IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 14, 2021, accepted April 25, 2021, date of publication April 28, 2021, date of current version May 6, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3076149

Fault Diagnosis for Variable Frequency Drive-Fed Induction Motors Using Wavelet Packet Decomposition and Greedy-Gradient Max-Cut Learning

SHAFI MD KAWSAR ZAMAN[1], (Member, IEEE), XIAODONG LIANG[2], (Senior Member, IEEE), AND WEIXING LI[3], (Senior Member, IEEE)

[1]Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada
[2]Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada
[3]School of Electrical Engineering, Dalian University of Technology, Dalian 116024, China

Corresponding author: Xiaodong Liang (xil659@mail.usask.ca)

This work was supported in part by the Natural Science and Engineering Research Council of Canada (NSERC) under Discovery Grant RGPIN-2016-04170.

ABSTRACT In this paper, a novel fault diagnosis method for variable frequency drive (VFD)-fed induction motors is proposed using Wavelet Packet Decomposition (WPD) and greedy-gradient max-cut (GGMC) learning algorithm. The proposed method is developed using experimental stator current data in the lab for two 0.25 HP induction motors fed by a VFD, subjected to healthy and faulty cases under various operating frequencies and motor loadings. The features are extracted from stator current signals using WPD by evaluating energy eigenvalues and feature coefficients at decomposition levels. The proposed method is validated by comparing with other graph-based semi-supervised learning (GSSL) algorithms, local and global consistency (LGC) and Gaussian field and harmonic function (GFHF). To enable fault diagnosis for untested motor operating conditions, mathematical equations to calculate features for untested cases are developed through surface fitting using features extracted from tested cases.

INDEX TERMS Graph-based semi-supervised learning, greedy-gradient max-cut, induction motors, variable frequency drive, wavelet packet decomposition.

NOMENCLATURE

VFD	Variable frequency drive
GSSL	Graph-based semi-supervised learning
GGMC	Greedy-gradient max-cut
LGC	Local and global consistency
GFHF	Gaussian field and harmonic function
SSL	Semi-supervised learning
l, u	Number of labeled and unlabeled inputs
G	Undirected graph
X	Set of vertices
E	Set of edges
W	Weight matrix
D	Vertex degree matrix
L	Normalized graph Laplacian
F	Continuous classification function
B	binary matrix
k-NN	k-nearest-neighbors
WPD	Wavelet packet decomposition
db	Daubechies wavelet
f, f_s	Fundamental frequency, sampling frequency
N	Length of a signal
N_{L_s}	Number of decomposition levels
D_j	Detail signal at level j where $j = 1, 2, 3, \ldots, 10$
E_j	Energy eigenvalue at a particular decomposition level j ($j = 1, 2, 3, \ldots, 10$)
$x_{j,k}$	WPD coefficients at level j and node k
$W(j, k)$	WPD Feature coefficients
LR	Label ratio

The associate editor coordinating the review of this manuscript and approving it for publication was Pinjia Zhang.

H	Healthy motor
UNB	Unbalance shaft rotation
BF	Bearing fault
BRB	Broken rotor bar
1BRB	One broken rotor bar
2BRB	Two broken rotor bars
3BRB	Three broken rotor bar
UV	Unbalanced voltage
d8	Decomposition level 8

I. INTRODUCTION

Induction motors are most widely used in various industrial sectors, either connected direct online or fed by variable frequency drives (VFDs). A survey for 0.75 kW to 150 kW induction motors in [1] reports common faults regularly occurred: 7% broken rotor bar faults, 21% stator winding faults, 69% bearing faults, and 3% shaft/coupling and other faults [1]. To improve reliability of critical industrial processes and reduce operational downtime, effective fault diagnosis for various electrical and mechanical faults in induction motors is essential.

VFDs are increasingly used in industry facilities by offering flexible production control and soft motor start-up. Compared to direct online induction motors, VFD-fed induction motors are affected by variable operating frequencies, harmonics generated at the drive output, and complex control systems. Harmonics may cause higher stress in bearings and windings of the motor; harmonics also cause poor signal to noise ratio when using stator currents for fault diagnosis; a fault diagnosis algorithm developed for a fixed motor operating frequency may become invalid due to varying drive output frequencies in VFD applications; control loops of VFDs may affect how electrical or mechanical variables of induction motors are coupled under faulty conditions [2]. Therefore, the introduction of VFDs requests significant changes in the field of induction motors fault diagnosis [2], but existing work mainly focus on direct online induction motors.

Some research progress has been made on induction motor fault diagnosis for VFD applications, mostly using signal-based techniques. Among the reported research in the literature, Refs. [3], [4] detect eccentricity faults in inverter-fed induction motors by creating mathematical models of the machine, but precise machine models under faulty conditions may be difficult to develop. Refs. [2], [5] rely on signal processing to achieve signature extraction for fault diagnosis.

With the advancement of artificial intelligence, machine learning have attracted increasing interests in fault diagnosis of induction motors. Machine learning-based fault diagnosis is data-driven, does not require machine models, and can achieve high accuracy through supervised learning, deep learning or semi-supervised learning. Most inverter-fed induction motors fault diagnosis methods reported in the literature use supervised learning [6].

Labeling a large amount of data samples using supervised learning and deep learning is time-consuming and expensive [7], [8], semi-supervised learning (SSL), on the other hand, only needs a small amount of labeled data to train a classification model [7], [8]. The graph-based semi-supervised learning (GSSL) is a transductive semi-supervised technique, and among the most popular and most effective semi-supervised learning strategies [8], [9]. GSSL is effective by propagating a small amount of initial labels to a large amount of unlabeled data [10]. GSSL exploits connectivity patterns between labeled and unlabeled samples to improve classification performance through the nearest neighbor graph, the information from labeled to unlabeled samples is propagated along the edges of the graph [8], [9].

Several SSL-based induction motors fault diagnosis methods are reported: semi-supervised smooth alpha layering [11], semi-supervised label consistent dictionary learning framework [12], information fusion strategy-based semi-supervised deep learning [13], deep semi-supervised method of multiple association layers networks [14], and manifold regularization-based SSL [15]. However, these papers are for direct online induction motors; and only vibration or the combination of vibration and torque signals are used as monitoring signals.

Although GSSL is considered among the most popular and effective SSL, to the authors' best knowledge, its application has not been used in fault diagnosis for VFD-fed induction motors. The authors conducted some preliminary work using GSSL in [16]–[18], but they are all for direct online induction motors. For VFD-fed induction motor's operation, the changing frequency at the VFD output might affect or even invalid an induction motor fault diagnosis approach that is functional at a fixed operating frequency. To overcome such concerns and fill in this research gap for VFD-fed induction motors, we propose a fault diagnosis method using Wavelet Packet Decomposition (WPD) and greedy-gradient max-cut (GGMC) learning (a GSSL algorithm) in this paper. The proposed method maintains a high classification accuracy using a small amount of labeled data; is effective when subjected to a changing motor operating frequency from the output of the VFD; can classify an individual fault or several simultaneous faults; and can detect faults for untested motor operating conditions with the untested frequency and motor loading.

The main contribution of this paper includes: 1) propose an effective fault diagnosis method for VFD-fed induction motors using a GSSL learning algorithm, GGMC; 2) develop a novel feature extraction technique using WPD by evaluating energy eigenvalues and feature coefficients at decomposition levels using the stator current signal; 3) to enable fault diagnosis for untested frequency and motor loading conditions using the proposed method, develop mathematical equations for feature calculation for such untested motor operating conditions through surface fitting using features extracted from experimental data of tested conditions.

The rest of this paper is arranged as follows: the proposed fault diagnosis method and the fundamental theory of GSSL are provided in Section II; experimental testing of the two

induction motors fed by a VFD is introduced in Section III; the novel feature extraction method using WPD is discussed in Section IV; fault classification results using the proposed method are shown in Section V; mathematical equations to calculate features for untested motor operating conditions are developed through surface fitting in Section VI; conclusions are drawn in Section VII.

II. THE PROPOSED METHOD

In this paper, a fault diagnosis method for VFD-fed induction motors is proposed, its implementation procedure is summarized in the following five steps. The flow chart of the proposed method is shown in Fig. 1.

