European Journal of Scientific Research
ISSN 1450-216X Vol.34 No.2 (2009), pp.237-251

http://www.eurojournals.com/ejsr.htm

Support Vector Regression Based S-transform for Prediction of Single and Multiple Power Quality Disturbances

M F Faisal
Department of Electrical, Electronic and Systems Engineering
Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia
E-mail: mfuadfa@tnb.com my
Tel: +603-89216590; Fax: +603-89216146

A Mohamed
Department of Electrical, Electronic and Systems Engineering
Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia
E-mail: azah@eng.ukm.my
Tel: +603-89216590; Fax: +603-89216146

A Hussain
Department of Electrical, Electronic and Systems Engineering
Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia

M Nizam
Department of Electrical, Electronic and Systems Engineering
Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia

Abstract

This paper presents a novel approach using Support Vector Regression (SVR) based S-transform to predict the classes of single and multiple power quality disturbances in a three-phase industrial power system. Most of the power quality disturbances recorded in an industrial power system are non-stationary and comprise of multiple power quality disturbances that coexist together for only a short duration in time due to the contribution of the network impedances and types of customers' connected loads. The ability to detect and predict all the types of power quality disturbances encrypted in a voltage signal is vital in the analyses on the causes of the power quality disturbances and in the identification of incipient fault in the networks. In this paper, the performances of two types of SVR based S-transform, the non-linear radial basis function (RBF) SVR based S-transform and the multilayer perceptron (MLP) SVR based S-transform, were compared for their abilities in making prediction for the classes of single and multiple power quality disturbances. The results for the analyses of 651 numbers of single and multiple voltage disturbances gave prediction accuracies of 86.1% (MLP SVR) and 93.9% (RBF SVR) respectively.

Keywords: Power Quality, Power Quality Prediction, S-transform, SVM, SVR

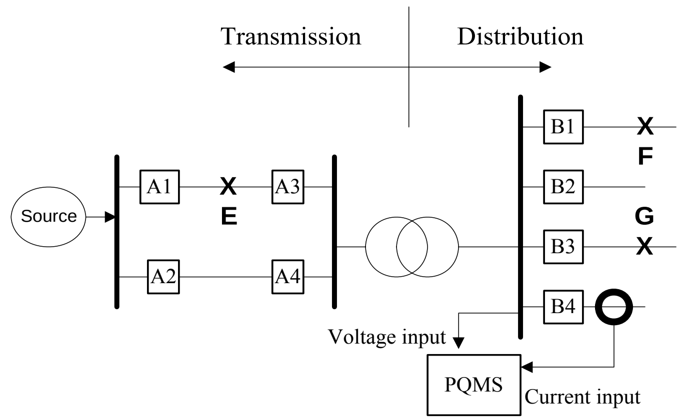
1. Introduction

Over the years, many people involved in power quality have defined and perceived power quality problems differently. The most accepted definition of power quality was made by Dugan et al. (1996), who defined it as “Any electrical power problem manifested in voltage, current, or frequency deviations that results in failure or mis-operation of customers’ operation." In its broadest sense, power quality is a set of boundaries that allows electrical systems to function in their intended manner without significant loss in performance or lifetime. This term is used to describe electric power that drives an electrical load and the load's ability to function properly with that electric power. Without the proper power, an electrical device (or load) may malfunction, fail prematurely or not operate at all. There are many ways in which electric power can be of poor quality and many more causes of poor power quality. The proliferation of electronic equipment in automated manufacturing process industries has brought power quality to the center stage of power supply system planning, design, and utilization.

Variations in the power quality, even for a very short time, that were not a concern before can now be very costly in terms of process shut-downs and electrical equipment malfunctions in manufacturing plants. Modern day customers demand high power quality for improved production output as well as for maintaining an optimal operating cost. As customers seek to increase efficient utilization of resources, power utilities must now strive to better understand power quality problems affecting their customers. There are many events, some man-made and others due to nature, that can cause power quality problems. Analysis of these events is usually difficult because the causes are often unclear. The causes of the disturbances may be related to either a switching operation within the manufacturing facility or a power system fault hundreds of kilometers away.

