

Data Analytics for Predictive Maintenance of Industrial Robots

Tawfik Borgi*, Adel Hidri†, Benjamin Neef‡, Mohamed Saber Naceur*

* Université de Tunis El Manar

Ecole Nationale d'Ingénieurs de Tunis, LTSIRS Laboratory, 1002, Tunis, Tunisia

Email: tawfik.borgi@gmail.com

† Université de Tunis El Manar

Ecole Nationale d'Ingénieurs de Tunis, LR-SITI Laboratory, 1002, Tunis, Tunisia

‡ Chair of Sustainable Manufacturing and Life Cycle Engineering Research

Institute of Machine Tools and Production Technology

Technische Universität Braunschweig, Langer Kamp 19b, 38106 Braunschweig, Germany

Abstract—The predictive maintenance of industrial machines is one of the challenging applications in the new era of Industry 4.0. Thanks to the predictive capabilities offered by the emerging smart data analytics, data-driven approaches for condition monitoring are becoming widely used for early detection of anomalies on production machines. The aim of this paper is to provide insights on the predictive maintenance of industrial robots and the possibility of building a condition-monitoring system based on the data analysis of robot's power measurements. A predictive modeling approach is proposed to detect robot manipulator accuracy errors based on robot's current data analysis for predictive maintenance purposes. An experimental procedure is also carried out to oversee the correlation between the robot accuracy error and a set of extracted features from current time-series, and to evaluate the proposed predictive modeling. The obtained results are satisfactory and prove the feasibility of building a data-driven condition monitoring of robot manipulators using the electrical power time-series data analysis.

Keywords- Data Analytics, Predictive maintenance, Robot manipulators

I. INTRODUCTION

In modern digital manufacturing environments, huge sets of voluminous data are massively generated from many involved areas and sources. These continuously growing data contain useful knowledge and information that could enhance the productivity of the whole manufacturing process, as a valuable asset for decision support in various areas, particularly in health monitoring and condition-based maintenance.

Condition-Based Maintenance techniques have been developed to allow scheduling maintenance actions based on machine's condition measured without the interruption of the normal machine operations. In fact, Condition Monitoring is a decision making strategy that allows real-time diagnosis of occurring failures and prognosis of future asset and machines/equipment health by continuous observation of the system and its components' condition [1]. It is strongly related to the concept of Prognostic and Health Monitoring (PHM) which has been initially introduced in medical field for disease and epidemiology prediction, but has been widely applied in manufacturing con-

text [2]. Prognosis consists in forecasting the likely outcome of a situation, and the straightforward application of PHM in the field of maintenance is the prediction of the Remaining Useful Life (RUL) of production machines. The recent advances of computerized control, information techniques and communication networks have made possible collecting huge volumes of operation and process's conditions data to be harvested in making an automated Fault Detection and Diagnosis (FDD) [3], and developing more efficient approaches for intelligent preventive maintenance activities, also termed as predictive maintenance [4].

Data-driven models for failure detection and predictive maintenance are generally based on data analysis and statistical techniques, such as the machine learning algorithms. Based on a huge historical data, predictive models can be built in the training stage, where relationships between data records and outputs are examined to generate a data-driven model. In the test stage, any new data will be evaluated using this model for decision support. The result of the decision making in the context of the predictive maintenance can generally be whether predicting the RUL of the asset or predicting the occurrence and type of the failure.

In many industrial and manufacturing environments, robots manipulators are key production machines and adopting successful strategies for faults detection and isolation in these tools is widely desired since it might prevent big loss of productivity [5]. In fact, industrial robots are manipulators designed to move parts and materials and perform various programmed tasks in production and manufacturing settings. Based on the analysis of the robot mechanical standards and the used technology of the electrical control system, many different reasons can be attributed to faults in these industrial robots. As an example, brakes failures are usually due to wear [6], while motor faults are generally due to short circuit [7]. Since the pattern of evolution of data and the time of this evolution vary across all these faults types, it is not easy to combine all these faults into a single fault-detection framework especially when a single model is being used. Therefore, data-driven methods can be used as an alternative black-box

approach for failure detection in robot manipulators.