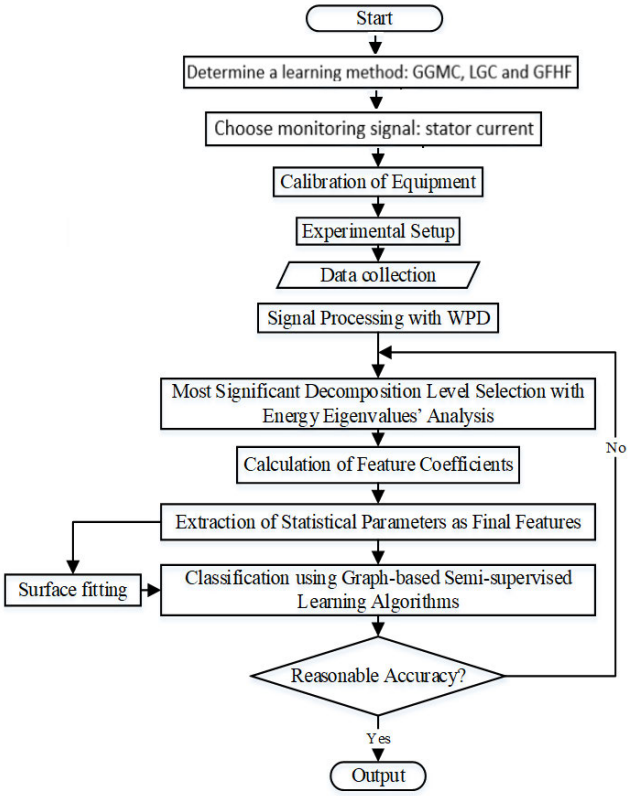


FIGURE 1. The flowchart of the proposed method.

Step 1: *Determine a learning method*. In this paper, GGMC learning, a GSSL algorithm, is chosen as the learning method. GGMC only requires a small amount of labeled data in a dataset. Its performance is evaluated by comparing with other two GSSL algorithms, local and global consistency (LGC) and Gaussian field and harmonic function (GFHF).

Step 2: *Choose monitoring signal and obtain dataset*. Although many published papers use vibration signals to monitor and detect faults for induction motors, stator currents measured in experiments in the lab are chosen as the monitoring signal in this paper due to its advantage of ease to measure. Experimental stator currents are measured for each healthy or faulty case of a motor under a specific operating frequency and a motor loading during each test.

Step 3: *Feature extraction using WPD*. In order to extract features, WPD is used in this paper by processing experimental stator current signals. The details is presented in Section IV.

Step 4: *Fault classification*. Using the proposed approach, an individual single- or multi-fault or several simultaneous faults are classified for VFD-fed induction motors.

Step 5: *Fault diagnosis for a untested frequency and motor loading condition*. To enable machine learning for untested motor operating conditions, feature calculation formulas for untested cases are developed through surface-fitting using features extracted from experimental data for the tested conditions

The basic theory of GSSL is introduced in this section. In GSSL, both labeled and unlabeled samples are treated as vertices in a graph, and pairwise edges between the vertices are built and weighed by similarities between the sample pairs. A small portion of vertices carries seed labels, these vertices are harnessed through the graph partition or the information propagation to predict labels for unlabeled vertices [10].

In a GSSL algorithm, for a dataset X, labeled samples are $\{(x_1, y_1), \ldots, (x_l, y_l)\}$, and unlabeled samples are $\{x_{l+1}, x_{l+u}\}$. Inputs of labeled samples are $X_l = \{x_1,\ldots,x_l\}$, and inputs for unlabeled samples are $X_u = \{x_{l+1},\ldots,x_{l+u}\}$, where l and u represent the number of labeled and unlabeled inputs, respectively. The labeled input X_l is associated with the initial seed labels $Y_l = \{y_1,\ldots,y_l\}$, $y_i \in \{1,2, \cdots, c\}$, c *is the number of class labels, and* $i = 1, 2,\ldots,l$. The goal of a GSSL algorithm is to determine the missing labels for unlabeled samples, $Y_u = \{y_{l+1}, \ldots, y_{l+u}\}$ for the unlabeled input X_u, where $l \ll n$ ($l + u = n$). GSSL constructs a weighted sparse graph G from the input data $X = X_l \cup X_u$. Following that, a labeling algorithm utilizes G and the labels Y_l to estimate $\hat{Y}_u = \{\hat{Y}_{l+1}, \ldots, \hat{Y}_{l+u}\}$, and $\hat{Y}_u$ are expected to match the true labels $Y_u = \{y_{l+1},\ldots,y_{l+u}\}$ as measured by a loss function [10].

Assuming $G = \{X, E, W\}$ is an undirected graph produced from the input data X. In the graph G, the set of vertices is $X = \{x_i\}$, and the set of edges is $E = \{e_{ij}\}$. Each input data sample x_i is a vertex, and the weight of edge e_{ij} is w_{ij}. Typically, a kernel function $k(\cdot)$ is used over pairs of points to compute weights. Weights for edges are used to generate a weight matrix, $W = \{w_{ij}\}$, and subsequently the vertex degree matrix $D = diag([d_1,\ldots,d_n])$ is defined as $d_i = \sum_{j=1}^{n} w_{ij}$. The graph Laplacian is defined as $\Delta = D - W$, and the normalized graph Laplacian is [10]

$$L = D^{-1/2}\Delta D^{-1/2} = I - D^{-1/2}WD^{-1/2} \tag{1}$$

They are operators in the function space f, which is used to define a smoothness measure in a graph over highly connected regions [10]. The smoothness measurement of the function f using L over a graph is defined

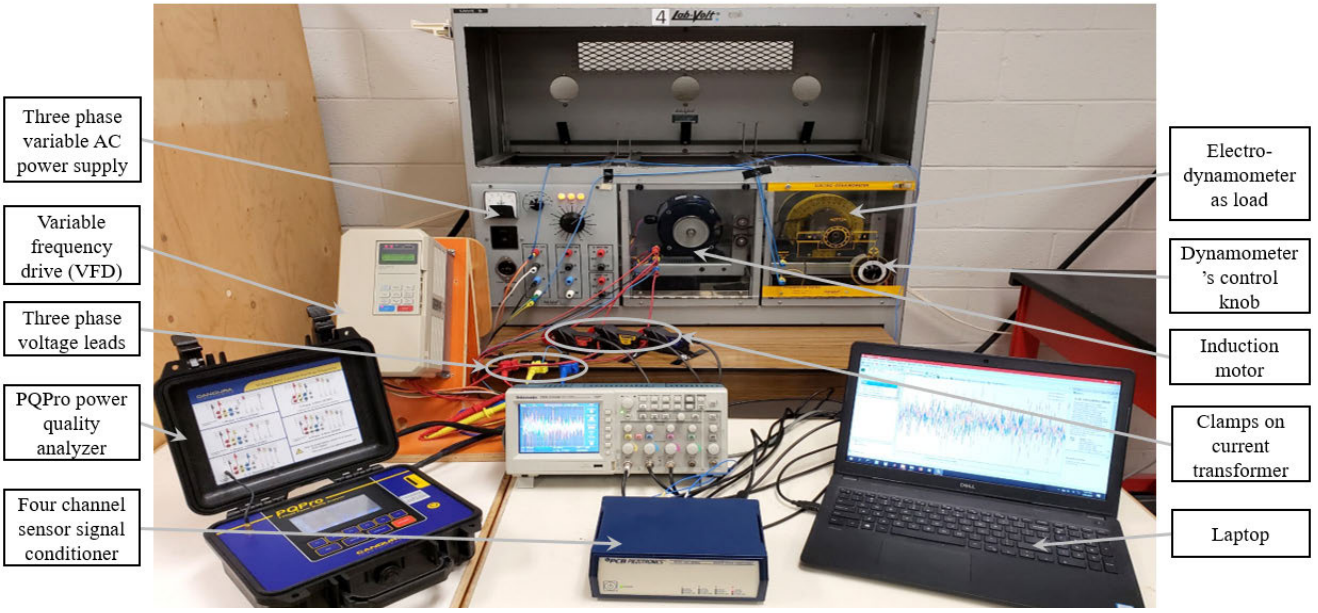


FIGURE 2. Experimental set-up in the lab.

by [10]

$$\langle f, Lf \rangle = \sum_i \sum_j w_{ij} \left\| \frac{f(x_i)}{\sqrt{d_i}} - \frac{f(x_j)}{\sqrt{d_j}} \right\|^2 \quad (2)$$

Finally, the label information is encoded into a label matrix $Y \in \{1, 0\}^{n \times c}$, where $y_{ij} = 1$ if x_i has a label $j \in \{1, 2, \ldots, c\}$, $y_i = j$; $y_{ij} = 0$ if not. The label prediction function f is applied to vertices X, the output values of f is saved in F = f (X). GSSL uses W along with the initial known label matrix Y_l to recover a continuous classification function F, by minimizing a predefined fitness function on the graph G [10].

The k-nearest-neighbors (k-NN) method [19] is adopted in this paper. The constructed graph is sparsified to improve efficiency, accuracy, and adaptive capability to noise. Graph sparsification will remove edges by finding a binary matrix $B \in B\ \{1,0\}^{n \times n}$ where $B_{ij} = 1$ means that an edge exists between nodes x_i and x_j, while $B_{ij} = 0$ means that the edge does not exist, assuming $B_{ii} = 0$. After a graph is sparsified and a binary matrix B is found, several schemes can be further applied to adjust the weights and generate the final weight matrix W. There are three potential weighting schemes, binary weighting, Gaussian kernel weighting, and locally linear reconstruction-based weighting. In this paper, the first two weighting schemes are considered [10], [20].

