Optimal Feature and Decision Tree-Based Classification of Power Quality Disturbances in Distributed Generation Systems

Prakash K. Ray, *Student Member, IEEE*, Soumya R. Mohanty, *Senior Member, IEEE*, Nand Kishor, *Senior Member, IEEE*, and João P. S. Catalão, *Senior Member, IEEE*

***Abstract*—Penetration of distributed generation systems in conventional power systems leads to power quality (PQ) disturbances. This paper provides an improved PQ disturbances classification, which is associated with load changes and environmental factors. Various forms of PQ disturbances, including sag, swell, notch, and harmonics, are taken into account. Several features are obtained through hyperbolic S-transform, out of which the optimal features are selected using a genetic algorithm. These optimal features are used for PQ disturbances classification by employing support vector machines (SVMs) and decision tree (DT) classifiers. The study is supported by three different case studies, considering the experimental setup prototypes for wind energy and photovoltaic systems, as well as the modified Nordic 32-bus test system. The robustness and precision of DT and SWM are performed with noise and harmonics in the disturbance signals, thus providing comprehensive results.**

***Index Terms*—Classification, decision tree (DT), distributed generation (DG), HS-transform (HST), power quality (PQ), support vector machines (SVMs).**

I. Introduction

RECENTLY, distributed generation (DG) has become widely used because of the advantages associated with this green form of energy [1]. Wind energy and photovoltaic (PV) systems are environmentally friendly and considered as leading energy resources for the future. However, high DG penetration offers technical and operational challenges to the power engineers, causing power quality (PQ) and stability problems.

Output stability is a prime concern when systems operate in an isolated or grid-connected mode as per the requirements [2]. Serious concerns have been raised in view of the impact from wind and PV resources on PQ. Therefore, these resources are usually integrated with other DG sources, such as fuel cell and diesel generator, alongside storage devices like batteries, flywheels, and ultracapacitors to augment the quality/stability of the system [3], [4].

The behavior of the PV system is influenced by various factors, including solar strength, the temperature of the cell, and possible shading [5]. Similarly, wind energy systems depend upon the input wind speed, tower shadow effect, among other factors [6].

These environmental factors along with variations in load, capacitor switching, charging of transformers, starting of induction machines, use of nonlinear loads, and welding transformers lead to PQ problems such as sag, swell, notch, harmonics, etc. In the past, many researchers had highlighted and studied the PQ problems due to the variations in linear/nonlinear load; however, the disturbances occurred due to environmental factors, such as deviations in wind speed and solar irradiance may also lead to various operational issues. This includes mal-operation of protective devices, failure and overloading of electrical equipment, instabilities, and so on [7], [8]. For example, when wind/PV systems are interfaced to the grid with the help of dc/dc and dc/ac converters, and maximum power point tracking controllers are incorporated into these systems, system complexity increases further to tackle PQ problems.

As a matter of fact, all these scenarios of disturbances must be addressed in order to make power system operations and control more robust with high penetration of renewable resources. The presence of DG in the system opens up a new challenge to power engineers and researchers to address the issues related to PQ disturbances.

In the past studies, PQ indices, such as peak values, crest factor, total harmonic distortion (THD), power factor, instantaneous frequency, and energy deviation, were calculated using frequency spectrum or Parseval's theorem for monitoring of the disturbances [7]. The techniques, such as fast Fourier transform (FFT), chirp Z-transform, Welch algorithm, and zoom FFT, have been widely used for monitoring of electrical parameter [8], [9]. But, sometimes these techniques lead to misclassification of the disturbances [9]. For example, FFT is not accurate in the analysis of nonstationary disturbances including voltage notch and transients.

Thus, time-frequency analysis, namely wavelet transform (WT), short-time Fourier transform (STFT), and Kalman filters, are being extensively employed for detection and classification problems [10], [11].

Manuscript received February 22, 2013; revised July 07, 2013; accepted August 10, 2013. This work was supported by FEDER funds (European Union) through COMPETE and by Portuguese funds through FCT, under Projects FCOMP-01-0124-FEDER-020282 (Ref. PTDC/EEA-EEL/118519/2010) and PEst-OE/EEI/LA0021/2013. The research leading to these results was supported by the EU Seventh Framework Programme FP7/2007-2013 under Grant 309048.

P. K. Ray is with the International Institute of Information Technology, Bhubaneswar 751003, India (e-mail: pkrayiiit@gmail.com).

S. R. Mohanty and N. Kishor are with the Motilal Nehru National Institute of Technology, Allahabad 211004, India (e-mail: soumya@mnnit.ac.in; nand_research@yahoo.co.in).

J. P. S. Catalão is with the University of Beira Interior, 6201-001 Covilha, Portugal, with INESC-ID, 1000-029 Lisbon, Portugal, and also with IST, University of Lisbon, 1049-001 Lisbon, Portugal (e-mail: catalao@ubi.pt).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSTE.2013.2278865

Yet, the main demerit of WT is its incapacity to sense disturbances when there is signal noise [11]. Hence, modified WT named S-transform (ST) was developed for improving the detection performance [12].

Curve-fitting, magnitude, and frequency estimation approaches, such as Kalman filtering, the recursive least squares algorithm, non-recursive Newton algorithm, and recursive Newton algorithm, were also used to detect PQ disturbances [13], [14]. But, when the system nonlinearities are considered, extended and unscented Kalman filters are proved to provide better results [15].

Similarly, the classification of PQ disturbances is studied by researchers using soft computing approaches such as artificial neural network (ANN), fuzzy and neuro-fuzzy systems [16]–[18]. But these classification methods are based on the raw disturbance signal data that might augment computational burden. Therefore, features such as mean, variance, standard deviation (STD), THD, energy, and entropy are extracted through STFT, ST, and Hilbert transform. These feature datasets are then fed to pattern recognition techniques, such as ANN, probabilistic neural network, neuro-fuzzy, DT, and support vector machines (SVMs), to classify the disturbances [19]–[27].

The authors in [27] have discussed the classification of PQ disturbances caused due to changes in load, solar irradiance, and wind speed in a renewable resources-based hybrid DG system. Based on the results reported in [27], the authors proposed features extraction using hyperbolic S-transform (HS-transform) [22] followed by the classification of PQ disturbances using decision tree (DT) [23], [24] or SVM [25], [26] in a system with DG. HS-transform is considered for its superior time/frequency resolution and proficiency in detecting and localizing the disturbances, even coexisting with noise.

