Multi Fault Diagnosis of Centrifugal Pumps with Time, Frequency and Wavelet-Based Features Using Support Vector Machines

Janani Shruti Rapur and Rajiv Tiwari(✉)

Department of Mechanical Engineering,
Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
rtiwari@iitg.ernet.in

Abstract. The centrifugal pumps (CPs) are most commonly chosen fluid machines for industrial and domestic applications. They constitute vital components to sustain the process flow of the plants, and hence their failure may lead to a significant monetary loss to the plant. Failures in CPs may be due to mechanical faults and/or fluid flow anomalies. In the current work, it is attempted to build an adaptable support vector machine (SVM) based algorithm to identify critical faults, like flow restrictions (with changing severities), impeller cracks, dry run and cover plate faults. Furthermore, co-existence of mechanical and fluid flow faults is studied. Experimentally generated CP vibration data and motor current data is utilized for the fault diagnosis. The classifier parameters are chosen optimally by means of cross-validation method along with grid-search technique, and effective fault feature selection is done using the wrapper model. Furthermore, a comparative analysis on the performance of the methodology for features extracted from three domains, namely: time, frequency and wavelets (wavelet packet transform, time-frequency domain) of the raw signals is presented at a range of CP operating speeds. The analysis results presented that the developed methodology could identify fifteen CP fault conditions successfully based on features from all three domains at all CP speeds.

Keywords: Centrifugal pump · Flow-induced faults · Mechanical faults
Support vector machine · Multiclass fault classification
Time-domain, frequency domain · Wavelet packet transform

1 Introduction

The maintenance of critical plant equipment (such as CPs) is essential for its sustained operation. In addition, a good maintenance strategy ensures the safety of the technicians handling the equipment. Centrifugal pumps (CPs) are popularly used hydraulic machinery in many industries. It is estimated that, in a standard chemical plant, the number of CPs used approximately equals its employee headcount [1]. Also, roughly, twenty-percent of the energy generated globally is utilized in driving the CPs [2]. Thus, it can be safely said that CPs are critical plant equipment and their efficient working is essential for the persistent functioning of the plant.

K. L. Cavalca and H. I. Weber (Eds.): IFToMM 2018, MMS 61, pp. 43–57, 2019.
https://doi.org/10.1007/978-3-319-99268-6_4

The CP commonly breaks-down because of mechanical faults or due to fluid pattern anomalies or a combination of both of these. Authors have found that the hydraulic and mechanical CP faults are not independent of one another, and therefore, one fault can lead to the occurrence or augmentation of the other. Hence, it is essential that the faults are identified at their formative stages, so that corrective actions may be timely initiated. The primary mechanical faults in a CP are: bearing failures, seal imperfections, rotor misalignments, impeller faults, and unbalance. The fluid inconsistencies include flow recirculation, dry-runs, and cavitation.

In this study, it is attempted to segregate complex fault combinations of a CP. Some of the fault groupings (only two at a time) studied are presented in Table 1. The faults considered include, CP suction blockages (SBs), CP discharge blockages (DBs), CP impeller defects (IFs), CP cover-plate faults (PCs). Faults such as bearing defects or seal defects are not considered in this research. In the table, green color represents the presence of a fault condition and red represents the absence of it.

SBs are formed owing to a decrease in the cross-section area of the CP inlet, which may be due to, (a) obstruction in the suction pipe due to debris in fluid or (b) pipe surface damages. Flow over an obstruction on the suction side causes *discharge recirculation* [3]. This recirculation causes: cavitation like damage on impeller vane and on the casing. This also results in axial movement of shaft, cracks and impeller shroud malfunctions, shaft malfunction, excessive vibrations and noise.

Table 1. Considered fault combinations in the present work

Fault combination	**SB (Suction blockage)**	**DB (Discharge blockage)**	**IF (Impeller faults)**	**PC (Pitted cover)**
Healthy CP	Red	Red	Red	Red
Only SB	Green	Red	Red	Red
Only DB	Red	Green	Red	Red
Only IF	Red	Red	Green	Red
Only PC	Red	Red	Red	Green
Both IF and SB	Green	Red	Green	Red
Both IF and DB	Red	Green	Green	Red
Both PC and SB	Green	Red	Red	Green
Both PC and DB	Red	Green	Red	Green
Dry run	Green	Red	Red	Red

In general, discharge of CP is throttled to get the requisite flow rate. This may be considered to be DB condition. The decrease in flow leads to excessive loading on impeller vanes and also causes *suction recirculation* [3]. This also causes: cavitation of the vane, vibrations, and noise.

Most applications a designed to have CPs primed. However, at some instances the machinists do not recognize the sump dry up or absence of priming fluid in CPs, resulting in *dry run*. It leads to extreme heat generation following which the bearing and seal malfunctions may result, and extreme vibrations. The mechanical faults

(PC and IF) may be a consequence of fabrication errors and/or damages due to cavitation. When the passage of fluid inside the CP is disturbed a *pseudo-fluid flow recirculation* [4] is resulted.

