurement.2015.11.047>. pages 56

ZHOU, D.; YIN, H.; FU, P.; SONG, X.; LU, W.; YUAN, L.; FU, Z. Prognostics for state of health of lithium-ion batteries based on Gaussian process regression. **Mathematical Problems in Engineering**, v. 2018, 2018. ISSN 15635147. Disponível em: <https://doi.org/10.1155/2018/8358025>. pages 30

ZHOU, Y.; HUANG, M. Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model. **Microelectronics Reliability**, Elsevier Ltd, v. 65, p. 265–273, 2016. ISSN 00262714. Disponível em: <https://doi.org/10.1016/j.microrel.2016.07.151>. pages 28

ZURITA, D.; CARINO, J. A.; DELGADO, M.; ORTEGA, J. A. Distributed neuro-fuzzy feature forecasting approach for condition monitoring. **19th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2014**, 2014. Disponível em: <https://doi.org/10.1109/ETFA.2014.7005180>. pages 24

Appendix

APPENDIX A – RECURSIVE ESTIMATION OF STATISTICAL MEASURES

This chapter provides closed formulas for the recursive computation of statistical measures, such as the mean, auto-correlation matrix, covariance matrix, and inverse covariance matrix. These computations allow the tracking of such measures in data-streams, enabling its application to real-time systems.

A.1 Recursive estimation of the mean

Let $\hat{\boldsymbol{\mu}}_n \in \mathbb{R}^m$ be the sample mean of a data stream composed of n observations $\mathbf{x}_i \in \mathbb{R}^m$, defined as

$$\hat{\boldsymbol{\mu}}_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i. \tag{A.1}$$

The recursion relation can be found after separating the last term in the summation, such that

$$\begin{aligned}\hat{\boldsymbol{\mu}}_n &= \frac{1}{n}\left(\sum_{i=1}^{n-1}\mathbf{x}_i + \mathbf{x}_n\right)\\ &= \frac{1}{n}\left[(n-1)\hat{\boldsymbol{\mu}}_{n-1} + \mathbf{x}_n\right],\end{aligned} \tag{A.2}$$

therefore, the recursive algorithm for the sample mean of a data stream that requires only the past value of the mean and the new observation at time instant n is given by

$$\hat{\boldsymbol{\mu}}_n = \left(1 - \frac{1}{n}\right)\hat{\boldsymbol{\mu}}_{n-1} + \frac{1}{n}\mathbf{x}_n. \tag{A.3}$$

A.2 Recursive estimation of the auto-correlation matrix

Let $\mathbf{X} \in \mathbb{R}^{m\times n}$ be a data set with n data points $\mathbf{x} \in \mathbb{R}^m$ with mean vector $\boldsymbol{\mu}_{\mathbf{X}} \triangleq \mathbb{E}[\mathbf{X}]$. The full auto-correlation matrix of $\mathbf{X}$ is given as

$$\mathbf{R}_{\mathbf{XX}} = \begin{bmatrix} \frac{1}{n}\sum_{k=1}^{n} x_{1k}\bar{x}_{k1} & \cdots & \frac{1}{n}\sum_{k=1}^{n} x_{1k}\bar{x}_{km} \\ \vdots & \ddots & \vdots \\ \frac{1}{n}\sum_{k=1}^{n} x_{mk}\bar{x}_{k1} & \cdots & \frac{1}{n}\sum_{k=1}^{n} x_{mk}\bar{x}_{km}, \end{bmatrix} \tag{A.4}$$

where x_{ij} and $\bar{x}_{ij}$ are the j-th elements of the i-th of $\mathbf{X}$ and $\mathbf{X}^\top$, respectively. Let $r_{ij,n}$ be the j-th element of the i-th row of the estimated auto-correlation matrix up until time instant n, defined as:

$$r_{ij,n} = \frac{1}{n}\sum_{k=1}^{n} x_{ik}\bar{x}_{kj}. \tag{A.5}$$

Similarly to the recursion shown in Section A.1, a recursion for $\hat{r}_{ij,n}$ can be computed as

$$\begin{aligned}\hat{r}_{ij,n} &= \frac{1}{n}\sum_{k=1}^{n} x_{ik}\bar{x}_{kj} \\ &= \frac{1}{n}\left(\sum_{k=1}^{n-1} x_{ik}\bar{x}_{kj} + x_{in}\bar{x}_{nj}\right) \\ &= \frac{1}{n}\left[(n-1)\hat{r}_{ij,n-1} + x_{in}\bar{x}_{nj}\right]. \end{aligned} \tag{A.6}$$

Therefore a recursion for the auto-correlation matrix is given as

$$\hat{\mathbf{R}}_n = \left(1 - \frac{1}{n}\right)\hat{\mathbf{R}}_{n-1} + \frac{1}{n}\mathbf{x}_n\mathbf{x}_n^\top. \tag{A.7}$$