Similarly, DT and SVM are selected because of their classification ability with higher accuracy. The optimal set of features, selected by GA, is fed to these classifier algorithms. DT partitions the input feature dataset into different classes based on thresholds, whereas classification by SVM is based on nonlinear decision boundaries called hyperplanes. Parameters of SVM are obtained by cross-validation.

The paper is structured as follows. Section II introduces the approach for classification. Section III addresses the influence of environmental factors on voltage signal. Optimal feature selection by GA is provided in Section IV. The results using DT/SVM are given in Section V. Section VI concludes the paper.

II. Classification Approach

The descriptions of HS-transform (for detection and feature extraction), DT, and SVM (for PQ disturbances classification) are detailed as follows.

A. HS-Transform

ST, a modified WT with phasor correction, is an invertible time-frequency localization multi-resolution analysis, which is formulated using WT and STFT. Its analysis considers a variable Gaussian window whose width is inversely proportional to frequency. Sometimes, ST with a Gaussian window fails to localize transient disturbances in the presence of noise. Thus, HS-transform with pseudohyperbolic Gaussian window provides enhanced time/frequency resolutions in low–high frequencies. A higher window asymmetry at low frequencies increases the width in frequency domain [22].

The hyperbolic window is expressed as

$$W_{hb} = \frac{2f_s}{\sqrt{2\pi(\alpha_{hb}+\beta_{hb})}} \exp\left\{-\frac{-f_s^2 X^2}{2}\right\} \tag{1}$$

where

$$X = \frac{(\alpha_{hb}+\beta_{hb})}{2\alpha_{hb}\beta_{hb}}(\tau - t - \xi) + \frac{(\alpha_{hb}-\beta_{hb})}{2\alpha_{hb}\beta_{hb}}\sqrt{(\tau - t - \xi)^2 + \lambda_{hb}^2}, \qquad 0<\alpha_{hb}<\beta_{hb}$$

$$\text{and } \xi = \frac{\sqrt{(\beta_{hb}-\alpha_{hb})^2\lambda_{hb}^2}}{4\alpha_{hb}\beta_{hb}}. \tag{2}$$

The discrete version of HS-transform is calculated and $G(m_F, n_F)$ denotes Fourier transform of hyperbolic window

$$G(m_F, n_F) = \frac{2f_s}{\sqrt{2\pi(\alpha_{hb}+\beta_{hb})}} \exp\left\{-\frac{-f_s^2 X_D^2}{2}\right\} \tag{3}$$

where

$$X_D = \frac{(\alpha_{hb}+\beta_{hb})}{2\alpha_{hb}\beta_{hb}}t + \frac{(\alpha_{hb}-\beta_{hb})}{2\alpha_{hb}\beta_{hb}}\sqrt{t^2+\lambda_{hb}^2}. \tag{4}$$

$H[m_F, n_F]$ is the frequency shifted Fourier transform $H[m_F]$ and is given by

$$H[m_F] = \frac{1}{N}\sum_{m_F=0}^{N-1} h(k)\exp(-i2\pi n_F k) \tag{5}$$

$$S[n_F, j] = \sum_{m_F=0}^{N-1} H(m_F+n_F)G(m_F, n_F)\exp(-i2\pi m_F j). \tag{6}$$

B. Decision Tree

Data classification is the process of checking the similarities in a dataset and to classify them into distinct classes.

DTs [23], [24] are widely used in the classification based on the choice of an attribute that maximizes and fixes data division. These attributes are split into several branches recursively, until the termination and classification is reached.

The mathematical illustration of DT algorithm is constructed upon the subsequent definitions

$$\bar{X} = \{X_1, X_2, \ldots, X_m\}^T \tag{7}$$

$$X_i = \{x_1, x_2, \ldots, x_{ij}, \ldots, x_{in}\} \tag{8}$$

$$S = \{S_1, S_2, \ldots, S_i, \ldots, S_m\} \tag{9}$$

where m is the available observations number, n is the independent variables number, S is the m-dimension vector of the variable forecasted from $\bar{X}$, X_i is the ith component vector of n-dimension autonomous variables, $x_{i1}, x_{i2}, \ldots, x_{ij}, \ldots, x_{in}$ are autonomous variables of pattern vector X_i, and T is the transpose notation vector.

The objective of DT is to forecast S based on the observation of $\bar{X}$. As several DTs with different accuracy levels can be built from $\bar{X}$, obtaining the optimal tree is challenging due to the large search space dimension. But suitable algorithms can be established for DT that reflect a tradeoff between complexity and accuracy. These algorithms create a DT by a sequence of local optimal decisions about which features/system parameters can be used to partition dataset $\bar{X}$. The optimal or rightly sized DT T_{k0} is built in accordance with the subsequent optimization problem

$$\hat{R}(T_{k0}) = \min_{k} \{\hat{R}(T_k)\}, \quad k = 1, 2, 3, \ldots, K \tag{10}$$

$$\hat{R}(T) = \sum_{t \in \hat{T}} \{r(t)p(t)\} \tag{11}$$

where $\hat{R}(T)$ is the error level in the misclassification of tree T_k, T_{k0} is the optimal DT for minimizing the error of misclassification, T is a binary tree $\in \{T_1, T_2, T_3, \ldots, T_k, t_1\}$, k is the index number of the tree, t is a tree node, with t_1 the root node, $r(t)$ is the re-substitution estimation of error in misclassifying in node t, and $p(t)$ is the probability that any case drops into node t.

T^L and T^R represent the sub-trees given on left/right partition sets. T is formed by features plane partitioning. Lattice L could be binary partitioned into conjointly exclusive left/right sets, as presented in Fig. 1(a) for 2-D binary classification boundaries presented in Fig. 1(b). The left set comprises lattice components having feature q values less than the threshold limit, whereas the right set comprises lattice components having feature q values surpassing the threshold limit.

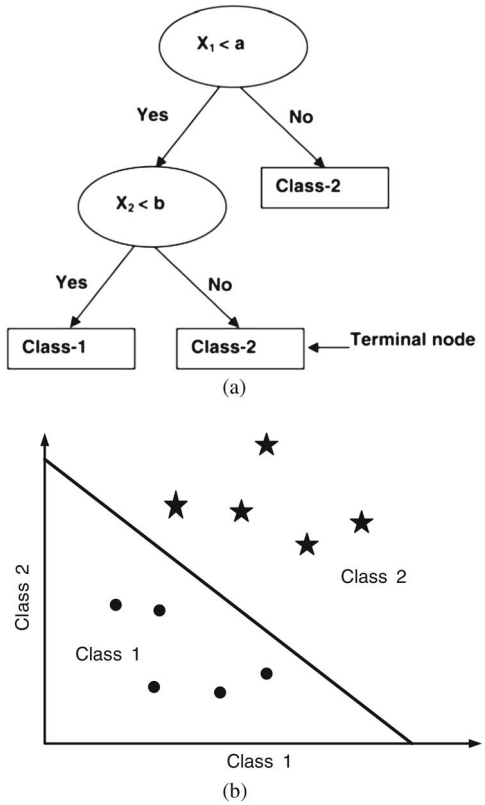


Fig. 1. (a) Threshold-based classification and (b) boundary-based classification.