The causes of each of the aforementioned faults are independent of one another. Nevertheless, an extended presence of one fault can initiate another. Or, they can even exist simultaneously and thus reduce the longevity of the machine. Therefore, the fault combinations are selected in such a way that the worst conditions may be represented. Such consideration of faults also helps in establishing the robustness of the designed technique.

The conventional maintenance strategies are based on, scheduled maintenance strategy or maintenance after occurrence of failure. The main drawbacks of such practices are, (i) complication in establishing the mode of failures, (ii) late identification of fault resulting in the propagation of fault and (iii) stoppage of plant process flow during maintenance. Hence, the contemporary maintenance strategies are based on intelligent plant maintenance. In this, a variety of operating parameters of critical equipment is acquired continuously and input to the designated machine learning. Any variation in the parameter values observed is attributed to corresponding failure mode.

Many researches who attempted CP fault diagnosis investigated independent existence of mechanical and fluid flow based faults. Muralidharan et al. researched on development of a technique to detect faults such as: cavitation, bearing defects, impeller cracks, and combination of bearing and impeller defects [5]. Sakthivel et al. proposed a methodology for detection of bearing defects, seal faults, impeller cracks, cavitation, and combination of bearing and impeller faults together [6]. Researchers also worked on identification of flow instabilities and cavitation (Chudina [7], Alfayez et al. [8], and Bordoloi and Tiwari [9]). However, the combined effects of the mechanical and fluid flow anomalies have rarely been considered or identified. In one such research, Perovic et al. [10] attempted classification of cavitation, discharge blockages, impeller cracks and coexisting discharge blockages and impeller cracks. In another work, Rapur and Tiwari [4] classified coexisting suction blockages and impeller faults. In both the mentioned studies, the classification accuracy dropped when fault combinations were considered. This may be because of the challenge in identifying intricate fault interactions.

There are numerous choices for machine learning (ML) algorithms that can be employed for CP fault diagnosis. They include decision tree algorithms [5], artificial neural networks (ANNs) [11], fuzzy logic based methods [10], support vector machines (SVMs) based methodologies [4, 9, 11, 12, 13] etc. The ANNs give high prediction performance when large fault data is available. Nevertheless, they succumb to local minima trap. In contrast to the other ML techniques, SVM is constructed based on the principle of structural risk minimization (SRM) and has a proper mathematical representation. Hence, it can arrive at a near global minimum and thus gives improved learning performance [14, 15]. Also, it develops a reliable classifier even with less fault data. Owing to these advantages, SVM is selected as the machine learning technique to develop the classification methods.

The data acquired from the CP can be used in three domains, i.e. in time, frequency, and time-frequency. The advantages or disadvantages of each of these domains rely on the behavior of the fault. In other words, it depends on the type of signal the fault produces. For example, a bubble explosion produces a short-lived transient signal, while a rotor crack or an unbalance produces a stationary-periodic signal. Hence, the choice of the domain plays a crucial role in the way the fault is interpreted and captured. Thus, this has a direct effect on the fault classification accuracy obtained. Therefore, in this current research, it is attempted to compare the fault classification accuracies obtained using each of these three domains, independently.

2 Centrifugal Pump Experimental Set-Up

Machine Fault Simulator (MFS^{TM}) laboratory setup as shown in Fig. 1 is used for data generation and experimentation. The pump is fixed on the base plate of the MFS^{TM} and is driven by a 'V' belt pulley drive using a variable speed induction motor. The CP is driven at distinct speeds: 30 Hz to 65 Hz in steps of 5 Hz. The healthy and faulty data from the CP is collected with the help of two accelerometers (tri-axial, one mounted on the bearing housing and the other on the pump casing) and three motor current sensors. Data-acquisition system (DAQ) loaded with LabVIEW software is used to configure the experimental data collection. Sampling rate of 5000 Hz with 5000 samples per dataset is used for the data collection. As previously discussed, even with small data SVM can build a reliable classifier. Therefore, hundred and fifty non-overlapping datasets for every considered fault condition are collected.

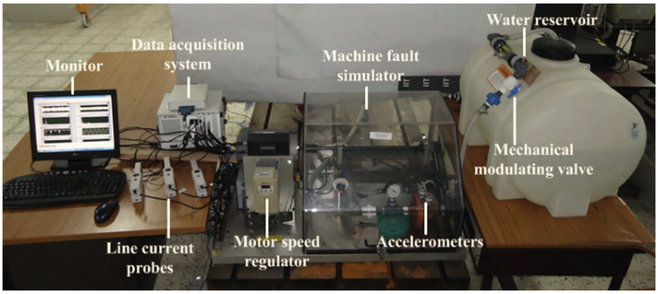

Fig. 1. CP laboratory experimental configuration

The faults on the CP are artificially created. For this effect, CPs with three different configurations are used, as shown in Table 2. Here also green color represents the presence and red represents the absence of the faults. On each mentioned configuration, blockages (suction and discharge) are administered over the pipe using modulating valves present on the inlet and outlet of the CP.