Eski et al.[8] deployed artificial neural networks with noise measurement and joint accelerometer data, to make vibration prediction on industrial robots. Bittencourt et al.[6] presented a data-driven approach for condition monitoring of systems operating in repetitive manner. The method considers changes in the distribution of data samples that have been obtained from several executions of one or multiple tasks, using the Kullback-Leibler divergence. Sathish et al.[9] proposed a failure detection approach for industrial robots based on the Principal Component Analysis (PCA) method. The historic data used in this approach are time-series data of up to 60 days of normal operation with some failure records collected from industrial robots.

In general, in any data-driven predictive maintenance solution, the key element of the whole approach is the data which have to be voluminous and contain records of damage periods in addition to the normal operation of the asset. Otherwise, maintenance actions must take place during the acquisition of the training data in order to make possible observing the degradation patterns when adopting anomaly detection approaches. From this perspective, classical data-driven approaches cannot be adopted in the case where no failure data are available in the historic data. The purpose of this paper is to draft a methodology for a data-driven condition monitoring of robot manipulator without historic failure data. The proposed approach aims at detecting the robot accuracy error based on electricity data analysis for predictive maintenance. The paper is organized as follows : section 2 specifies the accuracy issues of robot manipulators and introduces the proposed solution. Section 3 describes the proposed predictive modeling approach. The experimental results are given in section 4. Finally, this paper is ended with conclusions.

II. ROBOT ACCURACY ERROR DETECTION

The precision performance of manipulator can be described by two main values specified in the international standard ISO 9283 which defines the performance criteria of robot manipulators. These two values are the pose accuracy and pose repeatability. The pose repeatability measures the ability of the robot to move back to the same orientation and position, while the pose accuracy describes the ability of precisely moving to a desired position in the 3-dimensional space [10]. Besides the assembly or mechanical issues, errors in pose accuracy are generally affected by geometric and non-geometric parameters and mainly caused by defects in the kinematic model [10]. Robot Calibration process aims at establishing manipulators parameters' values that affect the accuracy and improving it by performing modifications in the positioning software [10]. This process consists of modeling, measurement, parameter identification and finally implementation of compensation. As a part of the calibration process, an accurate external measuring of the end-effector Cartesian pose of each set of the joint values is performed in various positions, usually deploying a Laser Tracker device[11].

However, the calibration procedure is a complex process that requires a specific knowledge and expertise in this field. In addition, laser tracker systems are expensive and their deployment for a continuous check and accuracy measurement might be very costly in terms of time and productivity. As an alternative, a smart-data approach is proposed in this context for condition monitoring and signal-based fault detection of manipulators to oversee the accuracy degradation through power time-series analysis and without using external position and accuracy measurement methods.

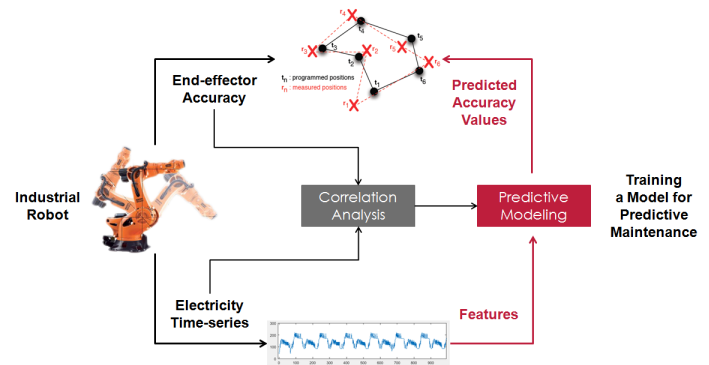


Fig. 1. Accuracy errors detection methodology

The proposed approach for accuracy degradation detection can be established by finding a correlation between the electrical signal measured from the robot and its accuracy values (see figure 1). Normally, the correlation that might exist between the fault (accuracy errors) and the used signal (the measured electrical signal) is determined based on a prior knowledge of the system and the accuracy performances of manipulators. However, we propose in this paper to find this correlation using a data-driven approach based on a predictive modeling methodology. We present in the following section the proposed predictive modeling approach.