Three GSSL algorithms (GGMC. LGC, and GFHF) are considered in this paper. LGC [10], [21] and GFHF [22] are univariate graph regularization-based algorithms, while GGMC is a bivariate label propagation algorithm [23], [24].

III. EXPERIMENTS IN THE LAB

In this paper, learning datasets are stator current signals measured in the lab for various healthy and faulty cases of an induction motor fed by a VFD. Experiments were conducted using the set-up shown in Fig. 2, including a VFD, an induction motor and the load.

The VFD is manufactured by Saftronics (Model: CIMR-G5U23P7F), a low voltage drive with a voltage source inverter and two-level pulse width modulation technique. The drive inputs are AC three-phase, rated at 200-220 V at 50 Hz (200-230 V at 60 Hz) and 21 A; the drive output are AC three-phase, rated at 0-230 V, 0-400 Hz and 17.5 A. In these experiments, the voltage per Hz control was selected, and the carrier frequency was set at 3,100 Hz for the drive.

Two three-phase squirrel-cage induction motors were used in the experiments, each rated at 4-pole, 0.25 HP, 208-230/460 V, and 1725 rpm (Model: LEESON 101649), named "Motor 1" and "Motor 2". The load was a dynamometer coupled with the motor shaft through a pulley, and the dynamometer's control knob was used to adjust the motor loading. An eight-channel power quality analyzer, PQPro by CANDURA Instruments, was used to measure the motor's three-phase stator currents. For each test, stator currents were recorded for 2 min with a sampling frequency of 15.38 kHz.

A healthy and five faulty cases for Motor 1 were tested in the lab (the considered faults are mainly mechanical faults): 1) a healthy motor (H); 2) an unbalance shaft rotation (UNB); 3) a bearing fault (BF); 4) a multi-fault with BF and UNB (BF+UNB); 5) a multi-fault with BF and one broken rotor bar (BRB) (BF+1BRB); and 6) a multi-fault with BF, UNB, and unbalanced voltage (UV) from the VFD output (BF+UNB+UV)).

A healthy and five faulty cases for Motor 2 were tested (the considered faults are mainly electrical faults): 1) a healthy

motor (H); 2) unbalanced voltage (UV) from the VFD output; 3) one BRB fault (1BRB); 4) two BRB fault (2BRB); 5) three BRB fault (3BRB), and 6) a multi-fault with three BRBs and UV (3BRB+UV).

In the lab, various faults were created as shown in Fig. 3. a BRB fault was created by drilling a hole (a 5 mm diameter and 18 mm depth) in a rotor bar. One hole was drilled for one BRB fault (Fig. 3(a)); two and three holes were drilled on adjacent rotor bars for two and three BRB faults (Figs. 3(b) and 3(c)), respectively, which is quite different from previous experiments done in the lab for two and three BRB faults for direct online induction motors with drilled holes 90° apart [16]–[18].

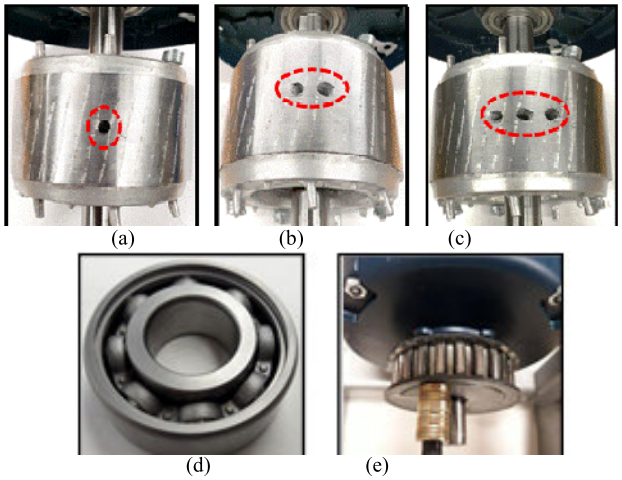


FIGURE 3. Faults created on the two motors in the lab: (a) 1BRB; (b) 2BRB; (c) 3BRB; (d) BF; and (e) UNB.

A general roughness type bearing fault (BF) was realized by the sandblasting process; the outer and inner raceway of the bearing became very rough (Fig. 3(d)). The UNB is due to uneven mechanical load distribution causing unbalanced shaft rotation, and it was created by adding extra weight on part of the pulley (Fig. 3(e)). The UV condition was created by adding an extra resistance at the second phase of VFD output.

For each of the above healthy or faulty cases of the two motors, the motor was tested under six output frequencies from the drive output (45, 50, 55, 60, 65, and 70 Hz). Under a specific operating frequency, six motor loadings (0, 20, 40, 60, 80, and 100%) were tested for each motor, which leads to a total of 432 tests. Table 1 provides the parameter settings of the VFD and the motor used in experiments.

TABLE 1. VFD and motor parameter settings used in experiments.

Parameters	Settings
VFD output frequency, Hz	45, 50, 55, 60, 65, 70
VFD carrier frequency, Hz	3,100
VFD control method	Voltage per Hz control
Motor loading, %	0, 20, 40, 60, 80, 100

IV. THE NOVEL FEATURE EXTRACTION USING WPD

In this paper, a novel feature extraction method using WPD is developed through MATLAB in two steps: 1) determine energy eigenvalues for all decomposition levels in order to choose the most significant decomposition level with most variations among healthy and faulty cases of the motor; and 2) use coefficients at the most significant decomposition level to extract features for machine learning.

WPD decomposes a signal into the details (high pass) and approximations (low pass) by the consecutive application of quadrature mirror filters with finite impulse response, which have pre-determined scaling and mother wavelet functions [25], [26]. For the decomposition level j, WPD produces 2^j sets of coefficients [25].

The structure of the WPD algorithm up to three resolution levels is shown in Fig. 4, where S represents the original signal, A and D represent approximation and detail signals, respectively, ($\downarrow$ 2) denotes downsampling by 2, and the subscript numbers represent the number of decomposition levels. All coefficients corresponding to all decomposition levels can be computed sequentially. The decomposition coefficient of the j^{th} level can be obtained by the $(j-1)^{\text{th}}$ level. Decomposition up to the j^{th} level results in the frequency ranges of all subspaces at the j^{th} level as $\left\{\left[0, \frac{f_s}{2^{(j+1)}}\right]; \left[\frac{f_s}{2^{(j+1)}}, \frac{2f_s}{2^{(j+1)}}\right]; \ldots; \left[\frac{(2^j-1)f_s}{2^{(j+1)}}, \frac{f_s}{2}\right]\right\}$; where f_s is defined as the sampling frequency [27].

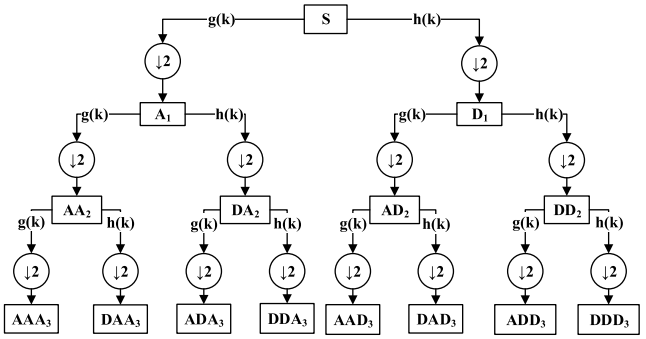


FIGURE 4. Wavelet packet tree generated from applying WPD on the signal S, up to the third resolution.

As an example, for the decomposition level 8 (d8), the corresponding frequency ranges of all subspaces at the 8^{th} level can be obtained by substituting $j = 8$.

In this paper, although three-phase stator currents were measured during experiments and available to be used, we found that one phase current was sufficient for induction motor fault diagnosis. Therefore, among the measured three-phase stator currents, only the 2^{nd} phase currents are processed by WPD and used for fault diagnosis in each case. The length of the stator current dataset contains 90,000 samples for each test. The dataset is further segmented into 10 data windows, each containing 9,000 data samples.

In this study, Daubechies wavelet of the order 44 (*db*44) is chosen as the mother wavelet for WPD.

TABLE 2. Energy eigenvalues at individual decomposition levels from 1 to 10 for motor 1 (60 Hz and 80% loading).

Healthy or faulty case	Energy eigenvalues at individual decomposition level from 1 to 10									
	1	2	3	4	5	6	7	8	9	10
BF	18.33	54.03	1.339	2.836	72.57	2.767	2003	4168	25.89	26.32
BF+1BRB	19.99	55.66	1.42	2.745	78.49	3.287	2286	3834	24.92	28.04
BF+UNB	16	49.88	1.18	2.972	83.54	3.182	2521	3663	27.82	33.64
BF+UNB+UV	19.42	50.37	1.707	3.039	53.93	2.309	72.04	1412	3.016	2.002
H	22.98	71.32	2.109	0.547	0.292	1.368	1402	3186	7.93	5.412
UNB	14.16	47.21	1.061	3.011	101	2.974	2033	4080	28.62	32.2

TABLE 3. Energy eigenvalues at individual decomposition levels from 1 to 10 for motor 2 (55 Hz and 100% loading).