Table 2. Different CP configurations used for experimentation

CP Type	Healthy impeller	Healthy cover plate
Healthy pump		
Pump with impeller defects		
Pump with cover plate defects		

The faults considered are shown in Fig. 2. The valve can be opened or closed to the desired amount based on the amount of flow restriction required. On each of the impeller vanes two notches are given to create the impeller faults. Randomly located pits of different sized are given on the cove plate to create the cover plate defects.

On the whole, the CP is administered with thirty-three faults. They include, (i) healthy condition (HP), (ii) Blockages - suction (SBz), (iii) Blockages - discharge (DBz), (iv) impeller faults (IF), (v) co-existing impeller faults and suction blockages (IFSBz), (vi) co-existing impeller faults and discharge blockages (IFDBz), (vii) cover plate faults (PC), (viii) co-existing cover plate faults and suction blockages (PCSBz), (ix) co-existing cover plate faults and discharge blockages (PCDBz) and (x) dry runs for each of the CP configurations (SB5, PCSB5, IFSB5). Here, SBz and DBz represent $(z/6)^{th}$ level of suction and discharge blockage, respectively, z = 1, 2, 3, 4 and 5. For SB5 conditions there is no priming fluid inside the CP. Hence this can be considered to be a dry run state. From these thirty-three faults, the classification performance of fifteen faults, including: HP, SB2, SB5, DB1, DB5, IF, IFSB2, IFSB5, IFDB1, IFDB5, PC, PCSB2, PCSB5, PCDB1, PCDB5 are presented. These faults represent low and high severity levels of faults in each CP configuration. This fault set will be represented as AF (all faults) fault set in the rest of the paper.

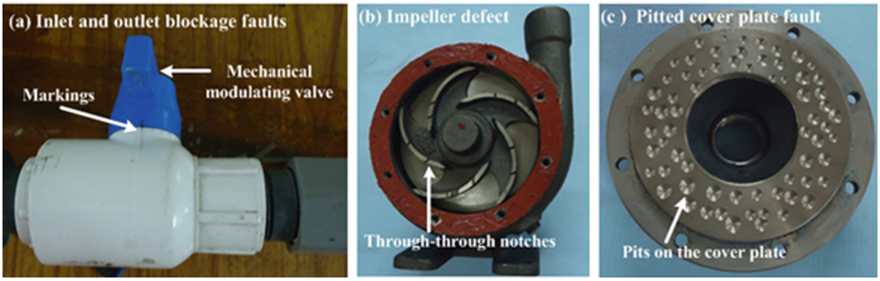

Fig. 2. Faults administered on the CP

3 CP Fault Segregation Methodology

This section briefly touches upon SVM classifier, deeper insights of SVM can be found in Refs. [14, 15]. For a binary classification using SVM, a hyperplane is constructed. This is oriented in such a way that the margin between the two classes of data is maximized, thus segregating them into positive class and negative class. The SVM construction is shown in Fig. 3.

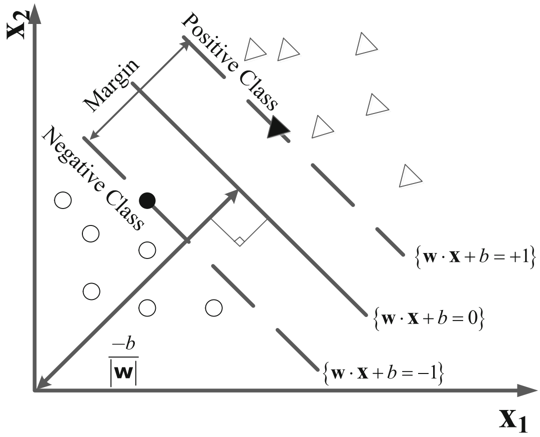

Fig. 3. Construction of an SVM classifier

If $\mathbf{x}_n$ is the closest point to the plane, it helps construct the hyperplane and thus divide the two classes of data. These are represented as black colored markers in Fig. 3. They are named *support vectors.* In continuation, the margin can be calculated to be $(\mathbf{w}^T\mathbf{w})^{-1}$, where $\mathbf{w}$ is the normal to the hyperplane. Hence, for given input vector $\mathbf{x}_p$ (p = 1, 2… N), N represents the quantity of samples; the optimal hyperplane can be obtained by solving the optimization problem,

$$\min\left[\frac{1}{2}\mathbf{w}^T\mathbf{w}+C\sum_{p=1}^{N}\xi_p\right] \tag{1}$$

$$\text{Subject to,}\quad y_p(\mathbf{w}^T\phi(\mathbf{x}_p)-b)\geq 1-\xi_p;\ \xi_p\geq 0 \tag{2}$$

where, p = 1, 2,…, N and y_p may be +1 or −1 depending on whether it is a positive class or a negative class, C is the penalty for error and the noise in slack variables is accounted by ξ_p. In the event the data is linearly inseparable, the training vector $\mathbf{x}_i$ is transformed to a higher dimensional space using a function φ, here SVM constructs a linear separating hyperplane. In addition, $K(\mathbf{x}_p,\mathbf{x})=\left\{\varphi^T(\mathbf{x})\varphi(\mathbf{x}_p)\right\}$ is called the *kernel* transformation.