III. PROPOSED PREDICTIVE MODELING APPROACH

In order to use the accuracy measurements in a signal-based FDD solution for predictive maintenance, we need to make an assumption about the existence of a certain correlation between the robot's accuracy error and some patterns in the measured electrical signals. In order to verify this hypothesis, we propose to build a model to predict accuracy errors by analyzing the electricity time-series data. To achieve this goal, we propose to build a regression model between features extracted from electricity time-series data and externally measured accuracy errors of the robot. The general flow of the proposed methodology is exposed in the block diagram of figure 2. The raw data used in this approach consist of the electrical power measured from the robot and a set of positions of its end-effector while performing a repetitive action during defined periods of time. These data undergo preprocessing and feature engineering operations to clean the data and extract the relevant descriptors and aggregate them into a single training set to train the predictive model. Once the model trained, only

electricity data will be used as input data, and the accuracy errors will be the outputs of the test stage. We describe these steps in details in the following sections.

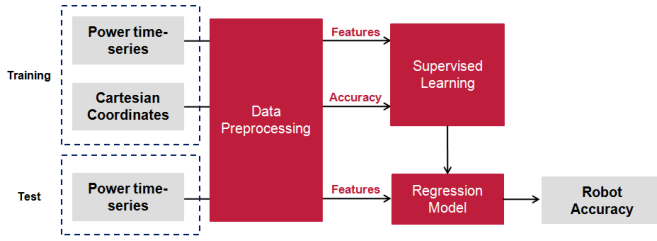


Fig. 2. Block diagram of the general approach

A. Data Acquisition and Preprocessing

The accuracy errors of the robot are not directly acquired, since no accuracy measurement system is available. Instead, the raw data in that case are the spatial positions of the robot's end effector being measured using an external position tracking device. For a given repetitive motion of the robot, a set of positions of the end-effector can be gathered in a first stage to mathematically model the trajectory of the programmed movement as illustrated in figure 3. Based on this model considered as a reference model, the trend of the accuracy errors can be measured by calculating the residuals between new positions and the model defined by reference positions.

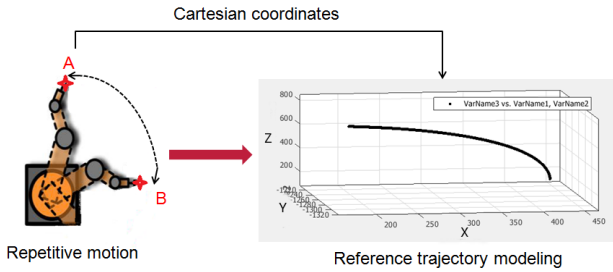


Fig. 3. Robot positions modeling.

On the other hand, the three phases of current in the robot PLC are measured during its operation using current clamps, and a set of time-series data is acquired. First data cleaning operations are performed to remove the non-relevant records (captured right before the robot starts moving or right after it stops) and only keep the needed values [12] (when the robot is performing the repetitive motion). Once the data cleaned, the power and positions data sets must be segmented into multiple sub-sets where each of them corresponds to some similar patterns related to the repetitive movement of robot. From each of these patterns of the electricity signal data, a set of features is extracted. On the other hand, for each sub-set of the Cartesian coordinates, a single accuracy value is calculated.

B. Feature Engineering

We propose to extract first a set of features from the electrical signal of the robot and analyze their correlation with

the accuracy error in order to only keep the relevant ones. We propose to use the following features : Mean, Standard deviation, Kurtosis, Skewness, Periodicity (Long period and Short period) and Energy [13].

• Mean and Standard Deviation:

These two measures are useful and simple statistical features. Mean (Eq. 1) describes the central tendency of the data while the Standard Deviation (Eq. 2) estimates the dispersion by showing the relation of the values to the mean of the data sample :

$$m = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - m)^2} \quad (2)$$

• Skewness :

Skewness can be described as the measure of the symmetry of a data set around the mean. It is defined as :

$$S = \frac{E(x - m)^3}{\sigma^3} \quad (3)$$

• Kurtosis :

Kurtosis describes the peakedness and how outlier-prone the distribution of a dataset is. It is defined as:

$$K = \frac{E(x - m)^4}{\sigma^4} \quad (4)$$

• Periodicity :

The periodicity is an important feature which gives insight about the seasonality and emphasizes the cyclic pattern of the time series. Autocorrelation is used to verify the existence of periods and determine their duration. The Autocorrelation is defined for different τ lags as :

$$ACF(\tau) = \frac{1}{N} \sum_{i=1}^N x(\tau) \cdot x(\tau + i) \quad (5)$$

We propose in this paper to pick the two first periods that we call the Long Period and the Short Period and use them as two different features.