C. Support Vector Machines

SVMs represent a statistical learning technique for the classification of patterns based on structural risk minimization method, being a suitable candidate due to its capacity to generalize high-dimensional feature spaces.

SVMs have provided better performance than other classical techniques, such as ANN and Bays classifier [25], being suitable for PQ disturbances classification [26]. It gives improved generalization capabilities comparatively to ANN because its training is supported on a sequential minimization technique.

For n-dimension inputs $s_i (i = 1, 2, \ldots, M)$, M is the sample number fitting to class 1 or class 2 with outputs $o_i = 1$ for class 1 and $o_i = -1$ for class 2, correspondingly.

The hyperplane is given as

$$f(s) = w^T s + b = \sum_{j=1}^{n} w_j s_j + b = 0 \tag{12}$$

where w is an n-dimension vector and b is a parameter.

The separating hyperplane position is defined by the w and b values (Fig. 2).

The constraints are $f(s_i) \geq 1$ if $o_i = 1$ and $f(s_i) \geq -1$ if $o_i = -1$, thus

$$o_i f(s_i) = o_i(w^T s + b) \geq +1, \quad \text{for } i = 1, 2, \ldots, M. \tag{13}$$

The geometrical distance is given as $\|w\|^{-2}$. The optimal hyperplane can be driven by the following optimization problem [25]:

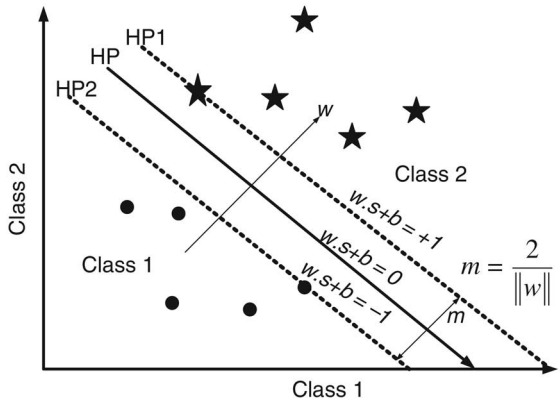


Fig. 2. Hyperplane of SVM for classification.

$$\text{Minimize } \frac{1}{2}\|w\|^{-2} + C\sum_{i=1}^{M} \xi_i \tag{14}$$

subject to

$$o_i(w^T s + b) \geq 1 - \xi_i, \quad \text{for } i = 1, 2, \ldots, M \tag{15}$$

$$\xi_i \geq 0 \quad \text{for all } i. \tag{16}$$

The optimal bias value b^* is given as

$$b^* = -\frac{1}{2}\sum_{\text{SVs}} o_i \alpha_i^* (v_1^T s_i + v_2^T s_i) \tag{17}$$

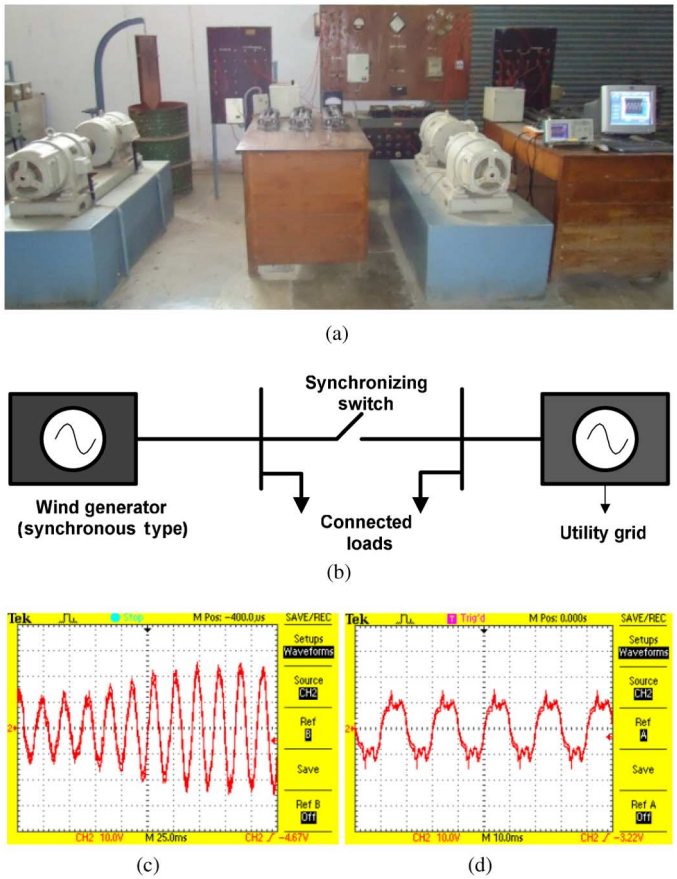


Fig. 3. (a) Experimental prototype for a wind energy system, (b) equivalent diagram, (c) swell due to increase in wind speed, and (d) notch with harmonics.

where v_1 and v_2 are random SVMs for class 1 and class 2, respectively.

The decision function corresponds to

$$f(s) = \sum_{\text{SVs}} \alpha_i o_i s_i^T s + b^*. \tag{18}$$

Unidentified data sample s is categorized as

$$s \in \left\{ \begin{array}{ll} \text{Class-1}, & f(s) \geq 0 \\ \text{Class-2}, & \text{otherwise} \end{array} \right\}. \tag{19}$$

The PQ disturbances classification is accomplished by applying a kernel function [27], such as Gaussian radial basis kernel function [28].

III. Effect of Environmental Factors on Voltage

The study is performed for three different test systems:

1) experimental setup prototype for a wind energy system;
2) experimental setup prototype for a PV system;
3) modified Nordic 32-bus test system with wind/PV penetration.

Environmental factors, such as solar/wind variations, significantly affect voltage at the point of common coupling (PCC), inducing sag, swell, and notch. The parameters of wind and PV system are considered from [11] and [27].

