Prediction Of Remaining Useful Life via Time Series Forecasting And Multi-Input Multi-Output Support Vector Machines

Hiago Henrique Gomes de Araújo, Monalisa Cristina dos Santos Moura, Isis Didier Lins, Márcio das Chagas Moura

Center for Risk Analysis Reliability and Environmental Modeling, Department of Production Engineering, Federal University of Pernambuco

Abstract

Condition-based Maintenance is one of the most modern policies discussed; it may be applied with the objective of avoiding waste of resources by extracting information about the equipment health. Diagnosis and prognostics are two essential steps in this policy, which are, respectively, when the current state of the machine is evaluated, and failures are isolated, whereas prognostics is a prediction regarding the future condition of the equipmentincluding its remaining useful life (RUL). The RUL is a critical measure since it represents the remaining function time until the equipment fails. Generally, this measurement is predicted using data-driven model (e.g. machine learning (ML)), which uses monitoring data to forecast its behavior and future state. Therefore, it is necessary extraction of data regarding an equipment health condition: temperature, vibration signals and acoustic emission are some examples of variables widely used with this objective. The Support Vector Machine (SVM) is a supervised learning algorithm used for classification and in RUL prediction context for regressive models. It is possible to predict RUL by selecting a health indicator and predict the moment at which it surpasses a pre-defined failure threshold. Hence, the RUL prediction may be treated as a time series forecasting problem. However, the standard Support Vector Regression (SVR) only works for one-dimensional outputs, difficulting the process of performing multi-step ahead forecasting of the health indicator. Thus, the SVR must be combined with forecasting strategies, and, in this paper, three of them are discussed: recursive, direct, and multi-input-multi-output (MIMO). This work proposes an approach that combines SVM with time series forecasting strategies to perform prediction of RUL. This proposed methodology was tested using a bearing vibration dataset, and results show that MIMO SVM is a feasible approach for this dataset.

1. Introduction

With the advancements in manufacturing systems and automation in industry, guarantee a reliable performance is necessary, not only to maintain the production rates but to ensure the safety of the system and its workers. Thus, an efficient maintenance policy is required, and, nowadays, a robust preventive maintenance strategy is the aim of most companies; however, it is one of the most expensive costs for industrial companies due to its complexity and high-level specifications [1]. Condition-based Maintenance (CBM), also known as predictive maintenance, is based on the analysis of data obtained by continuous equipment monitoring, and it became relevant due to advances in sensor technology and to an increased demand for integrated health management systems [2, 3].

CBM is one of the most modern maintenance policies, and it is applied with the objective to avoid waste of resources by anticipating the equipment behavior through extracting information about its current health. The first step of CBM is the diagnosis, which evaluates the current state of the equipment, isolate failures and analyze it. Then, the prognostics aim to predict the future condition of the equipment, including its Remaining Useful Life (RUL).

The bearing is an essential part of rotating machines being present in gearboxes, wind turbines, aircraft turbines, and other types of rotating equipment. Zhou, Habetler, and Haley [4] divide the condition monitoring of this equipment into seven categories: vibration, sound pressure, acoustic emission, current and laser analysis. The vibration analysis is one of the most used in this context [5]. In this paper, a bearing dataset was used to implement the interval prediction of RUL via support vector machines (SVM) and maximum entropy bootstrap (MEBOOT). SVM is a supervised learning algorithm, which is widely used in regressive models to predict the RUL [6]. RUL can be forecasted by selecting a health indicator and predicting the moment which it surpasses a pre-defined failure threshold.

The standard Support Vector Regression (SVR) is a one-dimensional method, not allowing multi-step ahead forecasting. Hence, in order to predict the health indicator using multi-step strategies, the SVR must be combined with forecasting strategies. As a proposed methodology, three of them are chosen: recursive, direct, and multi-input-multi-output (MIMO). The recursive approach is easy to compute and understand, the direct strategy is commonly more accurate than recursive, but it is computationally expensive, whereas MIMO avoids the computational cost and error propagation.

This paper proposes a method that combines SVR with MIMO. It was adopted to predict RUL of bearings using vibration signals. The ISO 10816:2009, which guides the measurement and evaluation of machine vibration [7], was used to define the failure threshold to predict vibration velocity. Particle Swarm Optimization (PSO) was used to improve SVM accuracy. A dataset from the PHM IEEE 2012 Data Challenge, fomented by the FEMTO-ST Institute, [8] was used to validate the model, and recursive and direct forecasting strategies were also applied for performance comparisons with the proposed model.