Numerous kernel choices include: linear kernels, sigmoid kernels, polynomial kernels Gaussian radial basis function (RBF) kernels, etc. The choice of these kernels and their parameters is instrumental to obtain good classification performance. Present work utilizes the Gaussian RBF kernel, which is extremely adaptable [11]. It is given by the formula,

$$K(\mathbf{x}_p,\mathbf{x})=\exp\left(-\gamma\|\mathbf{x}-\mathbf{x}_p\|^2\right) \tag{3}$$

As mentioned earlier, SVM is a binary classifier. Therefore, to do a multiclass classification using SVM, the multiclass problem is converted to a number of binary class classifications, using any of the three approaches, one versus one (OVO), one versus all (OVA), and direct acyclic graph (DAG) [16].

The current research utilizes a OVO approach. In this, to sort k groups of data, $k(k-1)/2$ dualistic (binary) classifiers are constructed. Every binary classifier trains from two dissimilar classes. Thereafter, a voting tactic is employed to classify the data while testing. For example, for one of the developed binary classifier (in a multiclass classification problem), while testing a sample $\mathbf{x}$, say that the sample belongs to the m^{th} class, and then the number of votes for class m pertaining to test sample $\mathbf{x}$ is incremented by 1. At the end, sample $\mathbf{x}$ is classified into a class that has maximum number of votes in its favour.

To select the optimum SVM as well as kernel parameters, a h-fold cross-validation (CV) technique is used. Using this method, the training samples of the algorithm are equally distributed into 'h' sets. From these h sets $(h-1)$ sets are employed to train the algorithm and the other dataset is employed for testing. This procedure is repeated for each set of the h sets. Thus, that parameters combination that gives best classification accuracies is selected to form the final classifier. In this research, the value of h is considered to be 5. To quantify the performance of the classifier, classification accuracy is calculated. It is given by,

$$\text{Classification accuracy} = \left(\frac{\text{no. of accurately predicted data}}{\text{total no. of testing data}}\right) \times 100\% \quad (4)$$

As mentioned in Sect. 1, the performance of the algorithm using features extracted from time, frequency, and time-frequency domains are compared in this study. For all the three domains SVM would be used for multiple fault classification.

The originally captured fault data is in time-domain. The power spectral density (PSD) of the time-domain data is extracted to convert it into the frequency domain. This frequency domain extraction assumes the signal to be stationary. However, in the wavelet transform the functions localized in both time domain and frequency domain called the 'wavelets' are used. The wavelet packet function $W^n_{p,q}$ is defined as,

$$W^n_{p,q}(t) = 2^{\frac{p}{2}} W^n(2^p t - q) \quad (5)$$

where, the integers p and q represent the index of scale, translation operations, respectively. Also, index n is a modulation parameter. First, two wavelet packet functions, namely, the scaling function and the mother wavelet function, are defined as,

$$W^0_{0,0} = \phi(t); \quad W^1_{0,0} = \psi(t) \quad (6)$$

When n takes values 2, 3,… the functions are given as,

$$W_{0,0}^{2n}(t) = \sqrt{2}\sum_{q} h(q)W_{1,q}^{n}(2t-q);\ W_{0,0}^{2n+1}(t) = \sqrt{2}\sum_{q} g(q)W_{1,q}^{n}(2t-q) \quad (7)$$

here, $h(q)$ and $g(q)$ represent the quadrature mirror filters related with the predefined scaling and mother wavelet functions respectively. After defining the filters, recursive decomposition is applied to the signal, to obtain its hierarchical multi-resolution representation. The choice of the mother wavelet and the corresponding scaling function determines how intricately the signal details are captured in the time-frequency domain. In the present work, a two-level decomposition using db6 wavelet is done for the wavelet packet transform study.

4 CP Fault Classification - Results and Discussions

The feature calculation and selection is a vital step for good machine learning. In this research the features are obtained from the measured current and vibration signals obtained from the experiments. Features like, mean (μ), variance (μ_2), standard deviation (σ), root mean square (*RMS*), and entropy (e) are calculated from the raw data (in each domain). These features have standard mathematical descriptions. Apart from these features, a new feature *reciprocal of standard deviation* (σ_R) has been defined. It is nothing but the inverse of (σ) value. This feature is especially useful when the (σ) value is high. This reduces the range of variation in (σ) while retaining the trend of it.

These features values are acquired from, (1) acceleration amplitude variation with respect to time data, i.e. $x(t)$, and (2) power spectral density amplitude versus frequency data, i.e. $X(\omega_i)$ with $i = 1, 2,\ldots, N$ in the frequency domain. Hence, the values of the features are different in each of these domains. Furthermore, the time-frequency domain features are extracted from wavelet packet transform coefficients.

On the whole, 150 datasets of features are generated from the 150 raw datasets for each fault at the specific operating condition. 50% of this data is used for training, and the rest 50% for testing.

4.1 Feature Selection and Classification

To begin with, six features, namely, mean, variance, standard deviation, root mean square, reciprocal of standard deviation and entropy are extracted from the data. To select the best feature or feature combinations, an accuracy based method called the wrapper model is used. This method is very simple and intuitive. The feature selection is done using the fault diagnosis results at a single speed of CP operation. Figure 4 shows the typical cross validation plot for a CP fault diagnosis.