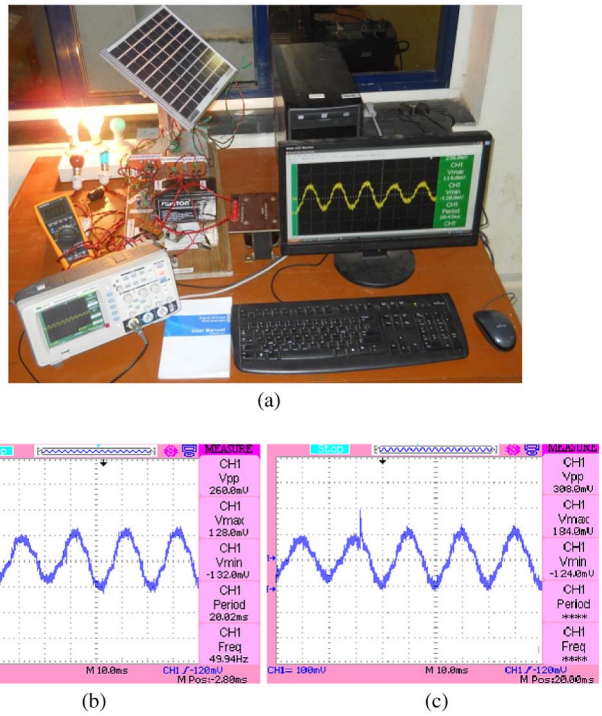


Fig. 4. (a) Experimental prototype for a PV system, (b) swell due to increase in solar irradiance, and (c) notch with swell.

A. Wind Energy System

A prototype setup of a wind energy system was experimentally developed in the laboratory [Fig. 3(a)] whose equivalent diagram is shown in Fig. 3(b). The experimental setup consists of two synchronous alternators driven by a dc motor as the prime mover. One of the generators acts as wind generator (WG), while the second one is considered equivalent to the grid. The variation in the speed of the prime mover is assumed to be equivalent to the wind speed variation input to the wind turbine generator. The swell in the voltage profile due to the increase in dc motor speed/wind speed can be seen in Fig. 3(c), whereas notch with harmonics can be seen in Fig. 3(d). The description of the disturbances (C1–C4) is presented in Table I PV system.

A prototype setup of a PV system was experimentally developed in the laboratory [Fig. 4(a)]. Based on the climatic conditions and sliding window arrangements in the laboratory, the solar irradiance is modified to generate PQ disturbances. The swell associated with the increase in solar irradiance can be seen in Fig. 4(b), whereas notch with swell due to sudden load change and solar irradiance can be seen in Fig. 4(c). The description of disturbances (C5–C7) is presented in Table I.

B. Modified Nordic 32-Bus Test System

Fig. 5 shows the modified Nordic 32-bus test system [28] considered in this study, which presents the effect of both PV and wind power penetration on PQ. The nominal power capacity is about 19 000 MW. The test system is changed by adding 600 MW from WGs and 400 MW from the PV system. PQ disturbances are created in this test system under varying solar irradiance/wind speed, and load up to $\pm 10\%$ of nominal values. The wind/PV models are taken from [11]. The description of disturbances (C8–C13) is presented in Table I.

TABLE I
SIGNAL EXPLANATION FOR PQ DISTURBANCES

Signal class	Signal description
C1	Sag due to increase in load in wind energy system
C2	Sag due to decrease in wind speed and increase in load in wind energy system
C3	Swell due to increase in wind speed and decrease in load in wind energy system
C4	Notch with harmonics due to sudden variation in load in wind energy system
C5	Sag due to decrease in solar irradiance in PV system
C6	Swell due to increase in solar irradiance in PV system
C7	Notch with swell due to sudden change in load and solar irradiance in PV system
C8	Sag due to decrease in solar irradiance in modified Nordic 32-bus test system
C9	Swell due to increase in solar irradiance in modified Nordic 32-bus test system
C10	Notch due to sudden increase in load in modified Nordic 32-bus test system
C11	Sag with harmonics due to decrease in wind speed and nonlinear load in modified Nordic 32-bus test system
C12	Swell with harmonics due to increase in wind speed and nonlinear load in modified Nordic 32-bus test system
C13	Notch with harmonics due to sudden increase in nonlinear load in modified Nordic 32-bus test system

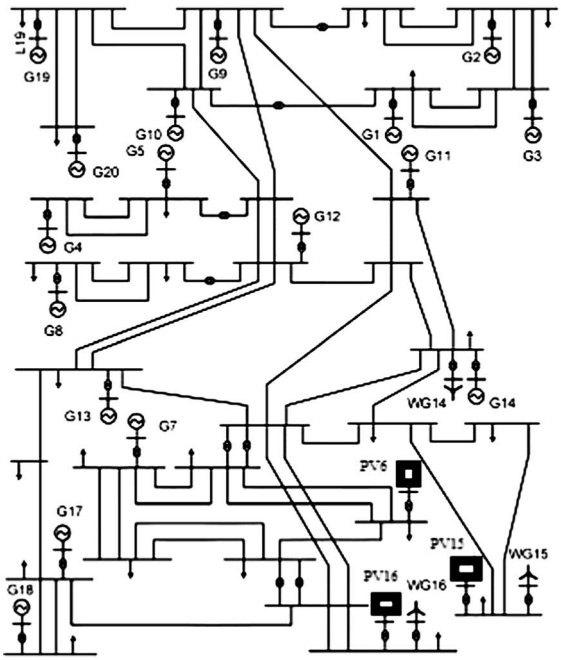


Fig. 5. Modified Nordic 32-bus test system [29].

IV. OPTIMAL FEATURE SELECTION BY GA

PQ disturbance signals, described in Table I, are processed through HS-transform, providing the HST matrix. The sampling frequency of retrieved signals at PCC is considered to be 3.2 kHz. Ten statistical features are obtained from the HST matrix, concerning energy of magnitude/phase contour, STD of magnitude/phase contour, mean of magnitude/phase contour, skewness of magnitude/phase contour, and kurtosis of magnitude/phase contour.

Optimal feature selection is accomplished by GA (Fig. 6). GA is a stochastic search algorithm based on the natural evolution process in which the fittest individual is considered to be winner in the competition. It begins with a random initial population in which the fittest individuals have a greater probability to be chosen to produce children for the following generation. The chosen individuals are enhanced by crossover/mutation. The algorithm is repeated till a converging criterion is met [30].