2. Theoretical Background

2.1 Forecasting Strategies

The multi-step ahead prediction of a time series $X = \{x_1, x_2, \dots, x_n\}$ aims to estimate the next H observations of X if $H > 1$ [9]. In this work, three forecasting strategies for multi-step-ahead predictions were considered: recursive, direct, and Multi-Input Multi-Output (MIMO).

The recursive strategy uses a one-step-ahead regression model to predict the values of interest. This model iteratively uses the past predicted values to provide a one-step-ahead forecast. This strategy is described as follows (Equations 1-3):

$$\hat{y}_{N+h} = \begin{cases} \hat{f}(x_N, x_{N-1}, x_{N-2}, \dots, x_{N-d+1}), & if\ h = 1 \quad (1) \\ \hat{f}(\hat{x}_{N+h-1}, \dots, \hat{x}_{N+1}, x_N, \dots, x_{N-d+1}), & if\ 2 \geq h \geq d \quad (2) \\ \hat{f}(\hat{x}_{N+h-1}, \hat{x}_{N+h-2}, \dots, \hat{x}_{N+h-d}), & if\ d+1 \geq h \geq H \quad (3) \end{cases}$$

Where h is the predicted step, and d is the autoregressive order. The recursive strategy is simple, easy to compute and understand; however, the prediction over predicts incurs on fast error propagation. In contrast to the recursive approach, the direct strategy consists of training a model for each future value x_{T+h} [10] (Equation 4):

$$\hat{x}_{t+h} = f_h(x_t, x_{t-1}, x_{t-2}, \dots, x_{t-d+1}), \qquad \text{with } 1 \geq \text{h} \geq \text{H} \tag{4}$$

Differently from the recursive strategy, the error of each prediction is not directly accumulated by the others. On the other hand, the building of H regression models incurs a significant increase in the computational cost. Furthermore, the estimates are independent of each other since all forecasting steps are computed individually.

Unlike the direct strategy, the MIMO approach avoids the assumption of conditional independence of future values, and it also prevents the error propagation observed in the recursive procedure [11]. Thus, the MIMO strategy preserves the relationship between the predicted values of a time series. The main idea of the MIMO approach is to find a function $f: \mathbb{R}^d \rightarrow \mathbb{R}^H$ that the forecasts of H future values could be directly obtained from a single autoregressive model based on d past values of the variable being considered (Equation 5):

$$\hat{x} = [\hat{x}_{N+H}, \hat{x}_{N+H-1}, \dots, \hat{x}_{N+1}] = f(x_N, x_{N-1}, \dots, x_{N-d+1}) \quad (5)$$

2.2 Support Vector Regression

SVM is a supervised machine learning technique introduced in 1995 [12], which uses data to construct the model, and is composed by two kinds of variables: the output or dependent (y) and the multi-dimensional inputs (x). The objective of this algorithm is to find a pattern to describe the dependence of the values of y within the values of x. SVR is the version of SVM used to solve a regression problem. Firstly, the training data set is used to solve the following convex minimization problem (Schölkopf and Smola (2002)) (Equations 6-10):

$$\min_{w,b,\xi,\xi^*} \frac{1}{2} w^T w + C * \sum_{i=0}^{l} (\xi_i + \xi_i^*) \quad (6)$$

Subject to:

$$y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i \quad (7)$$

$$w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \quad (8)$$

$$\xi_i \geq 0 \quad (9)$$

$$\xi_i^* \geq 0 \quad (10)$$

The objective function is a sum of two components, the first one is the capacity of the regression function, the smaller its value is, the more general will be the solutions found. The second part of the sum is associated with the difference between the original values and predicted values. The model restrictions show that the values of ξ_i and ξ_i^* are equal to the distance between y and a ε-insensitive tube around the constructed regression (**Error! Reference source not found.**), which is obtained by Equation 11:

$$\hat{y} = f(x) = w^T \phi(x) + b \quad (11)$$

In which $\phi(x)$ is a nonlinear transformation to the feature space. However, the scalar product $\Phi^T(x)\Phi(x)$ used in the solution of SVM demands a big computational effort, especially when x has a large number of dimensions. To solve this problem, kernel functions $K(x_i, x_j) = \Phi^T(x_i)\Phi(x_j)$ are applied to define a good mapping of the product, avoiding the hard computing of this value. Therefore, the regression function with kernel is found (2.10), based on the solution of the dual problem (2.11-2.15), which is obtained by the Karush-Kuhn-Tucker (KKT) resolution (Equations 12-17).