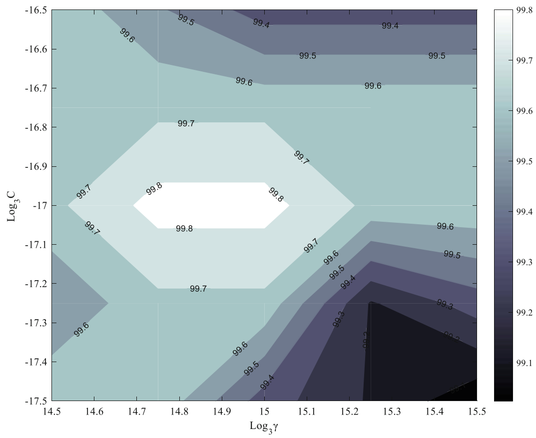

Fig. 4. Typical cross-validation plot for CP fault diagnosis

After CV and grid-search, the best grouping of SVM and kernel parameters (giving the best training accuracy) are selected. Additionally, these parameters are utilized to construct a final SVM classifier. Using these features and the final SVM classifier, fault diagnosis is performed at many CP speeds of operation in each of the three domains.

CP Fault Classification – Time-Domain. Here, the diagnosis is done using the features extracted from the time-domain. AF fault set is considered. The average classification accuracy at a given speed is described as average classification accuracies of all the faults. Also, the overall classification accuracy is the mean of average classification accuracies obtained at each speed. For selecting the best feature(s) among the six listed features, wrapper model is used. The results of the analysis are shown in Fig. 5. From the figure, it is observed that all the features give a accuracy of more than 90%. However, the μ and σ_R features give the best classification performance of approximately 100%. Therefore, both these features are finalized and will be used together.

Now, using these finalized features, the fault classification of the aforementioned faults is performed. The results of the fault classification are presented in Fig. 6. The overall classification accuracy of this classification over the entire speed range is found to be 99.9%. HP, SB2, SB5, DB5, IFDB1, IFDB5, PCSB5 and PCDB5 faults demonstrate cent percent classification at all the speeds. The other faults, however, miss the 100% classification at certain speeds. The average classification accuracy at each speed is more than 96%. The best performance of classification (with an average classification accuracy of 100%) is found at speeds of 30 Hz, 50 Hz, 55 Hz and 60 Hz. However, the classifier seems to be very adaptable and versatile to identify all the faults and is unaffected by the operating CP speed.

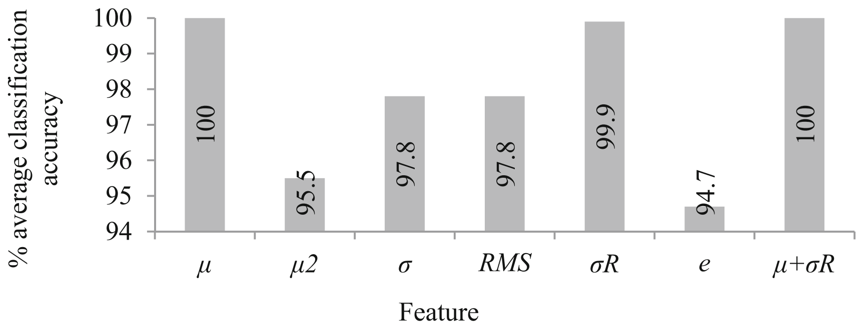

Fig. 5. Feature selection in time-domain

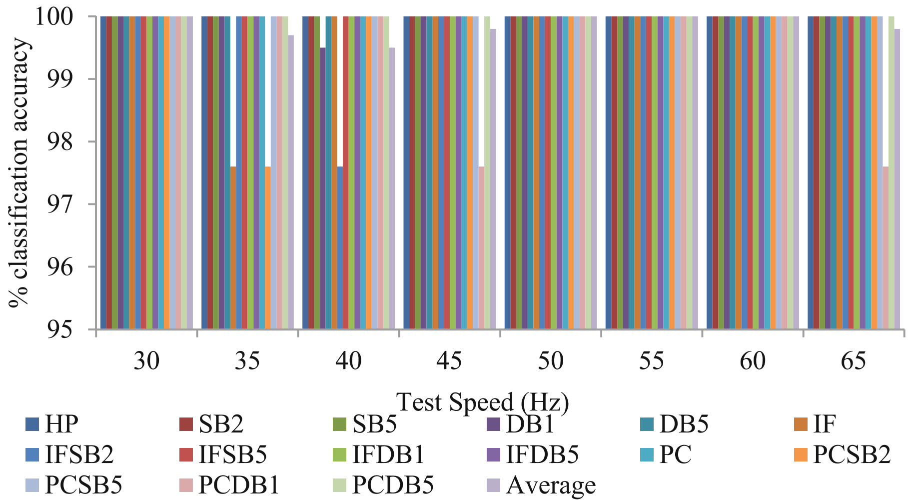

Fig. 6. Fault diagnosis based on time-domain features

CP Fault Classification – Frequency-Domain. The analysis in frequency domain caters to very good classification especially while dealing with faults that produce periodic signals. Impeller defects, for example, produce fault harmonics at, $BPF = n \times \omega$. Where n is the number of impeller blades, ω is the CP operating speed and BPF is the blade pass frequency. Therefore, frequency domain analysis is very useful. For the selection of the best features in the frequency domain, the wrapper model is used. The results of the fault classification for feature selection are presented in Fig. 7.