A.3 Recursive estimation of the covariance matrix

Let $\mathbf{X} \in \mathbb{R}^{m\times n}$ be a data set with n data points $\mathbf{x} \in \mathbb{R}^m$ with mean vector $\boldsymbol{\mu}_\mathbf{X} \triangleq \mathbb{E}[\mathbf{X}]$. The auto-covariance matrix of this data set is computed as in Gubner (2006, p. 335):

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_\mathbf{X} &= \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top] \\ &= \mathbb{E}[\mathbf{X}\mathbf{X}^\top - \mathbf{X}\mathbb{E}[\mathbf{X}]^\top - \mathbb{E}[\mathbf{X}]\mathbf{X}^\top + \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^\top] \\ &= \mathbb{E}[\mathbf{X}\mathbf{X}^\top] - \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^\top - \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^\top + \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^\top \\ &= \mathbf{R}_\mathbf{XX} - \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^\top, \end{aligned} \tag{A.8}$$

where $\mathbf{R}_\mathbf{XX} \triangleq \mathbb{E}[\mathbf{X}\mathbf{X}^\top]$ is the auto-correlation matrix. To find a recursion, the definition will be made in terms of the time instant n as

$$\hat{\boldsymbol{\Sigma}}_n = \mathbf{R}_n - \hat{\boldsymbol{\mu}}_n\hat{\boldsymbol{\mu}}_n^\top, \tag{A.9}$$

where the first term is given in (A.7) and the second can be computed using the mean recursion described in Section A.1, as:

$$\begin{aligned}\hat{\boldsymbol{\mu}}_n\hat{\boldsymbol{\mu}}_n^\top &= \left(\frac{n-1}{n}\hat{\boldsymbol{\mu}}_{n-1} + \frac{1}{n}\mathbf{x}_n\right)\left(\frac{n-1}{n}\hat{\boldsymbol{\mu}}_{n-1} + \frac{1}{n}\mathbf{x}_n\right)^\top \\ &= \left(\frac{n-1}{n}\right)^2 \hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top + \frac{n-1}{n^2}\left(\hat{\boldsymbol{\mu}}_{n-1}\mathbf{x}_n^\top + \mathbf{x}_n\hat{\boldsymbol{\mu}}_{n-1}^\top\right) + \frac{\mathbf{x}_n\mathbf{x}_n^\top}{n^2} \end{aligned} \tag{A.10}$$

where the first term of (A.10) can be developed as

$$\begin{aligned}\left(\frac{n-1}{n}\right)^2 \hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top &= \frac{n^2-2n+1}{n^2}\hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top \\ &= \frac{(n^2-n)+(1-n)}{n^2}\hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top \\ &= \frac{n-1}{n}\hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top - \frac{n-1}{n^2}\hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top, \end{aligned} \tag{A.11}$$

and the third term of (A.10) can be developed as

$$\begin{aligned}\frac{\mathbf{x}_n\mathbf{x}_n^\top}{n^2} &= \frac{\mathbf{x}_n\mathbf{x}_n^\top}{n^2} + \frac{\mathbf{x}_n\mathbf{x}_n^\top}{n} - \frac{\mathbf{x}_n\mathbf{x}_n^\top}{n} \\ &= \frac{\mathbf{x}_n\mathbf{x}_n^\top}{n} - \frac{n-1}{n^2}\mathbf{x}_n\mathbf{x}_n^\top. \end{aligned} \tag{A.12}$$

Replacing the developed terms (A.11) and (A.12) in (A.10), we have

$$\hat{\boldsymbol{\mu}}_n\hat{\boldsymbol{\mu}}_n^\top = \frac{n-1}{n}\hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top + \frac{\mathbf{x}_n\mathbf{x}_n^\top}{n} - \frac{n-1}{n^2}(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1})(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1})^\top. \tag{A.13}$$

Now, replacing (A.7) and (A.13) back in (A.9), we have

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_n &= \frac{n-1}{n}\hat{\mathbf{R}}_{n-1} - \frac{n-1}{n}\hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top + \frac{n-1}{n^2}(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1})(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1})^\top \\ &= \frac{n-1}{n}\left[\hat{\mathbf{R}}_{n-1} - \hat{\boldsymbol{\mu}}_{n-1}\hat{\boldsymbol{\mu}}_{n-1}^\top + \frac{1}{n}(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1})(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1})^\top\right] \end{aligned} \tag{A.14}$$

Therefore, the recursive recursion relation for updating the covariance matrix of a data stream is given as

$$\hat{\boldsymbol{\Sigma}}_n = \frac{n-1}{n}\left[\hat{\boldsymbol{\Sigma}}_{n-1} + \frac{1}{n}(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1})(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1})^\top\right]. \tag{A.15}$$