$$f(x) = \sum_i (\alpha_i - \alpha_i^*)\, K(x_i, x_j) + b \tag{12}$$

$$\max_{\alpha,\alpha^*} L_D = -\frac{1}{2}\sum_i \sum_j (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) x_j^T x_j - \sum_i (\varepsilon - y_i)\alpha_i - \sum_i (\varepsilon + y_i)\alpha_i^*, i, j = 1,2, \dots, l \tag{13}$$

subject to: (14)

$$\sum_i (\alpha_i - \alpha_j^*) = 0 \tag{15}$$

$$0 \leq \alpha_i \leq C \tag{16}$$

$$0 \leq \alpha_i^* \leq C \tag{17}$$

The most popular kernel function used on SVM is the Radial Basis Function (RBF), used in this work, which is given by Equation 18:

$$K(x_i, x_j) = exp(-\gamma \|x_i - x_j\|^2) \tag{18}$$

2.3 Multi-Dimensional Support Vector Regression

The difference between a simple and a multi-regression problem is the dimension of the output variable. In the second case, y is a vector in the $\mathbb{R}^H$ Space. This kind of problem is present in many applications, including the multi-step ahead time series forecasting using the MIMO strategy. Although, the standard SVM is not capable of doing this kind of regression, and to solve this limitation, [13] proposed a generalization of SVR, namely Multi-Dimension Support Vector Regression (MSVR).

The insensitive zones defined around the estimates are a significant difference between the two models. In MSVR, a novel approach is proposed in which a "hyper-spherical insensitive zone will allow us to equally treat every sample, being them penalized by the same factor if they lie outside this insensitive zone" [13]. In SVR, if a value is out of the "tube" defined by the insensitive function, then it is penalized by C and it is not otherwise. For MSVR, the penalization in the objective function is proportional to the number of dimensions where the sample is further than ε. The minimization problem related to the training phase of an MSVR presented below (Equations 19-21) [14]:

$$\min_{W,\ b} L_P = \frac{1}{2}\sum_{h=1}^{H} \|w_j\|^2 + C\sum_{i=1}^{n} L(u_i), \tag{19}$$

in which:

$$L(u) = \begin{cases} 0, & u \leq \varepsilon \\ u^2 - 2u\varepsilon + \varepsilon^2, & u > \varepsilon \end{cases} \tag{20}$$

$$u_i = \|y_i^T \phi^T(x_i) W - b^T\| \tag{21}$$

2.4 Particle Swarm Optimization

PSO is optimization method for nonlinear and continuous functions proposed by Eberhart and Kennedy [15], the construction of this metaheuristic was based on studies that simulated the motion of groups of animals such as flocks of birds and schools of fishes. The main idea is that the particles move through the search space based on the best position that they have individually found and the best position found by their neighbor particles.

The information about a particle is made up of three vectors, the position of the particle in the space $x_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$, the best position found by the particle individually $p_i = (p_{i1}, p_{i1}, \ldots, p_{iD})$, and its velocity $v_i = (v_{i1}, v_{i1}, \ldots, v_{iD})$. Bratton and Kennedy [16] define a standard for the parameters of the PSO. According to them the best topology is the local communication following a ring model, in which the particle neighborhood consists of two other particles. The proposed number of particles is 50, the initialization of the swarm must be non-uniform and the objective function is not evaluated when the particles are out of the boundary conditions.

At each iteration the velocity and current position of each particle are updated based on the following equations:

$$v_{id} = \chi[v_{id} + c_1\epsilon_1(p_{id} - x_{id}) + c_2\varepsilon_2(p_{gd} - x_{id})] \tag{22}$$

$$x_{id+1} = x_{id} + v_{id} \tag{23}$$

In this study the PSO was used to select the SVR and MSVR hyperparameters the data was separated in validation and a training set, the algorithm searches for the values that minimize the Normalized Mean Root Square error (NMRSE), which is a commonly used error function to this objective.