• Energy :

The periodicity of the signal is also reflected in the frequency domain. Based on Parseval's theorem, the energy can also be defined in the discrete-time domain as follows :

$$Energy = \sum_{i=1}^N |x(i)|^2 \quad (6)$$

The collected data of the robot end point positions must contain a huge number of Cartesian coordinates records. All these recorded three-dimensional points refer to the positions crossed by the end effector during its movement. Hence, a reference trajectory must be first represented by transforming all these spatial coordinates into a mathematical model.

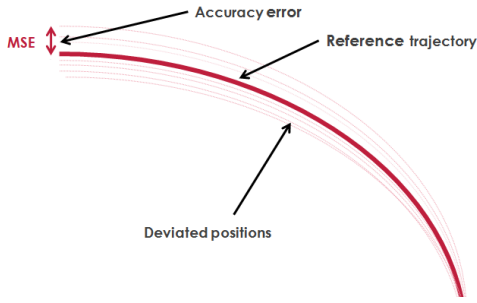


Fig. 4. Reference trajectory and deviated positions

Based on this model, the accuracy can be obtained by comparing the measured positions with the reference trajectory as illustrated in figure 4. The accuracy value is given by the mean square error (MSE) between programmed positions and measured positions :

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2 \quad (7)$$

C. Predictive Model and Correlation Analysis

We propose to use in our approach the *multiple linear regression* as a data model to predict accuracy values based on the extracted electrical features. A linear regression model with more than one independent or predictor variable is called a multiple linear regression model. For a vector X of observations on predictor variables, and a vector Y of responses, the general form of this model can be expressed as follows:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \epsilon_i, i = 1, \dots, N \quad (8)$$

where Y_i is the i^{th} response, x_{ij} is the j^{th} predictor variable measured for the i^{th} observation, β_k is the k^{th} coefficient and ϵ_i is the i^{th} noise term which is a random error.

Fitting this data model consists in estimating the values of β . Using the least squares approach in this context, the objective becomes to minimize the following :

$$\sum_i (Y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \dots - \beta_p x_{ip})^2 \quad (9)$$

over all the possible values of the coefficients.

However, before modeling this relationship between predictors and response, it is usually more convenient to perform a correlation analysis in order to verify the existence of a linear relationship between all pairs of used variables. The correlation coefficients, as a measure of the linear dependance between two random variables, are calculated for all the variables of the model. For each variable having N observations, the Pearson correlation coefficient between two variables A and B can be defined in terms of the covariance of A and B as follows :

$$\rho(A, B) = \frac{cov(A, B)}{\sigma_A \sigma_B} \quad (10)$$

where σ_A and σ_B are the standard deviation of A and B . A correlation matrix is then calculated to present the correlation coefficients for each pairwise variable combination. By examining this matrix, where each of its columns represents a separate quantity, the variables with strong linear relationship can be identified.

IV. EXPERIMENTAL PROCEDURE AND RESULTS

An experimental procedure was proposed in this paper to test and validate our approach and present the needed experimental results. The industrial robot used during this project is a KUKA KR 6-2 robot manipulator jointed-arm kinematics for continuous and point-to-point path tasks. It is a six-axis industrial robot powered by AC servomotors, with installed motor capacity of 8.8kW and Repeatability of 0.05 mm [14]. We also used an LTD800 Leica Laser Tracker. It is a coordinate measurement instrument used in many industrial fields, enabling 6DOF tracking for handheld and non-contact 3D-scanning [15]. The robot end-effector positions tracking using the Leica laser tracker device is illustrated in the figure 5.

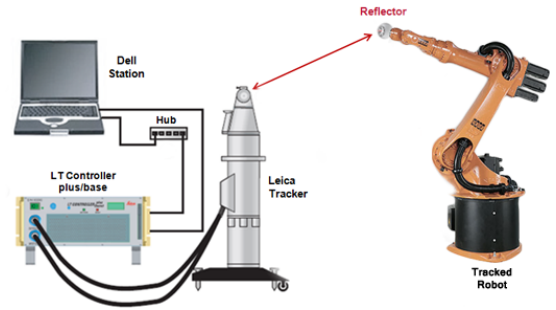


Fig. 5. Robot positions measurement using Leica Laser Tracker

In order to gather the needed data, an acquisition protocol is first established, based on which training datasets (current time-series and end-effector tracked positions) are acquired. The robot is first programmed to move repetitively between two points A and B for a defined period of time. Simultaneously, the laser tracker records the different positions of the end-effector during its movement with a rate of 100 Hz.