Healthy or faulty case	Energy eigenvalues at individual decomposition level from 1 to 10									
	1	2	3	4	5	6	7	8	9	10
1BRB	14.49	42.53	0.993	2.324	65.03	1.616	1118	2819	15.83	14.1
2BRB	14.88	44.38	0.9607	2.134	53.02	1.713	529.6	3424	6.183	5.608
3BRB	21.54	47.85	1.913	2.893	100.1	1.921	1346	2582	18.2	23.78
3BRB+UV	16.39	42.31	1.305	2.228	51.68	1.658	331	744.1	6.883	5.634
H	30.4	66.63	3.001	0.888	0.305	0.8804	1044	1698	7.178	15.54
UV	25.59	64.74	2.131	0.618	0.494	0.356	265.6	425.6	3.752	3.364

A. DETERMINE MOST SIGNIFICANT DECOMPOSOTION LEVEL USING ENERGY EIGENVALUES

The required number of decomposition levels for WPD is determined by the data-independent selection method [26]:

$$N_{L_s} = int\left(\frac{log\left(f_s/f\right)}{log\,(2)}\right) \tag{3}$$

where f_s is the sampling frequency of the stator current signal ($f_s \approx 15.38kHz$) and f is the fundamental frequency ($f = 60Hz$). According to [28], two additional levels are needed for better decomposition of the signal. So the required number of decomposition levels in this study can be calculated by

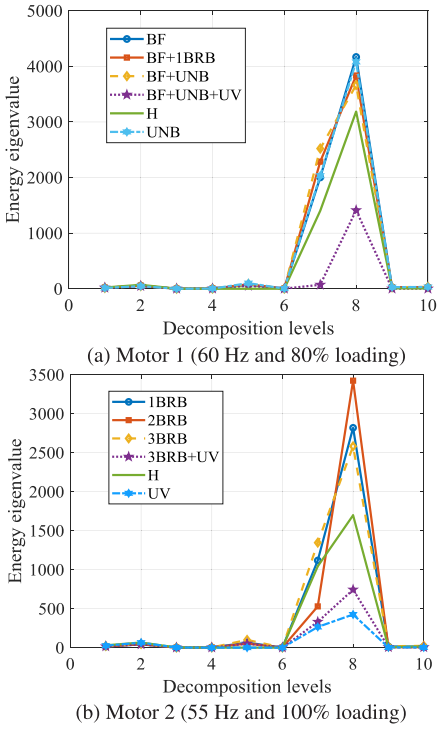
$$N_{L_s} + 2 = int\left(\frac{log\left(15380/60\right)}{log\,(2)}\right) + 2 = 10 \text{ levels} \tag{4}$$

The fault information in the stator current is distributed in frequency bands determined by WPD. Therefore, the energy eigenvalue at each frequency band or each node of the WPD tree can serve as a fault indicator. The energy eigenvalue at each decomposition level for a specific motor operating condition can be calculated as follows [26], [29]:

$$E_j = \sum_{n=1}^{n=N} \left|D_j(n)\right|^2 \tag{5}$$

where D_j is the detail signal at level $j (j = 1, 2, 3, \ldots, 10)$, and N is the length of the signal.

Fig. 5 shows energy eigenvalues vs. decomposition levels for healthy and faulty cases for Motor 1 at (60 Hz and 80% loading) and Motor 2 at (55 Hz and 100% loading). Tables 2 and 3 show the corresponding energy eigenvalues for the decomposition levels from 1 to 10 for Motor 1 at (60 Hz and 80% loading) and Motor 2 at (55 Hz and 100% loading). It is found that the most significant differences of energy eigenvalues among one healthy and five faulty cases of each motor occur at the decomposition level 8 ($d8$) ($j = 8$).



(a) Motor 1 (60 Hz and 80% loading)

(b) Motor 2 (55 Hz and 100% loading)

FIGURE 5. Energy eigenvalues vs. decomposition levels for healthy and faulty cases of the two motors under a specific motor operating condition.

Our study confirms that this remains true for other tested operating conditions. Therefore, $d8$ is selected as the most significant decomposition level.

B. EXTRACT STATISTIC FEATURES USING FEATURE COEFFICIENTS AT THE MOST SIGNIFICANT DECOMPOSITON LEVEL

A time-frequency spectrum of the stator current generated by WPD contains wavelet packet coefficients for all levels and nodes. Such spectrum is arranged in a way that low- and high-frequency components are located at small and large nodes,

TABLE 4. Features extracted by WPD for motor 1 (BF+1BRB) (under 60 Hz and 80% loading).

Features	Mean	Median	Median Absolute Deviation	Mean Absolute Deviation	L1 Norm	L2 Norm	Max Norm	Standard Deviation
s_1	0.004743	0.001888	0.633089	0.585094	5265.831	61.84523	1.159359	0.651925
s_2	0.002606	-0.00216	0.628503	0.583123	5248.065	61.66634	1.155044	0.650051
s_3	-0.00043	-0.00162	0.627964	0.583564	5252.078	61.71843	1.159359	0.650605
s_4	-0.00321	-0.01511	0.631201	0.585914	5273.399	61.91367	1.141556	0.652656
s_5	-0.00204	-0.00998	0.626615	0.58384	5254.685	61.70863	1.17986	0.650499
s_6	0.001059	-0.00405	0.625806	0.582364	5241.247	61.64405	1.156662	0.649821
s_7	0.002791	0.003237	0.630661	0.584339	5259.066	61.80127	1.163136	0.651473
s_8	-0.00102	-0.01403	0.625266	0.582462	5242.23	61.64345	1.157741	0.649814
s_9	0.003234	0.005395	0.628773	0.584343	5259.157	61.80536	1.16907	0.651514
s_{10}	0.002927	0.005934	0.627964	0.583113	5248.097	61.67684	1.167452	0.65016

TABLE 5. Features Extracted by DWT for motor 2 (3BRB+UV) (under 55 Hz and 100% loading).

Features	Mean	Median	Median Absolute Deviation	Mean Absolute Deviation	L1 Norm	L2 Norm	Max Norm	Standard Deviation
s_1	0.002563	0.001079	0.412708	0.34853	3136.754	38.69379	0.867497	0.407883
s_2	0.002814	0.00027	0.417294	0.349325	3143.87	38.77807	0.858865	0.40877
s_3	0.002291	0.000539	0.412169	0.348321	3134.866	38.66798	0.866418	0.407613
s_4	-7.67E-05	-0.00324	0.412708	0.348758	3138.826	38.71616	0.872891	0.408127
s_5	-0.00069	-0.00216	0.414866	0.349688	3147.213	38.79077	0.871273	0.408913
s_6	-0.00102	-0.00216	0.41109	0.348268	3134.439	38.66822	0.882063	0.40762
s_7	0.00168	-0.00108	0.412169	0.347733	3129.549	38.634	0.861562	0.407257
s_8	0.003467	0.002158	0.415945	0.349786	3148.082	38.77594	0.868036	0.408742
s_9	0.002862	0.003237	0.412708	0.349098	3141.943	38.71274	0.870734	0.408081
s_{10}	0.001141	-0.00432	0.411629	0.347762	3129.776	38.64699	0.869115	0.407396

TABLE 6. An individual fault classification accuracy for motor 1 at (55 Hz and 40% loading) (10 known labels).

GSSL Algorithms	Edge reweighting schemes for GSSL	Classification accuracy, %				
		H vs. BF	H vs. (BF+1BRB)	H vs. (BF+UNB)	H vs. (BF+UNB+UV)	H vs. UNB
LGC	Binary weighting	96.7	93.5	98.6	99.1	92.3
	Fixed Gaussian kernel weighting	96.7	93.7	97.1	98.9	93
GFHF	Binary weighting	97.5	93.7	98.8	100	98.9
	Fixed Gaussian kernel weighting	97.2	95	98.8	100	98.9
GGMC	Binary weighting	100	100	99.5	100	100
	Fixed Gaussian kernel weighting	100	100	99.2	100	100

respectively. For this frequency localization, the feature coefficient includes the effect of motor operating conditions, and is related to WPD coefficients $x_{j,k}$ at a certain level j and node k as follows [30], [31]:

$$W(j,k) = \sqrt{\sum_{n=1}^{n=M} \frac{x_{j,k}^2(n)}{M}} \tag{6}$$

where M is the number of coefficients determined by the decomposition level and the length of the signal N as follows:

$$M = N2^{-j} \tag{7}$$

Because $d8$ is selected as the most significant decomposition level, the number of nodes at the level 8 is $2^8 = 256$. Therefore, there are 256 feature coefficients for each data window at $d8$. Eight statistical parameters, "mean", "median", "median absolute deviation", "mean absolute deviation", "L1 norm", "L2 norm", "maximum norm", and "standard deviation", are extracted from the 256 feature coefficients to serve as features in this study.