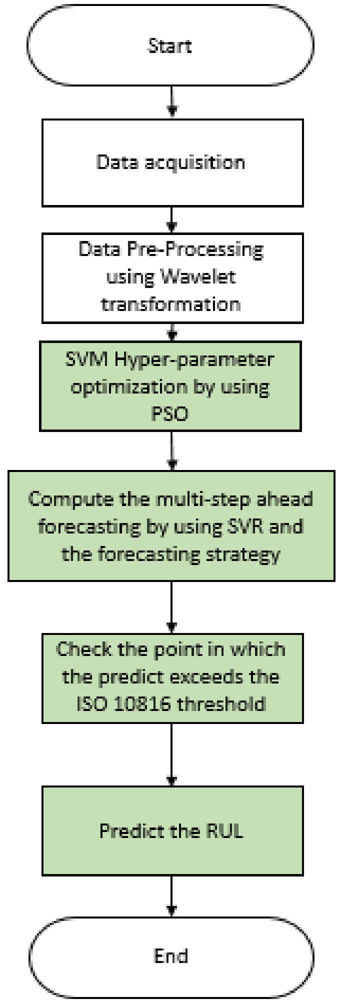
$$NMRSE = \sqrt{\frac{\sum_{i=1}^{\lambda}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{\lambda} y_i}}, \tag{24}$$

Where $(y_i - \hat{y}_i)$ is the difference between the real value in the validation data set and its prediction, and λ is the size of the validation data set. The main objective of the PSO search is to find the values of C, ε and σ that minimizes this objective function.

3. Methodology

The proposed methodology is presented in Figure 1, the green boxes represent the proposed steps of this paper. More details about data acquisition can be found in Nextoux *et al*, [17] and about the wavelet transform applied to pre-process the used time series in Santos [18].

Figure 1 - Flowchart of the applied method.



Source: This research.

The algorithms used to find the results were implemented using R. The package ksvm [19] was used to run the standard SVR, which was used to implement direct and recursive forecasting strategies. The Multi-Dimensional SVR optimization problem was solved by using an iterative procedure as shown in [20], also implemented in R.

The vibration data set used to test the proposed methodology was from the PHM IEEE 2012 Data Challenge using the PRONOSTIA platform [8]. PRONOSTIA is an experimentation platform dedicated to test and validate bearings prognostic methodologies [17]. This equipment provided an experimental data that characterize the degradation of real ball bearings along their useful life, and it is divided into three main parts: the rotating part (motor and its shafts), the degradation part that generates a load over the equipment, and the measurement part that has sensors of vibration and temperature signals.

The search for the best SVM hyperparameters was done by using the standard PSO [16]. Firstly, the last 50 points of the training dataset were selected to be the validation set. At each fitness evaluation in PSO, the SVR or MSVR model is trained by using the hyperparameters and the remaining training dataset. Thus, the prediction is made to validation dataset and the (NRMSE) is calculated, the algorithm search for a set of hyperparameters that minimizes this measure.

4. Results

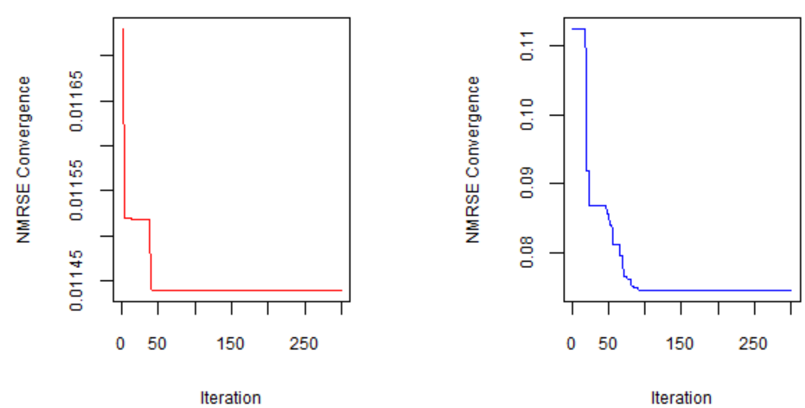
After data acquisition, the hyperparameters of the model is selected using PSO. Thus, 300 iterations of the PSO algorithm were made to find the best C, ε, and σ to perform the prediction by proposed strategies. According to the PSO results (Figure 2), there was a satisfactory convergence to the selected hyper-parameter values (Table 1). To the direct strategy, the PSO algorithm was executed for each of the *H* used models; the corresponding hyperparameter values are listed in table 2.