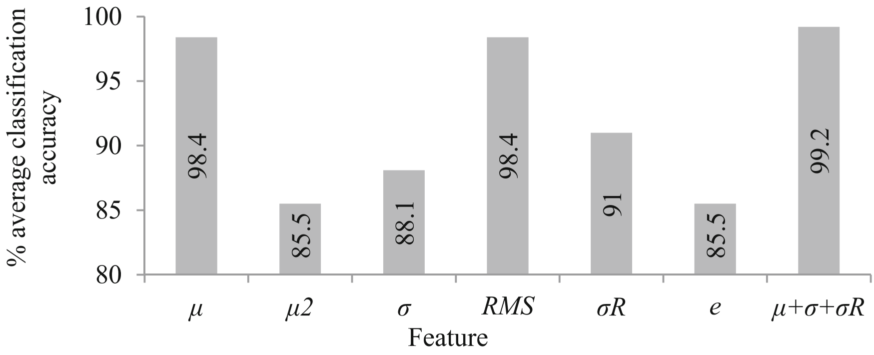

Fig. 7. Feature selection in frequency-domain

From the figure, it can be seen that μ, *RMS*, σ, and σ_R give a classification accuracy of more than 90%. However, when different permutations and combinations of these features were tried out, it was found that the combination of μ, σ and σ_R gave the best prediction. Hence, these three features are finalized.

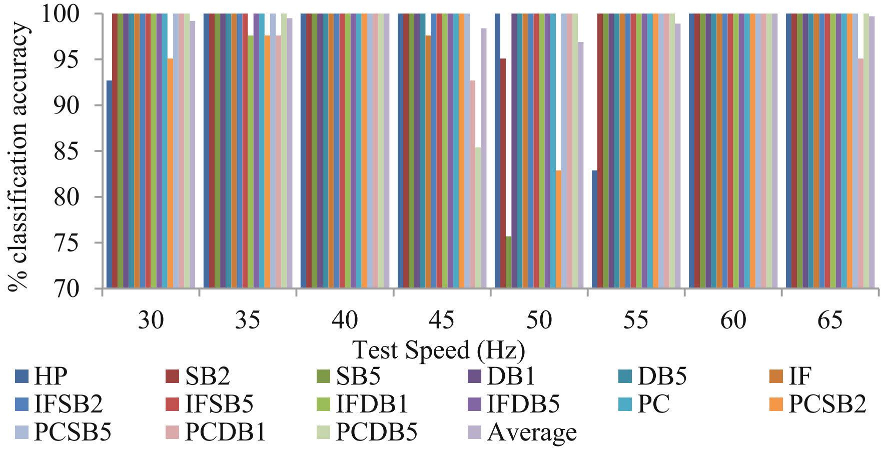

Fig. 8. Fault diagnosis based on frequency-domain features

With the finalized features the classification of the faults shown in the previous section is performed. The classification results are presented in Fig. 8. The classification shows more than 96.9% accuracy at all the speeds. The overall classification accuracy is 99.1%. Faults like, DB1, DB5, IFSB2, IFSB5, PC and PCSB5 faults give a 100% classification at all the speeds. Also, at 40 Hz and 60 Hz, the average classification accuracy is 100%. Here also, the classification performance is independent of the CP operating speed. Though recirculation is a transient phenomenon, the features from the frequency domain are able to classify the faults with good accuracy.

CP Fault Classification – Time-Frequency-Domain. In this section, the fault classification based on wavelet analysis is presented. db6 wavelet belonging to the Daubechies family is used for the signal decomposition. The fault classification results for feature finalization are presented in Fig. 9. From the figure, it may be observed that the classification using any of the six features is giving a good accuracy. However, the further trial revealed that the feature combination of μ, σ, e and RMS gave the best performance.