Tables 4 and 5 show the extracted features using the above approach for Motor 1 at (60 Hz and 80% loading) with a (BF+1BRB) fault, and Motor 2 at (55 Hz and 100% loading) with a (3BRB+UV) fault, respectively. One row in the table is obtained by processing sample stator current data within one data window, and ten rows correspond to ten data windows in a dataset.

V. FAULT CLASSIFICATION USING THE PROPOSED METHOD

The proposed method is used to classify an individual fault or five simultaneous faults for the two tested VFD-fed induction motors. Each motor has six defined classes: ("H", "UNB", "BF", "BF+UNB", "BF+1BRB", "BF+UNB+ UV") for Motor 1; and ("H", "UV", "1BRB", "2BRB", "3BRB", "UV+3BRB") for Motor 2.

TABLE 7. An Individual fault classification accuracy for motor 2 at (60 Hz and 20% loading) (10 known labels).

GSSL Algorithms	Edge reweighting schemes for GSSL	Classification accuracy, %				
		H vs. 1BRB	H vs. 2BRB	H vs. 3BRB	H vs. (3BRB+UV)	H vs. UV
LGC	Binary weighting	96.8	93.2	91	92.3	90
	Fixed Gaussian kernel weighting	96.8	89	93.2	92.5	90.4
GFHF	Binary weighting	97.8	100	93.8	94	91
	Fixed Gaussian kernel weighting	97.5	100	93.6	98.4	90.6
GGMC	Binary weighting	100	100	97	100	99.8
	Fixed Gaussian kernel weighting	100	100	97.4	100	100

TABLE 8. Fault classification accuracies for five simultaneous faults for motor 1 at (60 Hz and 80% loading) (30 known labels.

GSSL Algorithm	Edge reweighting schemes for GSSL	Classification accuracy in %
LGC	Binary weighting	77.13
	Fixed Gaussian kernel weighting	78.63
GFHF	Binary weighting	84.07
	Fixed Gaussian kernel weighting	89.5
GGMC	Binary weighting	93.13
	Fixed Gaussian kernel weighting	95.73

TABLE 9. Fault classification accuracies for five simultaneous faults for motor 2 at (60 Hz and 20% loading) (30 known labels).

GSSL Algorithm	Edge reweighting schemes for GSSL	Classification accuracy in %
LGC	Binary weighting	93.1
	Fixed Gaussian kernel weighting	94
GFHF	Binary weighting	97.8
	Fixed Gaussian kernel weighting	97.53
GGMC	Binary weighting	99.97
	Fixed Gaussian kernel weighting	98.9

TABLE 10. Classification average accuracies ± standard deviations with reference to label ratios for motor 1 at (60 Hz and 100% loading).

GSSL Algorithms	Edge weighting schemes for GSSL	Average accuracy ± standard deviation, %			
		Label ratio = 0.2	Label ratio = 0.3	Label ratio = 0.4	Label ratio = 0.5
LGC	Binary weighting	79.23 ± 3.17	81.88 ± 2.25	82.92 ± 1.95	84 ± 1.92
	Fixed Gaussian kernel weighting	80.29 ± 4.34	82.52 ± 3.08	83.64 ± 2.54	85.1 ± 2.32
GFHF	Binary weighting	85.9 ± 2.11	86.71 ± 2.06	88.97 ± 1.87	89.2 ± 1.44
	Fixed Gaussian kernel weighting	84.98 ± 2.14	87.43 ± 2.12	86.72 ± 1.84	87.03 ± 1.26
GGMC	Binary weighting	**87.83 ± 1.49**	**89.82 ± 1.32**	**90.94 ± 1.24**	**91.13 ± 0.63**
	Fixed Gaussian kernel weighting	86.06 ± 2.13	88.76 ± 2.01	89.42 ± 1.31	89.89 ± 1.19

TABLE 11. Classification average accuracies ± standard deviations with reference to label ratios for motor 2 at (65 Hz and 0% loading).

GSSL Algorithms	Edge weighting schemes for GSSL	Average accuracy ± standard deviation, %			
		Label ratio = 0.2	Label ratio = 0.3	Label ratio = 0.4	Label ratio = 0.5
LGC	Binary weighting	74.68 ± 7.57	76.95 ± 5.64	79.21 ± 4.61	80.7 ± 4.06
	Fixed Gaussian kernel weighting	75.42 ± 5.84	76.96 ± 4.39	78.17 ± 3.59	78.47 ± 3.11
GFHF	Binary weighting	83.9 ± 1.35	83.31 ± 1.31	84.06 ± 1.25	83.83 ± 1.01
	Fixed Gaussian kernel weighting	85.68 ± 2.47	86.14 ± 2.16	85.69 ± 1.84	87.4 ± 1.94
GGMC	Binary weighting	**93.33 ± 1.29**	**94.45 ± 1.28**	**94.6 ± 1.07**	**95.15 ± 0.95**
	Fixed Gaussian kernel weighting	89.17 ± 1.23	89.42 ± 0.76	90.03 ± 0.69	91.53 ± 0.65

A. CLASSIFICATION FOR AN INDIVIDUAL FAULT

An individual fault classification deals with one individual fault vs. the healthy case of the motor. An individual fault classification involves one faulty and one healthy cases, which requires 2 class labels within a data window; for a dataset with 10 data windows, the total number of class labels is $2 \times 10 = 20$ for a dataset.

In this case study, the number of known labels in the stator current datasets ranges from 2 (10% labeled data) to 10 (50% labeled data) for GGMC, LGC and GFHF. Tables 6 and 7 show the individual fault classification accuracy (with 10 known labels) for each of the five faults for Motor 1 at (55 Hz and 40% loading) and Motor 2 at (60 Hz and 20% loading). GGMC shows consistently

higher accuracy than LGC and GFHF in these chosen scenarios.

B. CLASSIFICATION FOR SEVERAL SIMULTANOUS FAULTS

To evaluate the fault classification performance dealing with several simultaneous faults, the proposed approach needs to classify accurately among five faults and one healthy case for each motor. A fault classification for five simultaneous faults for Motors 1 and 2 involves five faulty and one healthy cases, which requires 6 class labels within a data window; for a dataset with 10 data windows, the total number of class labels is $6 \times 10 = 60$ for a dataset.

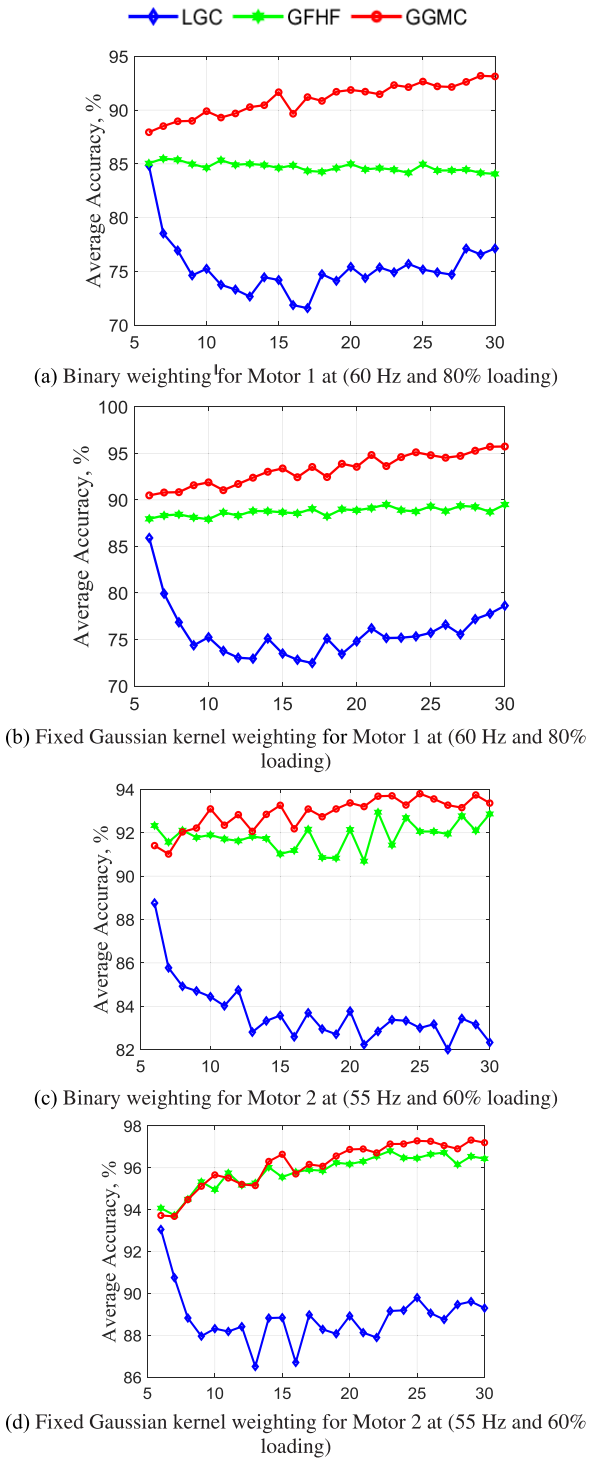
In this case study, the number of known labels in the stator current datasets ranges from 6 (10% labeled data) to 30 (50% labeled data) for GGMC, LGC and GFHF. Tables 8 and 9 show the accuracy (with 30 known labels) for Motor 1 at (60 Hz and 80% loading) and Motor 2 at (60 Hz and 20% loading), respectively, each motor has five simultaneous faults. In Table 8, GGMC's accuracy is 95.73% vs. LGC's 78.63% accuracy. GGMC shows significantly better performance than LGC and GFHF in these chosen scenarios.