Table 1 - MIMO and recursive forecasting hyperparameter selection.

Strategy	C	ε	σ
MIMO Strategy	399.9876230	0.2048704	3.6926108
Recursive Strategy	403.70144015	0.06736573	0.85978630

Source: This research.

Figure 1 - Particle Swarm Optimization Convergence of NMRSE to MIMO and Direct Strategy, respectively.



Source: This research.

Table 2 - PSO Results to direct forecasting.

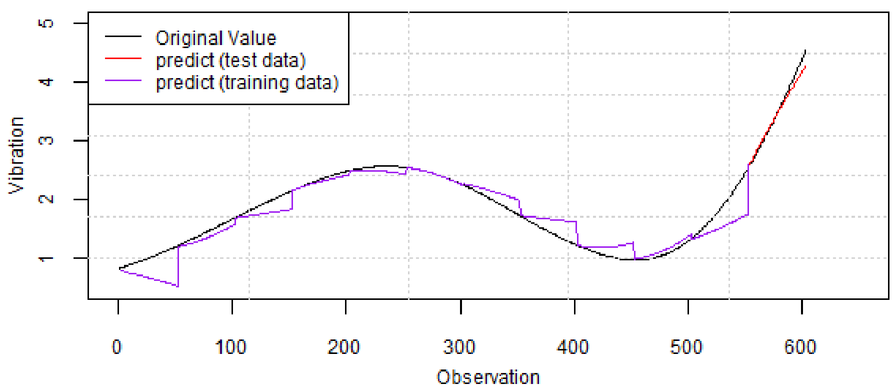
h	C	ε	σ	*h*	C	ε	σ
1	3831,335	0,137	2,497	**26**	29,418	0,203	2,728
2	0,276	0,237	2,649	**27**	36,610	0,139	2,676
3	508,265	0,097	2,735	**28**	24,782	0,261	2,760
4	0,689	0,123	3,224	**29**	27,071	0,111	2,727
5	96,205	0,114	1,826	**30**	28,389	0,229	2,746
6	2,231	0,264	2,038	**31**	29,181	0,125	2,782
7	25,038	0,095	2,620	**32**	31,336	0,270	2,719
8	29,736	0,044	2,645	**33**	29,822	0,107	2,782
9	27,865	0,142	2,667	**34**	29,135	0,182	2,732
10	37,415	0,220	2,608	**35**	28,881	0,170	2,681
11	19,467	0,068	2,736	**36**	38,161	0,149	2,681
12	19,417	0,131	2,769	**37**	34,610	0,118	2,702
13	18,922	0,037	2,839	**38**	38,565	0,048	2,686

14	13,936	0,019	2,754	**39**	35,369	0,130	2,758
15	15,681	0,075	2,731	**40**	18,479	0,086	2,803
16	22,924	0,170	2,788	**41**	33,518	0,141	2,668
17	18,409	0,261	2,834	**42**	39,016	0,053	2,670
18	29,623	0,203	2,748	**43**	30,091	0,246	2,722
19	33,763	0,152	2,741	**44**	38,759	0,160	2,667
20	19,241	0,070	2,730	**45**	34,089	0,131	2,664
21	18,312	0,166	2,742	**46**	28,673	0,263	2,711
22	28,092	0,114	2,673	**47**	47,014	0,245	2,578
23	32,004	0,141	2,729	**48**	38,527	0,191	2,588
24	28,484	0,166	2,781	**49**	43,370	0,198	2,588
25	31,703	0,075	2,776	**50**	33,004	0,144	2,644

Source: This research.