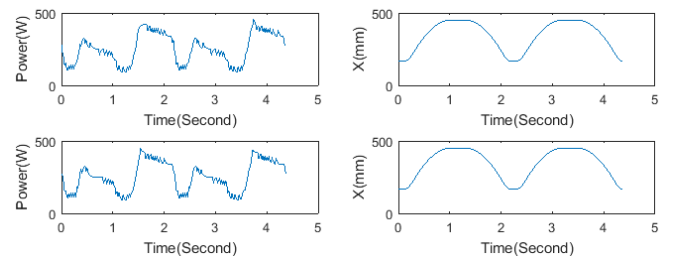


Fig. 6. Two consecutive profiles extracted from the Power and X coordinates data sets

The obtained raw data is then divided into 6 data sets : Power L1, Power L2, Power L3, X axis coordinates, Y axis

coordinates and Z axis coordinates. The reference trajectory is first defined. To build the reference spatial positions data set, the movement of the robot is tracked for a duration of 7 minutes with a capture rate of 1000 records per second, leading to data set of more than 420000 records of X, Y and Z axis coordinates values. The reference curve is obtained by applying a linear regression curve fitting on the XZ plane. Based on this model, the accuracy of new tracked positions of the robot end-effector is obtained by calculating the MSE values of these points with regard to the reference model.

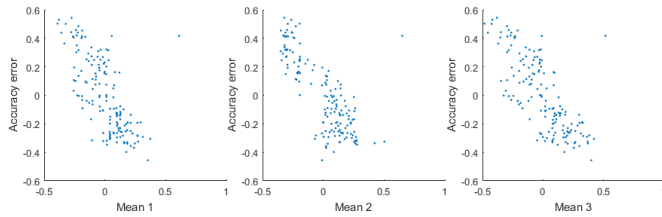


Fig. 7. Scatterplot of Mean of Current L1, L2, L3 and Accuracy Error

Once the data cleaned, the recorded power time-series and coordinates data are segmented into multiple profiles with regards to the robot movement. In fact, for each two consecutive movements of the robot's end effector from point A to point B and from point B to point A, we segment one profile from the power data set as well as from the positions data set. The figure 6 exposes examples of the plotting of two segmented profiles from power and X axis coordinates.

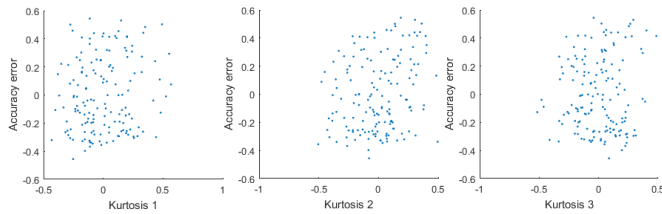


Fig. 8. Scatterplot of Kurtosis and Accuracy Error

For the training stage, we prepared a data set of current and robot positions data recorded during 11 minutes and 21 seconds of the robot movement. After we carried out the feature extraction, we obtained a feature vector of 21 columns (7 features from each of the 3-phase currents) and 155 rows (observations), corresponding to the different profiles.

Acct	Energy	Mean	STD	Kurtosis	Skweness	Long P	Short P
1.0000	-0.7314	-0.6919	-0.6864	0.1429	0.1829	-0.2659	-0.3125
-0.7314	1.0000	0.9897	0.9183	-0.1063	-0.4012	0.2939	0.3652
-0.6919	0.9897	1.0000	0.8875	-0.0967	-0.4341	0.3075	0.3802
-0.6864	0.9183	0.8875	1.0000	-0.0470	-0.3008	0.2118	0.2568
0.1429	-0.1063	-0.0967	-0.0470	1.0000	0.0984	-0.2203	-0.1776
0.1829	-0.4012	-0.4341	-0.3008	0.0984	1.0000	-0.1329	-0.1502
-0.2659	0.2939	0.3075	0.2118	-0.2203	-0.1329	1.0000	0.7947
-0.3125	0.3652	0.3802	0.2568	-0.1776	-0.1502	0.7947	1.0000

Fig. 9. Correlation values of features extracted from Power L1

In order to see how the features relate to accuracy error, we proceed with the correlation analysis. We examine the

direction and strength of the correlation of each feature with the accuracy using the visual representation (finding the line of best fit in the scatterplots) in a first step, and calculating then the Pearson coefficients of the correlation matrix. We give two examples of the obtained scatterplots of Mean and Accuracy errors in figure 7 and Kurtosis and Accuracy errors in figure 8, related to the 3-phase currents. Based on these plots, we can demonstrate that a relatively strong negative correlation exists between the Mean and Accuracy, while no particular pattern of the data points does exist in the Kurtosis and Accuracy scatterplot, showing that there is a little or no correlation.