To demonstrate the effect of the number of labels, Fig. 6 shows the classification accuracy vs. the number of labels for Motor 1 at (60 Hz and 80% Loading) and Motor 2 at (55 Hz and 60% Loading) for five simultaneous faults, where the number of labels ranges from 6 to 30. It also demonstrates the performance of the proposed method for different operating frequencies of the motors.

It is found that classification accuracies of the proposed approach using GGMC are not affected significantly by the number of labels, the GGMC curves are relatively flat, and the accuracy level of GGMC is significantly higher than GFHF and LGC. Fig. 6 also indicates that the proposed approach works well for the two different motor operating frequencies at 55 Hz and 60 Hz.

To further investigate the effect of the number of labels to a larger extent, Fig. 7 shows the classification accuracy for the number of labels ranging from 10% to 90% of the dataset for Motor 1 (60 Hz and 100% loading) and Motor 2 (65 Hz and 0% loading) for five simultaneous faults, where the label ratio (LR) is defined as: $LR = L/(L + U)$, L and U are the number of known and unknown labels, respectively. It is found that GGMC reaches the highest accuracy when LR $= 0.6$, further increasing LR slightly deteriorates GGMC's classification performance, and this trend is more pronounced for Motor 1 in Fig. 7(a). GGMC shows the superior performance compared to LGC and GFHF, a small amount of labels can achieve adequate accuracies. Fig. 7 also shows that the proposed approach works well for the two different motor operating frequencies at 60 Hz and 65 Hz.

Tables 10 and 11 show samples of classification average accuracies and standard deviations of the three GSSL algorithms ($LR = 0.2$, 0.3, 0.4 and 0.5) for Motor 1 at (60 Hz and 100% loading) and Motor 2 at (65 Hz and 0% loading), respectively. The average accuracy is obtained by averaging the accuracies of 100 iterations. GGMC shows less sensitivity towards changes in label ratios than LGC and GFHF. Binary weighting for GGMC performs better than fixed Gaussian kernel weighting in both tables.



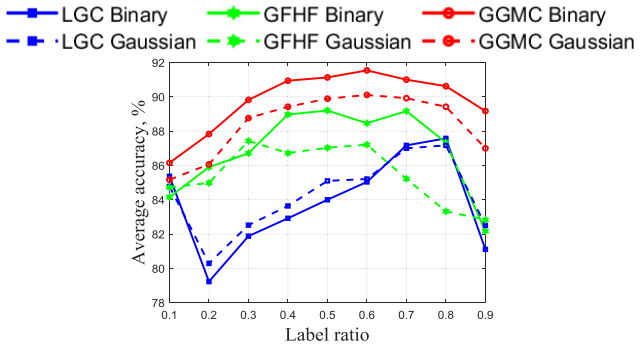
(a) Binary weighting for Motor 1 at (60 Hz and 80% loading)

(b) Fixed Gaussian kernel weighting for Motor 1 at (60 Hz and 80% loading)

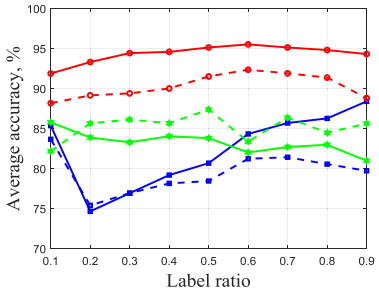
(c) Binary weighting for Motor 2 at (55 Hz and 60% loading)

(d) Fixed Gaussian kernel weighting for Motor 2 at (55 Hz and 60% loading)

FIGURE 6. Fault classification accuracy vs. the number of labels for the two motors.



(a) Motor 1 (60 Hz and 100% loading)



(b) Motor 2 (65 Hz and 0% loading)

FIGURE 7. Fault classification accuracy vs. label ratio.

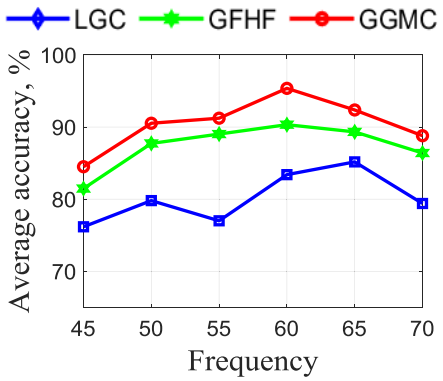
C. CLASSIFICATION ACCURACY FOR SEVERAL SIMULTANOUS FAULTS VS. VFD OUTPUT FREQUENCY

The influence of the motor operating frequency variations is one major concern regarding the developed fault diagnosis approach for VFD-fed induction motors. Is the proposed method effective for different motor operating frequencies? The previous examples in Figs. 5 and 6 show that the approach remains effective for the chosen frequencies, but they are just a few snapshots. To show a complete picture over the whole frequency band from 45 Hz to 70 Hz tested in the lab, Motors 1 and 2 are investigated as shown in Fig. 8, where classification accuracies for all five simultaneous faults for each motor using the three GSSL algorithms (with 30 known labels) vs. VFD output frequency ranging from 45 Hz to 70 Hz are displayed. It is found that the proposed approach using GGMC consistently shows higher accuracies than LGC and GFHF for both edge weighting schemes across the whole frequency band. Therefore, it is verified that the proposed method remains valid for potential motor operating frequencies in practical industrial applications.

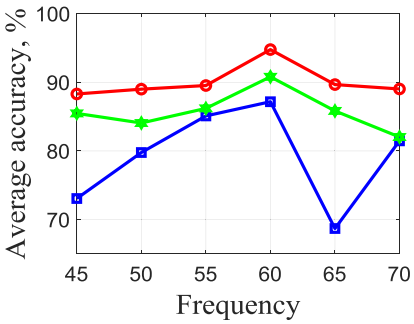
VI. FEATURES FORMULAS THROUGH SURFACE FITTING

A. FEATURES FORMULATION FOR UNTESTED MOTOR OPERATING CONDITIONS

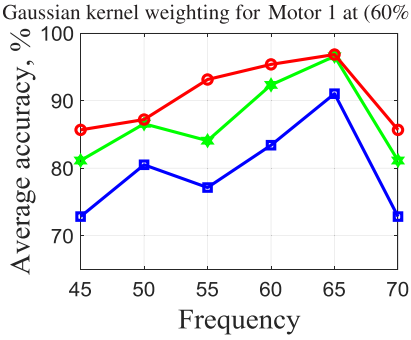
In this paper, the motors were tested in the lab at six VFD output frequencies and six motor loadings as shown in Table 1. In real life, the motor might run at other untested frequency or motor loading, and the test data would not be able to cover those operating conditions. To enable fault diagnosis of untested motor operating conditions, mathematical formulas



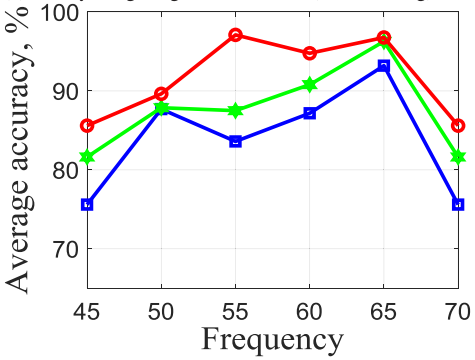
(a) Binary weighting for Motor 1 at (60% loading)



(b) Fixed Gaussian kernel weighting for Motor 1 at (60% loading)



(c) Binary weighting for Motor 2 at (40% loading)