A.4 Recursive estimation of the inverse covariance matrix

Considering the recursive estimation of the covariance matrix, as shown in (A.15), and the inversion matrix lemma, also known as the Sherman-Morrison-Woodbury identity (HENDERSON; SEARLE, 1981). The lemma states that, for a nonsingular matrix $\mathbf{A}$,

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})^{-1}\mathbf{VA}^{-1}, \tag{A.16}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{U} \in \mathbb{R}^{n\times k}$, $\mathbf{C} \in \mathbb{R}^{k\times k}$ e $\mathbf{V} \in \mathbb{R}^{k\times n}$. The lemma can be applied in the recursive formula (A.15) by considering $\mathbf{A} = \hat{\boldsymbol{\Sigma}}_{n-1}$, $\mathbf{U} = \boldsymbol{\Delta}_n \triangleq \mathbf{x}_n - \hat{\boldsymbol{\mu}}_{n-1}$, $\mathbf{C} = n^{-1}$ e $\mathbf{V} = \boldsymbol{\Delta}_n^\top$, such that:

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \left(\hat{\boldsymbol{\Sigma}}_{n-1} + \boldsymbol{\Delta}_n n^{-1}\boldsymbol{\Delta}_n^\top\right)^{-1}.$$

Therefore, the recursive relation to the inverse covariance matrix is given as

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_n^{-1} &= \frac{n}{n-1}\left(\hat{\boldsymbol{\Sigma}}_{n-1} + \boldsymbol{\Delta}_n n^{-1}\boldsymbol{\Delta}_n^\top\right)^{-1} \\ &= \frac{n}{n-1}\left[\hat{\boldsymbol{\Sigma}}_{n-1}^{-1} - \hat{\boldsymbol{\Sigma}}_{n-1}^{-1}\boldsymbol{\Delta}_n\left(n + \boldsymbol{\Delta}_n^\top\hat{\boldsymbol{\Sigma}}_{n-1}^{-1}\boldsymbol{\Delta}_n\right)^{-1}\boldsymbol{\Delta}_n^\top\hat{\boldsymbol{\Sigma}}_{n-1}^{-1}\right] \\ &= \frac{n}{n-1}\left(\hat{\boldsymbol{\Sigma}}_{n-1}^{-1} - \frac{\hat{\boldsymbol{\Sigma}}_{n-1}^{-1}\boldsymbol{\Delta}_n\boldsymbol{\Delta}_n^\top\hat{\boldsymbol{\Sigma}}_{n-1}^{-1}}{n + \boldsymbol{\Delta}_n^\top\hat{\boldsymbol{\Sigma}}_{n-1}^{-1}\boldsymbol{\Delta}_n}\right), \quad \text{for } n > 1. \end{aligned} \tag{A.17}$$

APPENDIX B – THE COVARIANCE MATRIX OF A PRODUCT TRANSFORMATION

In this chapter, the covariance matrix of a random vector transformed by a constant matrix multiplication is defined.

Theorem B.1. *Let* $\mathbf{X} \in \mathbb{R}^m$ *be a column vector of* m *jointly distributed random variables with finite second moments with covariance matrix* $\mathbf{\Sigma_X} \in \mathbb{R}^{m \times m}$ *and* $\mathbf{A} \in \mathbb{R}^{n \times m}$ *a constant matrix. The covariance matrix of the transformed random variable* $\mathbf{Y} = \mathbf{AX} \in \mathbb{R}^n$ *is* $\mathbf{\Sigma_Y} = \mathbf{A\Sigma_X A}^\top \in \mathbb{R}^{n \times n}$.

Proof. From the definition, the covariance matrix $\mathbf{\Sigma_X} = \mathbb{E}\left[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top\right]$. Considering the linearity of expectation, the covariance matrix of $\mathbf{Y}$ can be computed as:

$$\begin{aligned}
\operatorname{Cov}(\mathbf{AX}) &= \mathbb{E}\left[(\mathbf{AX} - \mathbb{E}[\mathbf{AX}])(\mathbf{AX} - \mathbb{E}[\mathbf{AX}])^\top\right] \\
&= \mathbb{E}\left[(\mathbf{AX} - \mathbf{A}\mathbb{E}[\mathbf{X}])(\mathbf{AX} - \mathbf{A}\mathbb{E}[\mathbf{X}])^\top\right] \\
&= \mathbb{E}\left[\mathbf{A}(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top\mathbf{A}^\top\right] \\
&= \mathbf{A}\mathbb{E}\left[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^\top\right]\mathbf{A}^\top \\
&= \mathbf{A\Sigma_X A}^\top
\end{aligned} \tag{B.1}$$

□