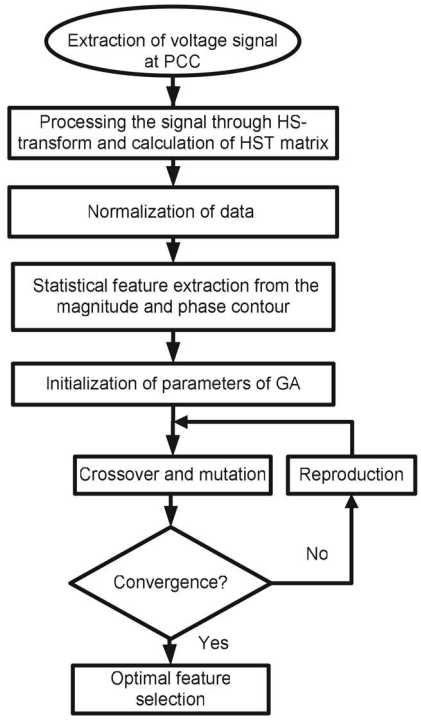


Fig. 6. Flowchart for the optimal feature selection by GA.

As illustrated in Fig. 6, the voltage signal is retrieved at PCC and is processed through HS-transform to extract the statistical features. The total number of features extracted for the 13 different PQ disturbances corresponds to 130 feature vectors. GA is used to select the most appropriate features, leading to improved classification accuracy. The GA encoding scheme of chromosome is composed of a binary string, representing the choice of one of the features from the existing 10 features in the 16 nodes. The decoding of the chromosomes is based on a discrete-type decoding technique. The mutation operator varies from 1 to 0 and vice-versa. Crossover probability is considered between 0.5 and 1, while mutation probability is considered between 0.1 and 1.

Usually, the following three steps are considered for selecting the optimal feature set [31], [32]:

Step 1) *Chromosome design:* The chromosomes are designed based on the conversion of bit strings from genotype to phenotype using the following equation [31]:

$$p = m_x + \frac{m_x - m_y}{2^l - 1} \times z \tag{20}$$

where p is the phenotype of bit string, m_x is the maximum value of the parameter, m_y is the minimum value of the parameter, z is the decimal value of bit string, and l is the length of bit string. For the chromosome representing the feature mask, the bit

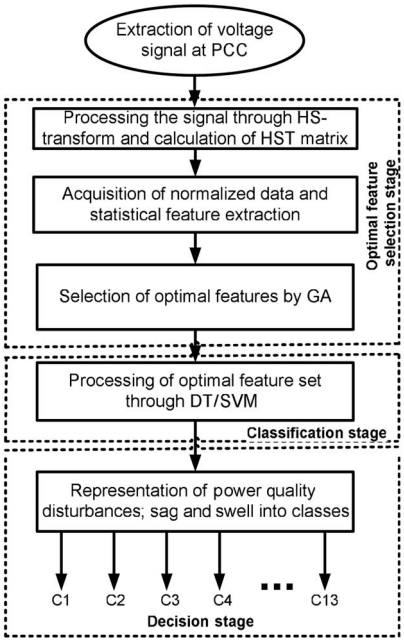


Fig. 7. Flowchart for the classification strategy using DT/SVM.

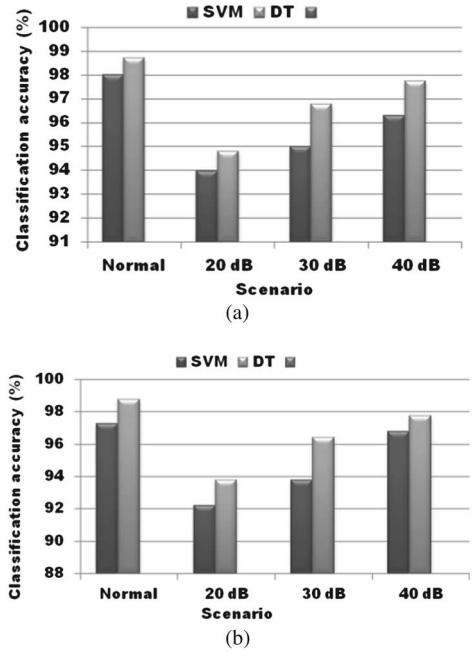


Fig. 8. Total classification accuracy for the PV system: (a) between sag and swell and (b) between swell and notch, in voltage signal.

TABLE II
CLASSIFICATION OF DISTURBANCES C1–C4 FOR THE WIND ENERGY SYSTEM
(A) NO-NOISE SIGNAL

	SVM				DT			
	C1	C2	C3	C4	C1	C2	C3	C4
C1	97	1	2	0	98	0	1	1
C2	2	96	1	1	2	98	0	0
C3	1	1	97	1	0	1	99	0
C4	1	0	2	97	1	1	0	98
	Total accuracy = 96.75				Total accuracy = 98.25			

(B) 20-DB NOISE SIGNAL

	SVM				DT			
	C1	C2	C3	C4	C1	C2	C3	C4
C1	94	2	2	2	94	4	1	1
C2	3	93	3	1	4	93	2	1
C3	4	1	93	2	1	1	96	2
C4	3	2	1	94	0	3	2	95
	Total accuracy = 93.4				Total accuracy = 94.5			

"1" indicates that the feature is selected, whereas "0" indicates that the feature is not selected.

Step 2) *Fitness function:* High classification accuracy, minimum number of selected features, and feature cost are the three important criteria used to design a fitness function based on

$$\text{fitness} = w_a \times \text{DT\_accuracy} + w_f \times \left(\sum_{i=1}^{n} c_i \times f_i \right) \tag{21}$$

where w_a is the DT classification accuracy weight, DT_accuracy is the classification accuracy of DT, w_f is the weight for the number of features, and c_i is the cost of feature i, where "1" indicates that feature i is selected, and "0" indicates that feature i is not selected.

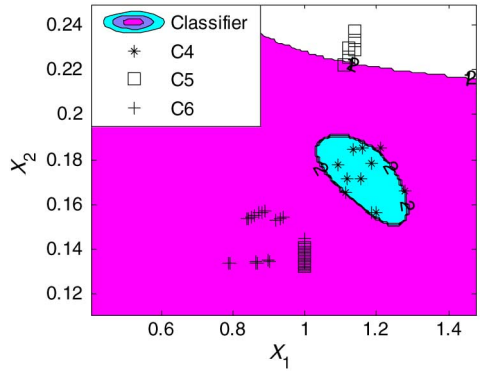


Fig. 9. Boundary plot to assess the classification performance of SVM.