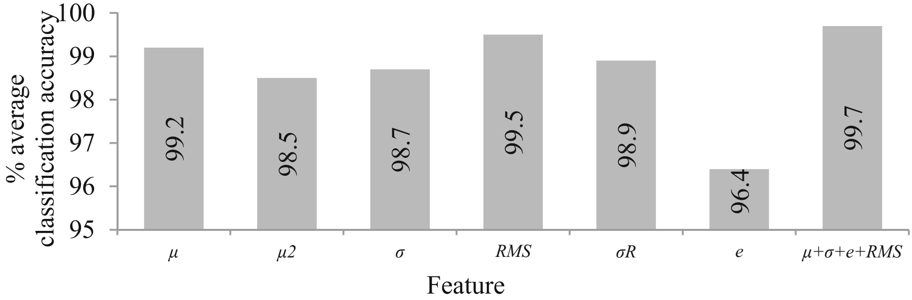

Fig. 9. Feature selection in time-frequency domain

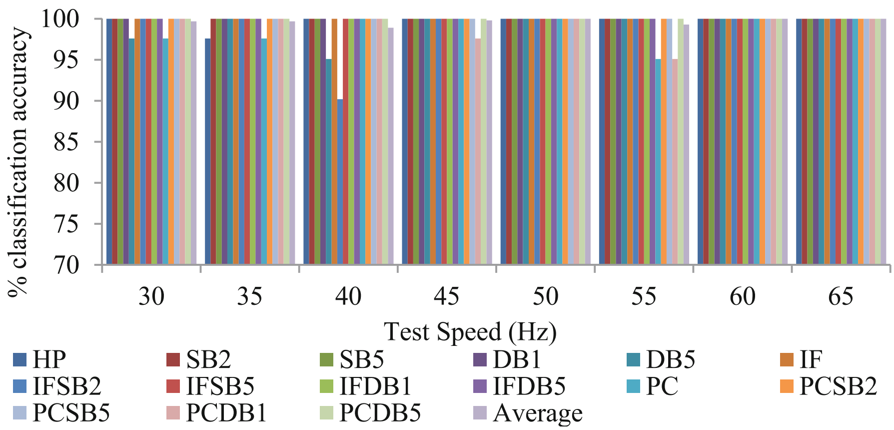

Fig. 10. Fault diagnosis based on time-frequency domain features

With the help of the finalized features, the classification of faults is performed and the results are presented in Fig. 10. At every speed the classification accuracy is above 98.9%. The overall classification accuracy, however, is found to be 99.7%. Faults such as SB2, SB5, DB1, IF, IFSB5, IFDB1, IFDB5, PCSB2, PCSB5, and PCDB5 show a 100% classification at all the speeds. Average classification accuracy is found to be 100% at 50 and 65 Hz speeds.

Figure 11 shows the comparison of the performance of the three domains. In general, it can be seen that features from all the three domains perform on comparable levels. However, owing to the transients produced due to the fluid-flow abnormalities, the prediction performance with frequency domain features is slightly lower than the other two domains.

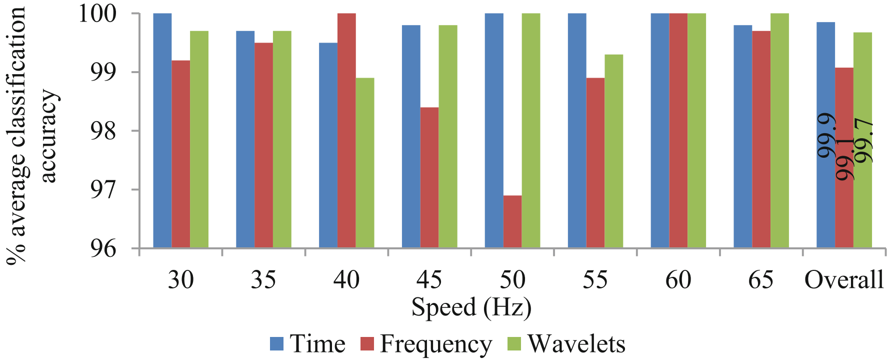

Fig. 11. Comparison of the classification accuracies obtained using features from time, frequency and time-frequency domains.

CP Fault Classification – Intermediate Speed. All the analyses performed so far considered the availability of history of fault data at the current operating condition of the CP. However, there may be cases in which the fault data may be unavailable at the current running condition of the CP. In case, the data is available at speeds higher and lower than the current condition, this method may be used. Here, the algorithm is trained with data at two distinct speeds and is tested with data at an intermediate speed. Figure 12 shows the comparison of the classification performances of the three domains. In figure, 30, 40/35 represents training with data obtained at 30 and 40 Hz CP speed and testing with data at 35 Hz, respectively.

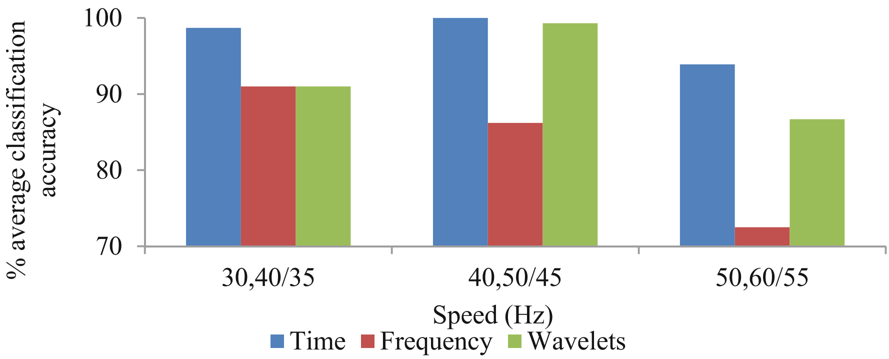

Fig. 12. Comparison of the classification accuracies obtained using features from time, frequency and time-frequency domains at intermediate speed.

From the figure it can be seen that the best classification performance is observed to be at 40, 50/45 case for time and time-frequency domain features, while it is at 30, 40/35 with the frequency domain features. However, with the increase in speeds the fault manifestation is expected to change largely, and therefore the classification performance with the three domains is decreasing at higher speeds. In total, the time domain features seem to be giving the most satisfactory performance with classification at intermediate speeds.