Linear regression model:
 $y \sim 1 + x1 + x2 + x3 + x4 + x5 + x6 + x7 + x8 + x9 + x10 + x11$

Estimated Coefficients:

	Estimate	SE	tStat	pValue
(Intercept)	-5.0614e-14	0.0073598	-6.8771e-12	1
x1	1.9161	0.44004	4.3544	2.5278e-05
x2	-4.8473	2.5882	-1.8728	0.063134
x3	4.8122	2.1852	2.2022	0.02925
x4	0.084753	0.093308	0.90832	0.36524
x5	-0.92275	0.50588	-1.8241	0.070232
x6	2.0387	0.59054	3.4523	0.00073128
x7	-2.1557	0.40833	-5.2794	4.7171e-07
x8	4.2066	2.5798	1.6306	0.10517
x9	-5.6189	2.3184	-2.4236	0.016616
x10	-0.05175	0.11375	-0.45495	0.64983
x11	-0.04415	0.067539	-0.65369	0.51436

Number of observations: 155, Error degrees of freedom: 143
 Root Mean Squared Error: 0.0916
 R-squared: 0.885, Adjusted R-Squared 0.876
 F-statistic vs. constant model: 99.7, p-value = 2.45e-61

Fig. 10. Specification of the obtained multiple linear regression model

The Pearson correlation coefficients of features extracted from Power L1 with Accuracy error are given in the correlation matrix given in figure 9. We only give the first rows of the matrix since no other relevant information can be found in the remaining rows. We proceed in the same way with features extracted from Power L2 and Power L3.

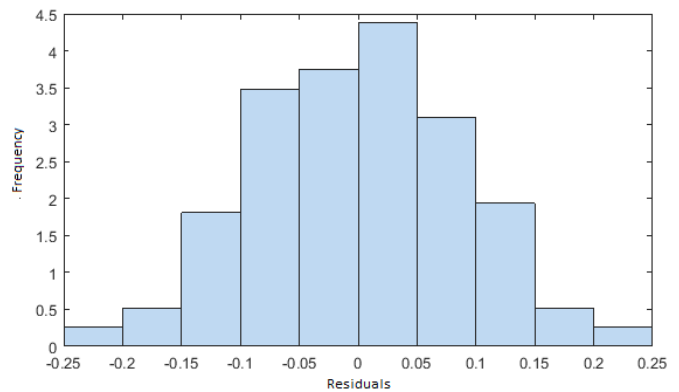


Fig. 11. Histogram of Residuals

Based on the analysis of the visual representations and calculated correlation coefficients, we can assume that the Mean, Standard Deviation and Energy are features with a relatively strong anti-correlation with Accuracy. The skewness extracted from Power L2 and Power L3 have besides shown a relatively

strong correlation with accuracy, while skewness extracted from Power L1 has a very weak correlation. We decide therefore to define our final feature vector based on these features (Mean, Standard deviation and Energy from the 3-phase currents, with Skewness from power L2 and power L3).

We finally fit our data in the multiple linear regression model. The obtained coefficients as well as the specification of the obtained model are given in figure 10. As a diagnostic plot to identify outliers, we expose in figure 11 a residuals histogram plot showing the range of the residuals of the obtained model and their frequencies. This histogram shows that the residuals have almost a normal distribution with no outliers, which highlights the goodness of fit of the obtained model.