Step 3) *Architectures for GA-based optimal feature:* This step consists of data processing, fitness evaluation, and genetic operation and convergence [32].

Fig. 7 shows the various steps for the classification of PQ disturbances.

V. CLASSIFICATION USING DT AND SVM

The classification capabilities of DT and SVM in the three test systems, under different operating scenarios, are detailed as follows.

A. Wind Energy System

The HS matrix from which the features are extracted is a matrix in which the rows denote frequency, whereas the columns denote time instant. In this study, 200 diverse tests are performed in the experimental prototype for four PQ disturbances. The final feature matrix size is 200 × 10, since 10 optimal features are

TABLE III
CLASSIFICATION OF DISTURBANCES C1–C7
(A) NO-NOISE SIGNAL

	SVM							DT						
	C1	C2	C3	C4	C5	C6	C7	C1	C2	C3	C4	C5	C6	C7
C1	95	1	1	1	2	0	0	96	1	2	1	0	0	0
C2	2	94	1	1	1	0	1	0	98	0	1	0	0	1
C 3	2	1	94	1	1	1	0	1	0	98	0	0	0	1
C4	1	1	2	95	1	0	0	1	1	0	97	1	0	0
C5	1	0	0	0	97	0	2	0	0	0		99	0	1
C6	0	1	0	1	1	95	2	0	1	0	0	0	99	0
C7	2	0	1	0	0	1	96	0	0	1	0	0	0	99
	Total accuracy = 95.15							Total accuracy = 98.0						

(B) 20-DB NOISE SIGNAL

	SVM							DT						
	C1	C2	C3	C4	C5	C6	C7	C1	C2	C3	C4	C5	C6	C7
C1	93	0	0	1	2	2	2	94	2	1	2	1	0	0
C2	3	91	1	4	1	0	0	2	93	2	1	0	0	2
C 3	4	3	89	1	0	2	1	0	0	94	1	2	2	1
C4	1	2	1	90	2	1	3	0	0	2	94	0	3	1
C5	1	2	2	1	91	2	1	0	1	2	0	95	0	2
C6	4	1	2	1	0	90	2	1	0	1	1	2	95	0
C7	1	3	1	2	1	0	92	1	2	0	1	0	0	96
	Total accuracy = 90.86							Total accuracy = 94.43						

selected using GA for each case. The classification accuracy for disturbances C1–C4 is provided in Table II. Elements in the diagonal denote properly classified percentage, whereas elements off the diagonal denote misclassification percentage. Total classification accuracy using DT and SVM is shown in the last row of the table.

The results reflect a better accuracy of DT compared to SVM in classifying the disturbances under different operating conditions.

B. PV System

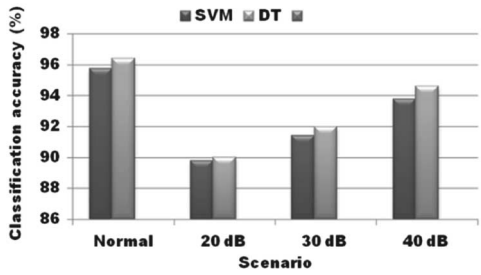
A similar procedure is adopted in this case, i.e., features are extracted by HS-transform. Out of the total features, only the optimal features are selected using GA, which are given as input to DT and SVM to classify the disturbances under no-noise, 40-, 30-, and 20-dB noise scenarios.

The results showing the classification accuracy between sag and swell disturbances are presented in Fig. 8(a), and the results showing the classification accuracy between swell and notch are presented in Fig. 8(b). The results show some degradation of classification accuracy with the increase of noise level from 40 to 20 dB, but still always above 92%. The classification performance of SVM is also tested using a boundary plot, presented in Fig. 9, clearly discriminating the disturbances based on the decision boundaries called hyperplanes.

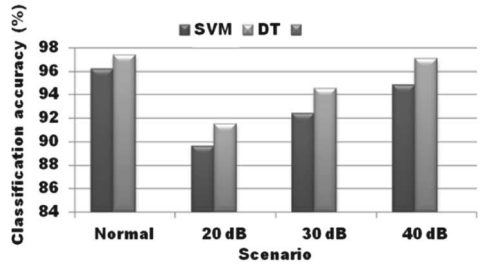
Next, a classification study is presented under no-noise and 20-dB noise using DT and SVM to check their ability to classify disturbances C1–C7. The classification results are shown in Table III. Both DT and SVM perform well without significant deterioration, but DT shows again better accuracy than SVM. This reflects the robustness of HS-transform combined with DT.

C. Modified Nordic 32-Bus Test System

The classification ability is also studied on the modified Nordic 32-bus test system, shown previously in Fig. 5.



(a)



(b)

Fig. 10. Total classification accuracy for the modified Nordic 32-bus test system: (a) between sag and swell and (b) between swell and notch, in voltage signal.

The different disturbances C8–C13, as mentioned in Table I, are classified using DT and SVM. The classification accuracy results are shown in Fig. 10(a) and (b) between sag and swell, swell and notch, respectively. The accuracy of DT as classifier is observed to be comparatively better than SVM. The results considering different combinations of features under different noise levels are shown in Table IV.

Classification accuracy augments as the features number increases. Moreover, as the noise level increases, the accuracy diminishes, but still always above 93% for SVM and 95% for DT, respectively. The classification accuracy with different hybrid techniques under different noise levels is shown in Table V. The

TABLE IV
CLASSIFICATION RESULTS WITH DIFFERENT COMBINATION OF DISTURBANCES

No. of features	Optimal/non-optimal features (Combination of features)	Classification accuracy by SVM					Classification accuracy by DT				
		No-noise	40 dB	30 dB	20 dB	10 dB	No-noise	40 dB	30 dB	20 dB	10 dB
1	F1	93.9	93.0	90.4	88.3	87.2	93.0	91.0	90.8	87.5	86.9
3	F2-F4-F6	95.0	93.3	92.0	91.3	89.2	95.0	94.6	94.0	93.5	91.1
6	F1-F3-F5-F7-F9-F11	95.8	94.3	94.0	91.6	90.1	97.0	96.8	95.4	93.7	92.8
9	F1-F2-F3-F4-F5-F6-F7-F8-F9	96.3	95.1	94.0	92.1	91.7	98.3	97.1	96.0	93.5	93.0
11	F1-F2-F3-F4-F5-F6-F7-F8-F9-F10-F11	97.0	96.3	95.0	93.7	92.0	99.0	98.2	97.8	95.7	94.3
13	F1-F2-F3-F4-F5-F6-F7-F8-F9-F10-F11-F12-F13	98.8	97.6	96.2	95.0	93.2	99.6	99.0	98.6	96.8	95.0