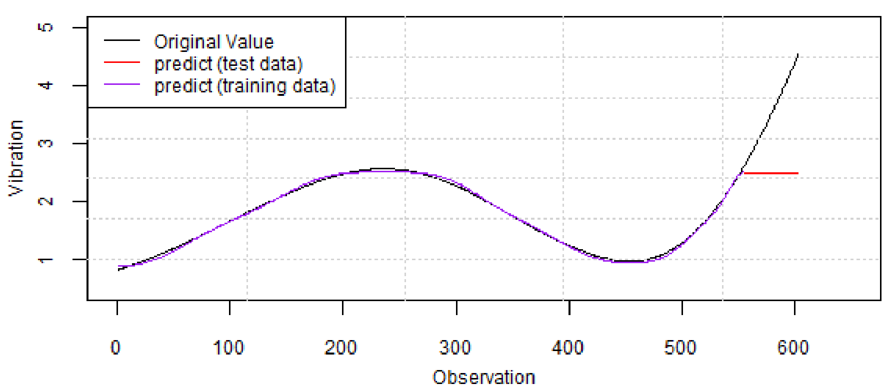
After selecting the models with PSO, a 50-step ahead prediction was obtained for each of the analyzed forecasting strategies. The result is presented in figures 3 to 5.

Figure 3 - MIMO Vibration Forecasting.



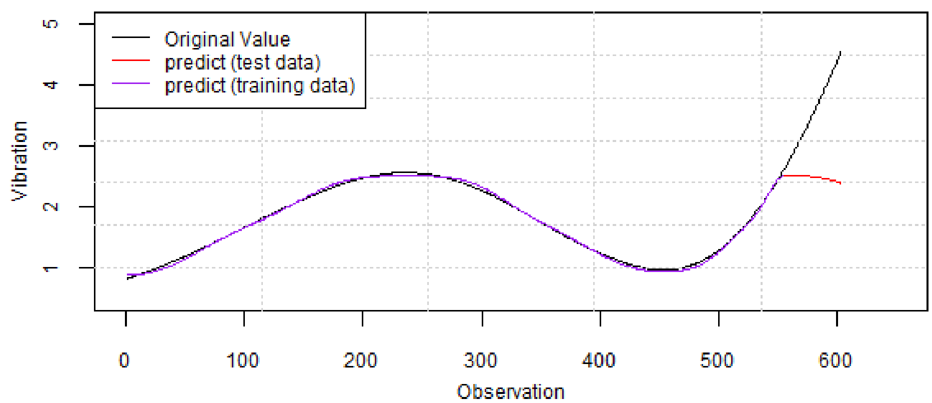
Source: This research.

Figure 4 - Recursive Vibration forecasting.



Source: This research.

Figure 5 - Direct time series Vibration forecasting.



Source: This research.

It is possible to conclude that only the MIMO strategy had a good trade-off between training error and accuracy compared to the test. To other strategies, the regression values of the training data were almost equal to the original time series, but the forecasting of the testing data set was extremely inaccurate. The values presented in Table 3 show that the Mean Average Error (MAE) and Mean Average Percentual Error (MAPE) were the smallest for the MIMO strategy. Thus, the MIMO is the best of the three applied strategies to all prediction horizons

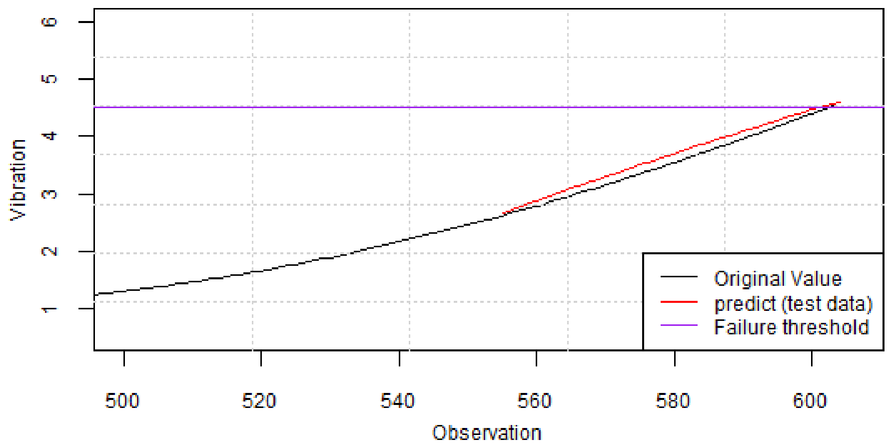
Table 3 - Forecasting MAE and MAPE.