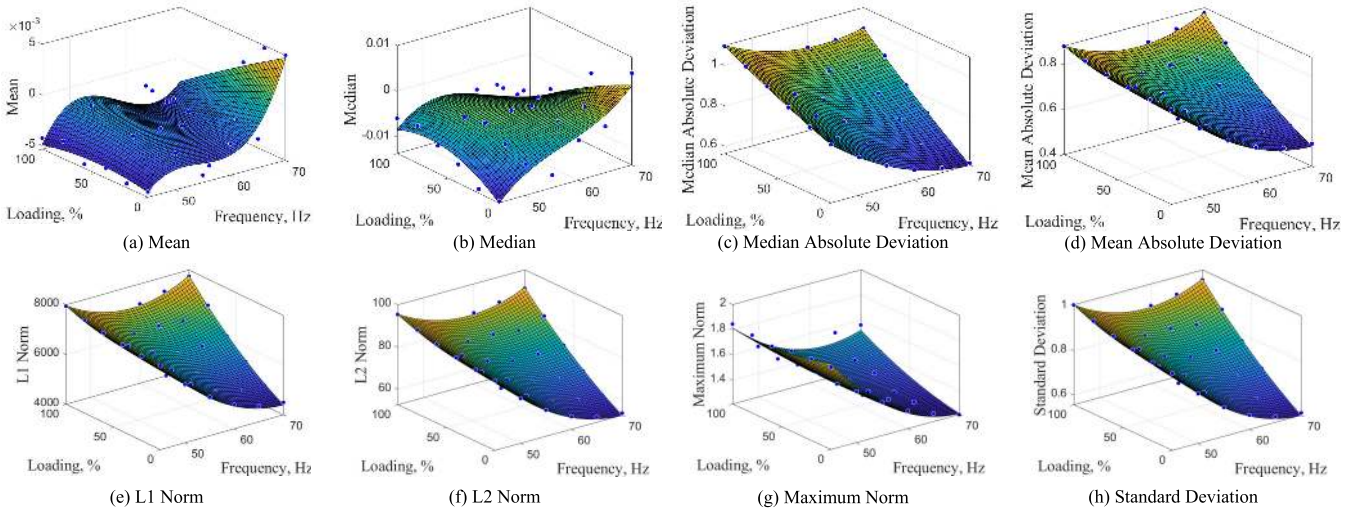


(d) Fixed Gaussian kernel weighting for Motor 2 at (40% loading)

FIGURE 8. Fault classification accuracy vs. VFD output frequency for two motors.

for features calculation are derived through the surface fitting technique using features extracted from experimental data of the tested cases.

Polynomial equations are developed through surface fitting to calculate the eight features: ''mean'', ''median'', ''median



(a) Mean (b) Median (c) Median Absolute Deviation (d) Mean Absolute Deviation

(e) L1 Norm (f) L2 Norm (g) Maximum Norm (h) Standard Deviation

FIGURE 9. Features vs. VFD output frequencies and motor loadings through the surface fitting for Motor 2 with a "3BRB" fault.

TABLE 12. Developed feature formulas for motor 2 with a "3BRB" fault.

Features	Equations	R-square
Mean	$f(x,y) = -0.3928 + 0.02159x - 0.000816y - 0.0003977x^2 + 3.3E-05xy - 6.112E-07y^2 + 2.435E-06x^3 - 3.182E-07x^2y + 7.929E-09xy^2$	0.7056
Median	$f(x,y) = -0.3522 + 0.01727x - 0.0002027y - 0.0002931x^2 + 2.217E-05xy - 3.359E-06y^2 + 1.7E-06x^3 - 2.96E-07x^2y + 4.273E-08xy^2$	0.6007
Median Absolute Deviation	$f(x,y) = 2.721 - 0.06078x - 0.001381y + 0.0004276x^2 + 7.523E-05xy + 4.67E-06y^2$	0.9973
Mean Absolute Deviation	$f(x,y) = 3.134 - 0.07523x - 0.006529y + 0.00053516x^2 + 0.0001278xy + 1.31E-05y^2$	0.9927
L1 Norm	$f(x,y) = 2.82E+04 - 677.1x - 58.77y + 4.816x^2 + 1.15xy + 0.1179y^2$	0.9927
L2 Norm	$f(x,y) = 307.3 - 7.215x - 0.5447y + 0.05114x^2 + 0.01176xy + 0.001096y^2$	0.9956
Maximum Norm	$f(x,y) = 6.35 - 0.1456x - 0.009188y + 0.001033x^2 + 0.0001506xy + 1.638E-05y^2$	0.9779
Standard Deviation	$f(x,y) = 3.239 - 0.07606x - 0.005743y + 0.0005391x^2 + 0.000124xy + 1.155E-05y^2$	0.9956

absolute deviation'', ''mean absolute deviation'', ''L1 norm'', ''L2 norm'', ''maximum norm'', and ''standard deviation''. The VFD output frequency and motor loading are two independent variables, and each statistical feature is the dependent variable. R-square values and relative errors between experimental and calculated features are used to evaluate accuracy of the developed fitting equations.

Table 12 shows the developed formulas for Motor 2 with a ''3BRB'' fault, where x is the frequency in Hz, y is the motor loading in %, and $f(x, y)$ represents the feature. Polynomial 32 equations are adopted for ''mean'' and ''median'', and Polynomial 22 equations are adopted for other six features. R-squares value close to 1 indicates a good fit. Among eight features, only ''mean'' and ''median'' have relatively lower

TABLE 13. Relative errors between experimental and calculated features for motor 2 with a "3BRB" fault at (50 Hz and 80% loading).

Features Name	Experimental feature	Calculated feature	Relative error, %
Mean	-0.001486	-0.001469	1.16
Median	-0.005412	-0.005390	0.4
Median Absolute Deviation	0.978175	0.971328	0.7
Mean Absolute Deviation	0.786897	0.783120	0.48
L1 Norm	7071.19	7037.96	0.47
L2 Norm	85.3049	84.8784	0.5
Maximum Norm	1.6345	1.6247	0.6
Standard Deviation	0.89881	0.89423	0.51

R-squares values, the rest six features has R-squares values very close to 1.

To validate the accuracy of the mathematical equations, relative errors between the feature directly extracted from experimental data and calculated from the developed equations are shown in Table 13 for Motor 2 with a ''3BRB'' fault at (50 Hz and 80% Loading). The highest relative error is 1.16%.

Fig. 9 depicts surface fitting models of the features versus the motor loading and the operating frequency of Motor 2 with a 3BRB fault. Dots in Fig. 9 are features processed by WPD using experimental data, and the fitting equation determines the surface. The similar procedure applies to other faults to develop feature formulas for untested motor operating conditions.

B. AN EXAMPLE OF FAULT DIAGNOSIS USING FITTING EQUATIONS

To demonstrate the effectiveness of the developed feature fitting equations for untested motor operating conditions, eight features of the following three untested conditions are

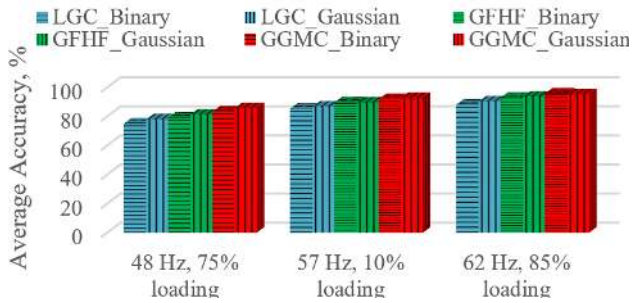


FIGURE 10. Fault classification for three untested conditions.

computed: (48 Hz and 75% loading), (57 Hz and 10% loading) and (62 Hz and 85% loading) for Motor 2 with five simultaneous faults. Fault classification based on the calculated features using GGMC, LGC, and GFHF are carried out. Fig. 10 shows classification accuracies for the three untested conditions. It is found that by using the proposed method, we are able to detect faults under a tested or untested motor operating frequency and motor loading for VFD-fed induction motors.

VII. CONCLUSION

In this paper, a fault diagnosis approach for VFD-fed induction motors is proposed using WPD and GGMC learning algorithm. As a critical step in this development, feature extraction from experimental stator currents using WPD is conducted by evaluating energy eigenvalues and feature coefficients at decomposition levels. The proposed approach remains valid for classification of an individual fault or several simultaneous faults, and also remains valid under different motor operating frequencies. To enable fault diagnosis for untested motor operating conditions, the calculation formulas for features are developed through surface fitting using features extracted from experimental data under tested motor operating conditions.

REFERENCES

[1] A. Bonnett and C. Yung, "Increased efficiency versus increased reliability," *IEEE Ind. Appl. Mag.*, vol. 14, no. 1, pp. 29–36, Jan. 2008.

[2] F. Briz, M. W. Degner, P. Garcia, and A. B. Diez, "High-frequency carrier-signal voltage selection for stator winding fault diagnosis in inverter-fed AC machines," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4181–4190, Dec. 2008.

[3] R. N. Andriamalala, H. Razik, L. Baghli, and F.-M. Sargos, "Eccentricity fault diagnosis of a dual-stator winding induction machine drive considering the slotting effects," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4238–4251, Dec. 2008.

[4] G. Bossio, C. D. De Angelo, J. Solsona, G. O. Garcia, and M. I. Valla, "Application of an additional excitation in inverter-fed induction motors for air-gap eccentricity diagnosis," *IEEE Trans. Energy Convers.*, vol. 21, no. 4, pp. 839–847, Dec. 2006.

[5] K. Teotrakool, M. J. Devaney, and L. Eren, "Adjustable-speed drive bearing-fault detection via wavelet packet decomposition," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 8, pp. 2747–2754, Aug. 2009.

[6] X. Huang, T. G. Habetler, and R. G. Harley, "Detection of rotor eccentricity faults in a closed-loop drive-connected induction motor using an artificial neural network," *IEEE Trans. Power Electron.*, vol. 22, no. 4, pp. 1552–1559, Jul. 2007.