TABLE V
CLASSIFICATION ACCURACY WITH DIFFERENT HYBRID TECHNIQUES

Classification technique	No-noise		40 dB noise		30 dB noise		20 dB noise	
	Non-optimal feature	Optimal feature by GA	Non-optimal feature	Optimal feature by GA	Non-optimal feature	Optimal feature by GA	Non-optimal feature	Optimal feature by GA
WT with SVM	96.0	97.0	95.2	95.9	93.5	95.2	91.7	93.6
WT with DT	96.8	97.1	95.7	96.5	94.2	96.2	92.9	94.9
S-transform with SVM	97.2	97.9	96.0	95.1	92.9	94.3	90.5	93.1
S-transform with DT	97.4	98.3	96.7	97.1	96.0	96.9	94.2	95.2
HS-transform with SVM	97.7	98.8	96.4	97.3	96.0	96.9	93.4	95.8
HS-transform with DT	98.7	99.5	97.9	98.7	97.0	97.8	95.9	96.1

TABLE VI
CLASSIFICATION ACCURACY WITH DIFFERENT PERCENTAGES OF DG PENETRATION

Classification technique	5% DG penetration		10% DG penetration		15% DG penetration		20% DG penetration	
	Non-optimal feature	Optimal feature by GA	Non-optimal feature	Optimal feature by GA	Non-optimal feature	Optimal feature by GA	Non-optimal feature	Optimal feature by GA
WT with SVM	96.4	97.2	96.2	97.0	95.8	96.4	95.0	96.0
WT with DT	97.0	97.4	96.4	97.2	96.0	97.3	95.7	96.4
S-transform with SVM	97.5	97.8	96.5	97.6	96.4	97.6	96.2	96.8
S-transform with DT	97.4	98.4	97.0	98.2	96.7	97.8	96.4	97.6
HS-transform with SVM	98.0	98.6	97.8	98.4	97.4	98.4	96.8	98.1
HS-transform with DT	98.8	99.8	98.0	99.4	97.7	99.0	96.8	98.8

best classification accuracy corresponds to HS-transform combined with DT, reaching a maximum of 99.5% with optimal feature selection by GA under the no-noise scenario.

Further, the accuracy of the proposed approach is tested for different percentages of DG penetration in the 32-bus test system, as presented in Table VI. It is concluded from this tabular result that with the increase in penetration level, the accuracy of the proposed approach is decreased by only a very small percentage, which is an important additional feature.

VI. CONCLUSION

This work presented a study on the detection and classification of PQ disturbances obtained in three different test systems (experimental prototypes for wind/PV systems and modified Nordic 32-bus test system with wind/PV penetration) using WT/ST/HS-transform and DT/SVM. The PQ disturbances are created not only due to load change, but also due to changes in environmental factors like wind speed and solar irradiance. The disturbances were detected using WT, ST, and HS-transform, of which HS-transform showed improved performance. Also, the classification of PQ disturbances under different operating conditions was tested by extracting features using WT/ST/HS-transform followed by the selection of optimal features by GA. The optimal features were fed to DT/SVM to classify the disturbances. Using optimal features from GA, along with the HS-transform, significantly improves the classification strategy. A thorough comparative assessment in terms of classification accuracy also leads to conclude that the proficiency of DT is better than SVM. The best classification accuracy was obtained with HS-transform combined with DT and optimal feature selection by GA, reaching 99.5% on the modified Nordic 32-bus test system under no-noise, and 96.1% with a 20-dB noise scenario.

REFERENCES

[1] G. Pepermansa, J. Driesenb, D. Haeseldonckxc, R. Belmansc, and W. D'haeseleer, "Distributed generation: Definition, benefits and issues," *Energy Policy*, vol. 33, no. 6, pp. 787–798, Apr. 2005.

[2] V. H. Mendez, J. Rivier, J. I. de la Fuente, T. Gomez, J. Arceluz, J. Marin, and A. Madurga, "Impact of distributed generation on distribution investment deferral," *Int. J. Elect. Power Energy Syst.*, vol. 28, no. 4, pp. 244–252, May 2006.

[3] S.-K. Kim, J.-H. Jeon, C.-H. Cho, J.-B. Ahn, and S.-H. Kwon, "Dynamic modeling and control of a grid-connected hybrid generation system with versatile power transfer," *IEEE Trans. Ind. Electron.*, vol. 55, no. 4, pp. 1677–1688, Apr. 2008.

[4] P. Ray, S. Mohanty, and N. Kishor, "Proportional integral controller based small-signal analysis of hybrid distributed generation systems," *Energy Convers. Manage.*, vol. 52, no. 4, pp. 1943–1954, Apr. 2001.