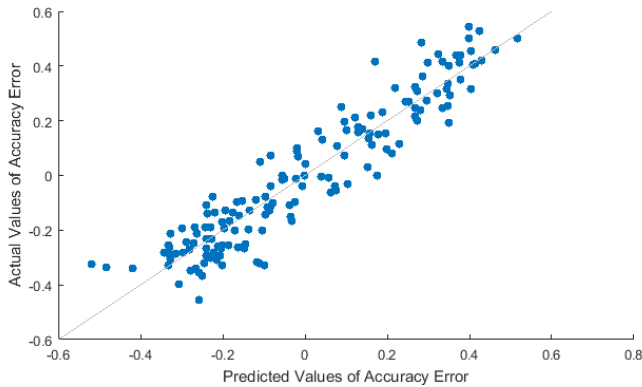


Fig. 12. Plot of Predicted vs Observed Values of Accuracy Error

To better understand the accuracy of our model, we plot the predicted values of accuracy error versus its observed values, as illustrated in figure 12. This plot shows a strong correlation between the model's prediction and its actual results, which reflects the satisfactory accuracy of the model, despite the small size of the used training data. The obtained RMSE is 0.0916 and thus an MSE of 0.008, which is very close to 0. We can then conclude that the obtained model is very useful for predicting accuracy values based on the specified set of features.

V. CONCLUSION

In this study, a data-driven accuracy error detection approach has been proposed for predictive maintenance of robot manipulators. The specified predictive modeling methodology relies on data analytics of the robot power time-series to predict the accuracy error. To evaluate the proposed approach, an experimental procedure has been carried out. The experimental results have demonstrated the existence of a relatively strong correlation between a set of electrical features of the robot 3-phase current and its accuracy values. The obtained predictive model has also shown a satisfactory goodness of fit and high usefulness for predicting accuracy errors based on measured power data. The specified features used within this paper can be employed for a further enhanced predictive modeling based on huge datasets of power time-series, to allow an efficient predictive maintenance of industrial robots.

REFERENCES

- [1] Ying Peng, Ming Dong, and Ming Jian Zuo. Current status of machine prognostics in condition-based maintenance: a review. *The International Journal of Advanced Manufacturing Technology*, 50(1):297–313, 2010.
- [2] Jay Lee, Fangji Wu, Wenyu Zhao, Masoud Ghaffari, Linxia Liao, and David Siegel. Prognostics and health management design for rotary machinery systems: reviews, methodology and applications. *Mechanical Systems and Signal Processing*, 42(12):314 – 334, 2014.
- [3] Xuewu Dai and Zhiwei Gao. From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *IEEE Transactions on Industrial Informatics*, 2013.
- [4] Recent advances and trends in predictive manufacturing systems in big data environment. *Manufacturing Letters*, 1(1):38–41, oct 2013.
- [5] R.J. Patton Jie Chen. Robust model-based fault diagnosis for dynamic systems. *The International Series on Asian Studies in Computer and Information Science*, 1999.
- [6] André Carvalho Bittencourt, Kari Saarinen, and Shiva Sander-Tavallaey. A data-driven method for monitoring systems that operate repetitively - applications to wear monitoring in an industrial robot joint. In *Proc. of the 8th IFAC SAFEPROCESS*, volume 8, Mexico City, Mexico, 2012.
- [7] Cesare Fantuzzi, Cristian Secchi, and Antonio Visioli. On the fault detection and isolation of industrial robot manipulators, 2003.
- [8] Sertac-Savas Sahin Yildirim Ikbal Eski, Selcuk Erkaya. Fault detection on robot manipulators using artificial neural networks. *Robotics and Computer-Integrated Manufacturing*, 2010.
- [9] V Sathish, Srini Ramaswamy, and Sachit Butail. Training data selection criteria for detecting failures in industrial robots. *IFAC-PapersOnLine*, 49(1):385 – 390, 2016.
- [10] M. Abderrahim, A. Khamis, L. Moreno, and S. Garrido. *Accuracy and Calibration Issues of Industrial Manipulators*. INTECH Open Access Publisher, 2006.
- [11] R. Loser M. Kleinkes. *Laser Tracker and 6DoF measurement strategies in industrial robot applications*. 2006.
- [12] Salvador Garca, Julin Luengo, and Francisco Herrera. *Data Preprocessing in Data Mining*. Springer Publishing Company, Incorporated, 2014.
- [13] Teemu Räsänen and Mikko Kolehmainen. *Feature-Based Clustering for Electricity Use Time Series Data*, pages 401–412. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [14] Kuka Roboter GmbH. *KR 6 Robot Specification*. 2003.
- [15] Leica Geosystems. *Leica Laser Tracker for Hand-Tools*. 2003.