[7] Y. Zhao, R. Ball, J. Mosesian, J.-F. de Palma, and B. Lehman, "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2848–2858, May 2015.

[8] D. Wu, X. Luo, G. Wang, M. Shang, Y. Yuan, and H. Yan, "A highly accurate framework for self-labeled semisupervised classification in industrial applications," *IEEE Trans. Ind. Informat.*, vol. 14, no. 3, pp. 909–920, Mar. 2018.

[9] W. Lin, Z. Gao, and B. Li, "Shoestring: Graph-based semi-supervised classification with severely limited labeled data," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 4173–4181.

[10] W. Liu, J. Wang, and S.-F. Chang, "Robust and scalable graph-based semisupervised learning," *Proc. IEEE*, vol. 100, no. 9, pp. 2624–2638, Sep. 2012.

[11] R. Razavi-Far, E. Hallaji, M. Farajzadeh-Zanjani, and M. Saif, "A semi-supervised diagnostic framework based on the surface estimation of faulty distributions," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1277–1286, Mar. 2019.

[12] W. Jiang, Z. Zhang, F. Li, L. Zhang, M. Zhao, and X. Jin, "Joint label consistent dictionary learning and adaptive label prediction for semisupervised machine fault classification," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 248–256, Feb. 2016.

[13] R. Razavi-Far, E. Hallaji, M. Farajzadeh-Zanjani, M. Saif, S. H. Kia, H. Henao, and G.-A. Capolino, "Information fusion and semi-supervised deep learning scheme for diagnosing gear faults in induction machine systems," *IEEE Trans. Ind. Electron.*, vol. 66, no. 8, pp. 6331–6342, Aug. 2019.

[14] K. Zhang, B. Tang, Y. Qin, and L. Deng, "Fault diagnosis of planetary gearbox using a novel semi-supervised method of multiple association layers networks," *Mech. Syst. Signal Process.*, vol. 131, pp. 243–260, Sep. 2019.

[15] J. Yuan and X. Liu, "Semi-supervised learning and condition fusion for fault diagnosis," *Mech. Syst. Signal Process.*, vol. 38, no. 2, pp. 615–627, Jul. 2013.

[16] S. M. K. Zaman, X. Liang, and H. Zhang, "Graph-based semi-supervised learning for induction motors Single- and multi-fault diagnosis using stator current signal," in *Proc. IEEE/IAS 56th Ind. Commercial Power Syst. Tech. Conf. (I&CPS)*, Jun. 2020, pp. 1–10.

[17] S. M. K. Zaman, X. Liang, and L. Zhang, "Greedy-gradient max cut-based fault diagnosis for direct online induction motors," *IEEE Access*, vol. 8, pp. 177851–177862, Sep. 2020.

[18] S. M. K. Zaman and X. Liang, "An effective induction motor fault diagnosis approach using graph-based semi-supervised learning," *IEEE Access*, vol. 9, pp. 7471–7482, Jan. 2021.

[19] M. Maier, U. V. Luxburg, and M. Hein, "Influence of graph construction on graph-based clustering measures," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1025–1032.

[20] T. Jebara, J. Wang, and S.-F. Chang, "Graph construction and b-matching for semi-supervised learning," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 441–448.

[21] Y. Xiu, W. Shen, Z. Wang, S. Liu, and J. Wang, "Multiple graph regularized graph transduction via greedy gradient max-cut," *Inf. Sci.*, vol. 423, pp. 187–199, Jan. 2018.

[22] X. Zhu, Z. Ghahramani, and J. D. Lafferty, "Semi-supervised learning using Gaussian fields and harmonic functions," in *Proc. 20th Int. Conf. Mach. Learn. (ICML)*, 2003, pp. 912–919.

[23] J. Wang, T. Jebara, and S.-F. Chang, "Semi-supervised learning using greedy max-cut," *J. Mach. Learn. Res.*, vol. 14, no. 1, pp. 771–800, Mar. 2013.

[24] Z.-J. Zha, T. Mei, J. Wang, Z. Wang, and X.-S. Hua, "Graph-based semi-supervised learning with multiple labels," *J. Vis. Commun. Image Represent.*, vol. 20, no. 2, pp. 97–103, Feb. 2009.

[25] A. A. Yusuff, C. Fei, A. A. Jimoh, and J. L. Munda, "Fault location in a series compensated transmission line based on wavelet packet decomposition and support vector regression," *Electr. Power Syst. Res.*, vol. 81, no. 7, pp. 1258–1265, Jul. 2011.

[26] A. Bouzida, O. Touhami, R. Ibtiouen, A. Belouchrani, M. Fadel, and A. Rezzoug, "Fault diagnosis in industrial induction machines through discrete wavelet transform," *IEEE Trans. Ind. Electron.*, vol. 58, no. 9, pp. 4385–4395, Sep. 2011.

[27] W. Ting, Y. Guo-Zheng, Y. Bang-Hua, and S. Hong, "EEG feature extraction based on wavelet packet decomposition for brain computer interface," *Measurement*, vol. 41, no. 6, pp. 618–625, Jul. 2008.

[28] J. Antonino-Daviu, M. Riera-Guasp, J. Roger-Folch, F. Martínez-Giménez, and A. Peris, "Application and optimization of the discrete wavelet transform for the detection of broken rotor bars in induction machines," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 2, pp. 268–279, Sep. 2006.
[29] C. Gan, J. Wu, S. Yang, Y. Hu, and W. Cao, "Wavelet packet decomposition-based fault diagnosis scheme for SRM drives with a single current sensor," *IEEE Trans. Energy Convers.*, vol. 31, no. 1, pp. 303–313, Mar. 2016.
[30] A. Sadeghian, Z. Ye, and B. Wu, "Online detection of broken rotor bars in induction motors by wavelet packet decomposition and artificial neural networks," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 7, pp. 2253–2263, Jul. 2009.
[31] Z. Ye, B. Wu, and A. Sadeghian, "Current signature analysis of induction motor mechanical faults by wavelet packet decomposition," *IEEE Trans. Ind. Electron.*, vol. 50, no. 6, pp. 1217–1228, Dec. 2003.

SHAFI MD KAWSAR ZAMAN (Member, IEEE) was born in Naogaon, Bangladesh. He received the B.Sc. degree in electrical and electronic engineering from the Islamic University of Technology, Gazipur, Bangladesh, in 2014, and the M.Eng. degree in electrical engineering from the Memorial University of Newfoundland, St. John's, NL, Canada, in 2020.

He was a Visiting Research Student with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, Canada, from September 2019 to July 2020. From February 2015 to August 2018, he served as a Lecturer with the Department of Electrical and Electronic Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh. In September 2020, he joined the College of the North Atlantic (Ridge Road Campus), St. John's, where he is currently an Instructor of Electronic Engineering Technology. His research interests include conditioning monitoring and fault diagnosis of electric machines using machine learning techniques.

XIAODONG LIANG (Senior Member, IEEE) was born in Lingyuan, Liaoning, China. She received the B.Eng. and M.Eng. degrees from Shenyang Polytechnic University, Shenyang, China, in 1992 and 1995, respectively, the M.Sc. degree from the University of Saskatchewan, Saskatoon, Canada, in 2004, and the Ph.D. degree from the University of Alberta, Edmonton, AB, Canada, in 2013, all in electrical engineering.

From 1995 to 1999, she served as a Lecturer at Northeastern University, Shenyang. In October 2001, she joined Schlumberger, Edmonton, and was promoted to be a Principal Power Systems Engineer with this world's leading oil field service company, in 2009. After serving Schlumberger for almost 12 years, from 2013 to 2019, she was with Washington State University Vancouver, USA, and the Memorial University of Newfoundland, St. John's, Canada, as an Assistant Professor and later as an Associate Professor. In July 2019, she joined the University of Saskatchewan, Saskatoon, Canada, where she is currently an Associate Professor and the Canada Research Chair in Technology Solutions for Energy Security in Remote, Northern, and Indigenous Communities. Her research interests include power systems, renewable energy, and electric machines.

Dr. Liang is a registered Professional Engineer in the province of Saskatchewan, Canada.

WEIXING LI (Senior Member, IEEE) was born in Zhengyang, China. He received the B.S. degree in applied physics and the M.S. and Ph.D. degrees in electrical engineering from the Harbin Institute of Technology, Harbin, China, in 1999, 2001, and 2007, respectively.

From July 2001 to September 2019, he was with the Harbin Institute of Technology, where he was an Assistant Professor and later an Associate Professor. From 2008 to 2010, he served as a Visiting PDF at the University of Alberta, Edmonton, AB, Canada, and the Lawrence Berkeley National Laboratory, UC Berkeley, USA. In September 2019, he joined the School of Electrical Engineering, Dalian University of Technology, Dalian, China, where he is currently a Professor. His research interests include power systems analysis, VSC-HVDC, and renewable energy.