[5] G. Chicco, J. Schlabbach, and F. Spertino, "Experimental assessment of the waveform distortion in grid connected photovoltaic installations," *Solar Energy*, vol. 83, no. 7, pp. 1026–1039, Jul. 2009.
[6] J. Bialasiewicz and E. Muljadi, "The wind farm aggregation impact on power quality," in *Proc. IEEE Annual Conf. Ind. Electron.*, Paris, France, 2006, pp. 4195–4200.
[7] P. Caramia, G. Carpinelli, and P. Verde, *Power Quality Indices in Liberalized Markets*, Chichester, U.K.: Wiley, 2009.
[8] D. Granados-Lieberman, R. J. Romero-Troncoso, E. Cabal-Yepez, R. O. Osornio-Rios, and L. A. Franco-Gasca, "A real-time smart sensor for high-resolution frequency estimation in power systems," *Sensors*, vol. 9, no. 9, pp. 7412–7429, Sep. 2009.
[9] H. Siahkali, "Power quality indexes for continue and discrete disturbances in a distribution area," in *Proc. IEEE Power and Energy Conf.*, Johor Bahru, Malaysia, pp. 678–683, 2008.
[10] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," *IEEE Trans. Power Del.*, vol. 11, no. 2, pp. 924–930, Apr. 1996.
[11] P. Ray, N. Kishor, and S. Mohanty, "Islanding and power quality disturbance detection in grid connected hybrid power system using wavelet and S-transform," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1082–1094, Sep. 2012.
[12] P. Ray, S. Mohanty, and N. Kishor, "Disturbance detection in grid connected distributed generation system using wavelet and S-transform," *Elect. Power Syst. Res.*, vol. 81, no. 3, pp. 805–819, Mar. 2011.
[13] S. Santoso, W. M. Grady, E. J. Powers, J. Lamoree, and S. C. Bhatt, "Characterization of distribution power quality events with Fourier and wavelet transforms," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 247–254, Jan. 2000.
[14] T. Lin and A. Domijan, "Recursive algorithm for real-time measurement of electrical variables in power systems," *IEEE Trans. Power Del.*, vol. 21, no. 1, pp. 15–22, Jan. 2006.
[15] P. K. Dash and M. V. Chilukuri, "Hybrid S-transform and Kalman filtering approach for detection and measurement of short duration disturbances in power networks," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 2, pp. 588–594, Apr. 2004.
[16] A. K. Ghosh and D. L. Lubkeman, "The classification of power system disturbance waveforms using a neural network approach," *IEEE Trans. Power Del.*, vol. 10, no. 1, pp. 109–115, Jan. 1995.
[17] M. Reaz, F. Choong, M. Sulaiman, F. Mohd Yasin, and M. Kamada, "Expert system for power quality disturbance classifier," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1979–1988, Jul. 2007.
[18] A. Elmitwally, S. Farghal, M. Kandil, S. Abdelkader, and M. Elkateb, "Proposed wavelet-neuro-fuzzy combined system for power quality violations detection and diagnosis," *IEE Proc. Gener. Transm. Distrib.*, vol. 148, no. 1, pp. 15–20, Jan. 2001.
[19] A. Ghosh and D. Lubkeman, "The classification of power system disturbance waveforms using a neural network approach," *IEEE Trans. Power Del.*, vol. 10, no. 1, pp. 109–115, Jan. 1995.
[20] W. M. Lin, C. H. Wu, C. H. Lin, and F. S. Cheng, "Detection and classification of multiple power-quality disturbances with wavelet multiclass SVM," *IEEE Trans. Power Del.*, vol. 23, no. 4, pp. 2575–2582, Oct. 2008.
[21] S. Mishra, C. Bhende, and B. Panigrahi, "Detection and classification of power quality disturbances using S-transform and probabilistic neural network," *IEEE Trans. Power Del.*, vol. 23, no. 1, pp. 280–287, Jan. 2008.
[22] B. Biswal, P. Dash, and B. Panigrahi, "Power quality disturbance classification using fuzzy C-Means algorithm and adaptive particle swarm optimization," *IEEE Trans. Ind. Electron.*, vol. 56, no. 1, pp. 210–220, Jan. 2009.
[23] J. R. Quinlan, "Induction of decision trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, Mar. 1986.
[24] F. Esposito, D. Malerba, and G. A. Semeraro, "A comparative analysis of methods for pruning decision trees," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 5, pp. 476–491, May 1997.
[25] S. Ekici, "Classification of power system disturbances using support vector machines," *Expert Syst. Appl.*, vol. 36, no. 6, pp. 9859–9868, Aug. 2009.
[26] R. Salat and S. Osowski, "Accurate fault location in the power transmission line using support vector machine approach," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 879–886, May 2004.
[27] P. Ray, S. Mohanty, and N. Kishor, "Classification of power quality disturbances due to environmental characteristics in distributed generation system," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 302–313, Apr. 2013.
[28] M. Glavić and T. Van Cutsem, "Wide-area detection of voltage instability from synchronized phasor measurements. Part II: Simulation results," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1417–1425, Aug. 2009.
[29] A. Zertek, G. Verbic, and M. Pantos, "A novel strategy for variable-speed wind turbines'participation in primary frequency control," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 791–799, Oct. 2012.
[30] D. C. Das, A. K. Roy, and N. Sinha, "GA based frequency controller for solar thermal–diesel–wind hybrid energy generation/energy storage system," *Int. J. Elect. Power Energy Syst.*, vol. 43, no. 1, pp. 262–279, Dec. 2012.
[31] C. L. Huang and C. J. Wang, "A GA-based feature selection and parameters optimization for support vector machines," *Expert Syst. Appl.*, vol. 31, no. 31, pp. 231–240, Aug. 2006.
[32] B. K. Panigrahi and V. R. Pandi, "Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm," *IET Gener. Transm. Distrib.*, vol. 3, no. 3, pp. 296–306, Mar. 2009.

Prakash K. Ray (S'10) received the Ph.D. degree from Motilal Nehru National Institute of Technology (MNNIT), Allahabad, India.

Currently, he is working as Assistant Professor in the Department of Electrical and Electronics Engineering, International Institute of Information Technology (IIIT), Bhubaneswar, India. His research area includes distributed generations, renewable energy resources and digital signal processing, and soft computing applications in power system.

Soumya R. Mohanty (M'08–SM'13) received the Ph.D. degree from Indian Institute of Technology (IIT), Kharagpur, India.

Currently, he is an Assistant Professor in the Department of Electrical Engineering, Motilal Nehru National Institute of Technology (MNNIT), Allahabad, India. His research area includes digital signal processing applications in power system relaying and power quality, pattern recognition applications to distributed generation-based system.

Nand Kishor (SM'13) received the Ph.D. degree from Indian Institute of Technology (IIT), Roorkee, India.

Currently, he is an Associate Professor with the Department of Electrical Engineering, Motilal Nehru National Institute of Technology (MNNIT), Allahabad, India. Since August 2012, he has been working as a Marie Curie Experienced Researcher (Marie Curie Fellow) at the Electrical Engineering Department, Aalto University, Finland. His research area includes artificial intelligence (AI) applications in power systems, distributed generations, wireless sensor networks, and digital signal processing applications in power system.

João P. S. Catalão (M'04–SM'12) received the M.Sc. degree from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2003, and the Ph.D. degree from the University of Beira Interior (UBI), Covilha, Portugal, in 2007.

He is currently a Professor at UBI, Director of the M.Sc. in Electromechanical Engineering (EUR-ACE label) and Researcher at INESC-ID. He is the Primary Coordinator of the EU-funded FP7 project SiNGULAR. He is an author or coauthor of more than 200 papers published in journals, book chapters, and conference proceedings, having supervised more than 20 post-docs, Ph.D., and M.Sc. students. His research interests include power system operations and planning, distributed renewable generation, demand response, and smart grids.

Prof. Catalão is an Editor of IEEE TRANSACTIONS ON SMART GRID, an Editor of IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, and an Associate Editor of *IET Renewable Power Generation*. Also, he is the Editor of the book titled *Electric Power Systems: Advanced Forecasting Techniques and Optimal Generation Scheduling* (Boca Raton, FL, USA: CRC Press, 2012).