5 Conclusions

The condition monitoring of rotating machinery is crucial for safe operation of the system and automation [17]. The research demonstrated a comparative study of CP fault diagnosis in three domains, viz. time, frequency and time-frequency using features extracted from CP vibration and motor line current signatures. The hydraulic and mechanical faults existing independently and in combinations with changing severities are considered. In order to establish the versatility of the offered methodologies, the diagnoses are performed at a range of distinct CP speeds. Features like mean (μ), variance (μ_2), standard deviation (σ), root mean square (RMS), and entropy (e) obtained from the raw data (in each of the domains). The best performing features are then selected in each of the domains separately using wrapper model. Finally, SVM is employed to do the multi class fault diagnosis. The results of developed methodologies show that they could successfully classify all the mechanical and hydraulic faults in all the domains. However, the classification accuracies obtained using time-domain features slightly outperformed the other two domains. The algorithm was able to classify faults at intermediate test speeds as well, with satisfactory accuracy. The future scope of this work can be to use real-time data from industries to check the applicability of the present methodology. Also, the performance of other time-frequency domain approaches, like, Hilbert-Huang transform (HHT), Pseudo Wigner-Ville and Continuous Wavelet Transform could also be investigated.

Acknowledgements. The authors are profoundly grateful to the financial support and infrastructure provided by IIT Guwahati in helping this research take its current shape. Likewise, they would like to recognize and appreciate the LIBSVM tool, which was very beneficial in carrying out present study: https://www.csie.ntu.edu.tw/~cjlin/libsvm/ [18].

References

1. Hennecke, F.W.: Reliability of pumps in chemical industry. In: Pump Users International Forum, Karlsruhe, Germany (2000)
2. Hart, R.J.: Pumps and their systems-a changing industry. In: Proceedings of the International Pump Users Symposium, Houston, Texas, U.S.A., pp. 141–144 (2002)
3. Frazer, H.: Flow recirculation in centrifugal pumps. In: ASME Meeting (1981)
4. Rapur, J.S., Tiwari, R.: Experimental time-domain vibration-based fault diagnosis of centrifugal pumps using support vector machine. ASCE-ASME J. Risk Uncertain. Eng. Syst. Part B: Mech. Eng. **3**(4), 044501–044507 (2017)

5. Muralidharan, V., Sugumaran, V.: Feature extraction using wavelets and classification through decision tree algorithm for fault diagnosis of mono-block centrifugal pump. Measurement **46**, 353–359 (2013)
6. Sakthivel, N.R., Sugumaran, V., Babudevasenapati, S.: Vibration based fault diagnosis of monoblock centrifugal pump using decision tree. Expert Syst. Appl. **37**, 4040–4049 (2010)
7. Chudina, M.: Noise as an indicator of cavitation in a centrifugal pump. Acoust. Phys. **49**, 463–474 (2003)
8. Alfayez, L., Mba, D., Dyson, G.: The application of acoustic emission for detecting incipient cavitation and the best efficiency point of a 60 kW centrifugal pump: case study. NDT & E Int. **38**(7), 354–358 (2005)
9. Bordoloi, D.J., Tiwari, R.: Identification of suction flow blockages and casing cavitations in centrifugal pumps by optimal support vector machine techniques. J. Braz. Soc. Mech. Sci. Eng. **39**, 1–12 (2017)
10. Perovic, S., Unsworth, P.J., Higham, E.H.: Fuzzy logic system to detect pump faults from motor current spectra. In: Conference Record of the 2001 IEEE Industry Applications Conference. 36th IAS Annual Meeting (Cat. No. 01CH37248), vol. 1, pp. 274–280 (2001)
11. Azadeh, A., Saberi, M., Kazem, A., Ebrahimipour, V., Nourmohammadzadeh, A., Saberi, Z.: A flexible algorithm for fault diagnosis in a centrifugal pump with corrupted data and noise based on ANN and support vector machine with hyper-parameters optimization. Appl. Soft Comput. **13**, 1478–1485 (2013)
12. Rapur, J.S., Tiwari, R.: A compliant algorithm to diagnose multiple centrifugal pump faults with corrupted vibration and current signatures in time-domain, p. V002T05A007 (2017)
13. Rapur, J.S., Tiwari, R.: Automation of multi-fault diagnosing of centrifugal pumps using multi-class support vector machine with vibration and motor current signals in frequency domain. J. Braz. Soc. Mech. Sci. Eng. **40**, 278 (2018)
14. Widodo, A., Yang, B.S.: Support vector machine in machine condition monitoring and fault diagnosis. Mech. Syst. Signal Process. **21**, 2560–2574 (2007)
15. Vladimir, V.N., Vapnik, V.: The Nature of Statistical Learning Theory. Springer, Heidelberg (1995)
16. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classification, pp. 1–16 (2003)
17. Tiwari, R.: Rotor systems: Analysis and Identification. CRC Press, Boca Raton (2017)
18. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**, 1–27 (2011)