		Forecasting horizon				
Strategy		**1-3**	**4-15**	**16-30**	**30-45**	**45-50**
Direct	**MAE**	0,10693	0,36752	0,8831	1,546667	2,0309
	MAPE	4,0723 %	12,607 %	25,86 %	38,467 %	45,741 %
Recursive	**MAE**	0,11767	0,38559	0,89396	1,5112	1,9450
	MAPE	4,480 %	13,23 %	26,202 %	37,602	43,809 %
MIMO	**MAE**	0,02598	0,04011	0,020536	0,10849	0,22944
	MAPE	0,991 %	1,392 %	0,616 %	2,662 %	5,163 %

Source: This research.

After computing the accuracy measurements, the RUL was predicted by identifying in which point of time the forecast exceeds the degradation threshold defined by the ISO 10816 (4.5 mm/s). The results are exposed in Figure 6 and in Table 5. The direct and recursive strategies were not accurate enough to reach the failure threshold. Therefore, the Average Percentage Error (APE) of the RUL prediction using the MIMO strategy is 0,19%. The point predict of the RUL was accurate, but it was made with only 490 seconds before the event happened. Furthermore, the proposed approach is dependent on the used health indicator, so it is necessary to evaluate if the pre-defined boundary really represents if the equipment is on the verge of failure.

Figure 6 – RUL Prediction Via Mimo SVM and Failure Threshold.



Source: This research.

Table 5 - RUL predicts Accuracy with MIMO strategy.

RUL Predict	Real RUL	APE
21980 s	22020 s	0,19%

Source: This research.

5. Conclusions

In this work, the ISO 10816 was used to guide the definition of the vibration velocity as a health indicator and its corresponding degradation boundaries. Furthermore, the MIMO, direct and recursive time series forecasting strategies were used jointly with SVM to provide multi-step ahead forecasting of vibration velocity.

The results show that MIMO-SVM is the better alternative to predict bearing RUL. The corresponding APE was only 1.65%. The results were not favorable to the use of direct or recursive strategies as they did not attain the predefined vibration velocity threshold. Although, the evidence shows that the model selection for the other two strategies was led to an overfitting model this indicates the need of testing other metaheuristics to search for the best SVM hyperparameters such as the fruit fly optimization algorithm, independence cohort intelligence and genetic algorithm [21-23]. Furthermore, the complexity of the used time series also indicates the necessity of developing a different health indicator and failure threshold as shown by Soualhi, Medjer, and Zerhouni [24]

Another important limitation of the present study is that interval estimates were not considered. Thus, the decision-maker does not have option to evaluate the dispersion of the model output. One solution to this problem is the use of non-parametric approaches such as bootstrap [25]. The uncertainty analysis can include the use of interval along with point estimates, to provide a measure of the variability of the estimates.

The proposed method of accuracy can be tested in other PHM contexts; it could be applied in other bearings data sets. For example, NASA bearing fault data set [26] or Case Western Reserve university bearing data set. The MIMO-SVM forecasting strategy for RUL prediction can also be tested in data sets related to other equipment such as batteries, although it would be necessary to change the health indicator and failure threshold.

References

[1] 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, p. 1483, (2006).

[2] AHMAD, R.; KAMARUDDIN, S. "An overview of time-based and condition-based maintenance in industrial application". *Computers & Industrial Engineering*, v. 63, n. 1, p. 135, (2012).

[3] KHAN, S.; YAIRI, T.. "A review on the application of deep learning in system health management". *Mechanical Systems and Signal Processing*, v. 107, p. 241, (2018).

[4] ZHOU, W., HABETLER, T.G., HARLEY, R.G. "Bearing condition monitoring methods for electric machines: A general review". *Diagnostics for Electric Machines, Power Electronics and Drives*. IEEE, 2007, SDEMPED, p. 3. (2007)

[5] ORHAN, Sadettin; AKTÜRK, Nizami; CELIK, Veli. "Vibration monitoring for defect diagnosis of rolling element bearings as a predictive maintenance tool: Comprehensive case studies". *Ndt & E International*, v. 39, n. 4, p. 293, (2006).

[6] ZHANG, L., LIU, Z., LUO, D., LI, J., and HUANG, H. "Review of remaining useful life prediction using support vector machine for engineering assets". *In 2013 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE)*, (IEEE), p. 1793. (2013).

[7] ISO 10816-1: 2009 (E) – Mechanical vibration – Evaluation of machines vibration by measuments on non-rotating parts – Part 1: General Guidelines.

[8] IEEE PHM 2012 Prognostic Challenge. Scoring of results and application procedure. *Web page: http://www.femto-st.fr/f/d/IEEE-Challenge-Appli.pdf*. (2012)

[9] TAIEB, S. B. "A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition". *Expert systems with applications*, v. 39, n. 8, p. 7067, (2012).

[10] SORJAAMA A, HAO J, REYHANI N, Ji Y, LENDASSE A. "Methodology for long-term prediction of time series". *Neurocomputing*. v.10, p.2861, (2007).

[11] BONTEMPI, G.; TAIEB, S.B.; LE BORGNE, Y.A.. "Machine learning strategies for time series forecasting". In: *European Business Intelligence Summer School*. Springer, Berlin, Heidelberg. p. 62. 2012

[12] VAPNIK, V., *The Nature of Statistical Learning Theory*. Springer (Vol. 8), 2000.

[13] PÉREZ-CRUZ F., CAMPS-VALLS G., SORIA-OLIVAS E., J.J. PÉREZ-RUIXO, A.R. FIGUEIRAS-VIDAL, A. Artés-Rodríguez, "Multi-dimensional function approximation and regression estimation". In *Artificial Neural Networks—ICANN 2002 Springer Berlin*, Heidelberg, p. 757. (2002).

[14] SÁNCHEZ-FERNÁNDEZ, M., DE-PRADO-CUMPLIDO, M., ARENAS-GARCÍA, J., & PÉREZ-CRUZ, F. "SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems". *IEEE transactions on signal processing*, v. *52*, n. 8, p. 2298-2307. (2004)

[15] EBERHART, R. KENNEDY, J.,, "Particle swarm optimization". *Proceedings, IEEE International Conference*, v. 4, p. 1942, (1995).

[16] NECTOUX, Patrick et al. "PRONOSTIA: An experimental platform for bearings accelerated degradation tests". In: *IEEE International Conference on Prognostics and Health Management, PHM'12***.** p. 1-8. (2012).

[17] BRATTON, D., & KENNEDY, J., "Defining a Standard for Particle Swarm Optimization". 2007 *IEEE Swarm Intelligence Symposium*, p. 120, (2007).

[18] SANTOS, M, C, M. Metodologia de prognóstico de falha de sistemas via support vector machines e técnicas de pré-processamento de dados. Trabalho de conclusão de curso – Departamento de Engenharia de Produção, Universidade Federal de Pernambuco, Pernambuco.2017

[19] KARATZOGLOU, A., SMOLA, A., HORNIK, K., & ZEILEIS, A. "kernlab-an S4 package for kernel methods in R". *Journal of statistical software*, v. 11, n. 9, p. 1. (2004)

[20] TUIA, D., VERRELST, J., ALONSO, L., PÉREZ-CRUZ, F., & CAMPS-VALLS, G. "Multioutput support vector regression for remote sensing biophysical parameter estimation". *IEEE Geoscience and Remote Sensing Letters*, v.8, n.4, p. 804. (2011).

[21] LIJUAN, W., & GUOHUA, C. Seasonal SVR with FOA algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow. *Knowledge-Based Systems*, v. 110, p. 157. (2016).

[22] ALADEEMY, M.; TUTUN, S.; KHASAWNEH, M.T. A new hybrid approach for feature selection and support vector machine model selection based on self-adaptive cohort intelligence. *Expert Systems with Applications*, v. 88, p. 118-131, (2017).

[23] ALANTARY, D.; YALKOWSKY, S. "Comments on prediction of the aqueous solubility using the general solubility equation (GSE) versus a genetic algorithm and a support vector machine model". *Pharmaceutical development and technology*, v. 23, n. 7, p. 739, (2018).

[24] SOUALHI, A., MEDJHER, K., ZERHOUNI, N., "Bearing Health Monitoring Based on Hilbert-Huang Transform, Support Vector Machine, and Regression". *IEEE Transaction on Instrumentation and Measurement***,** vol. 64, n. 1, 2015.

[25] EFRON, Bradley; TIBSHIRANI, Robert J. *An introduction to the bootstrap*. CRC press, 1994.

[26] LEE, J. QIU, H. YU, G. LIN, J. REXNORD TECHNICAL SERVICES. 'Bearing Data Set', IMS, University of Cincinnati. NASA Ames Prognostics Data Repository. 2007