The increasing use of technologies that are sensitive to minor voltage disturbances has created a need for the implementation of an online power quality monitoring system (PQMS). The PQMS will enable the power utility to perform continuous monitoring of the power systems in order to evaluate the level of the quality of the offered electrical powers, whether they are within pre-specified standards or not and also to obtain the necessary information to detect potential system problems and respond faster to any customer complaints. The general configuration of a PQMS is shown in Figure 1. The PQMS will receive voltage input from the common distribution bus and current input from only one feeder i.e. feeder B4. When a fault occurs either on a transmission or distribution network, it typically draw energy from the power networks. Depending on the relationship between the fault location and monitoring location, either a voltage sag or voltage interruption event will happen. When a fault happens at location E on a transmission line, voltage sag rather than a voltage interruption is observed in the PQMS. When a fault occurs at location F on a parallel distribution circuit, voltage sag will also be observed in the PQMS. Thus, voltage sag will appear when a fault occurs in the power networks. As soon as the fault is cleared by a circuit breaker, the voltage restores to the pre-sag value. The duration of a voltage sag event is associated with the time required for a relay and circuit breaker to detect and clear a fault. The PQMS will record all the behavior of both the voltage and current waveforms during the whole process of fault occurrences and fault clearing.

Figure 1: Architecture of the PQMS



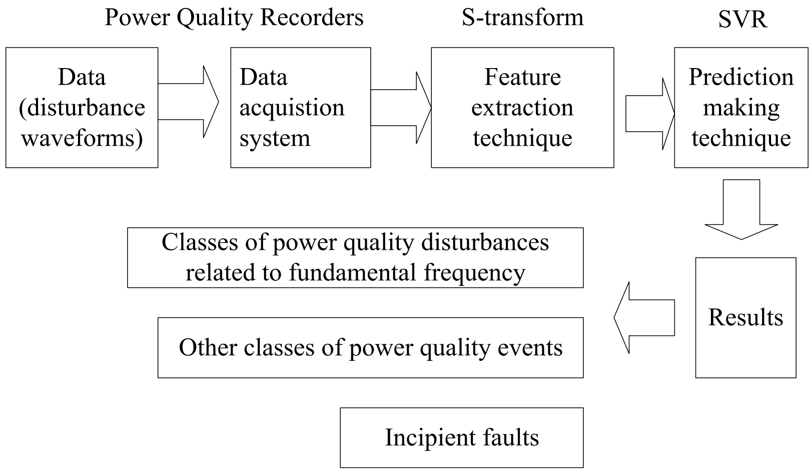
The power quality waveforms i.e. voltages and currents, recorded by the PQMS can provide excellent information for the identification of the types and causes of the power quality disturbances. These signatures can also be used to detect the existence of partial discharges and incipient fault and the needs to conduct conditioned-based maintenance (CBM) activities. The CBM activities, i.e., thermal imaging and ultrasound scanning are necessary to pinpoint the root causes of the partial discharges and incipient faults. Incipient fault occurs when damage or contamination progressively weakens the integrity of the network components over time and leads to insulation failure. These faults are predictable and avoidable if the degradation processes are known by analyzing the disturbance data. In future papers, the author will present the effectiveness of this novel approach in the detection of partial discharges and incipient faults in the networks.

To interpret all the power quality disturbance data will require a high level of engineering expertise. To analyze the causes of the power quality disturbances is also a nontrivial task, mainly due to the huge volume of disturbance records. Therefore, to solve both data volume and lack of expertise problems, a technique that can perform automatic identification of single and multiple power quality disturbances is required.

2. A Novel Approach to Predict Single and Multiple Power Quality Disturbances

In this paper a novel approach to perform automatic prediction of single and multiple power quality disturbances is presented. The block diagram for the new approach shown in Figure 2 was developed using the S-transform and the Support Vector Regression (SVR) techniques. The whole process will start with the recording of power quality disturbance data using power quality recorders in the PQMS. These data will then be processed by the S-transform and features that can characterize the disturbances will be extracted. These features will then be applied to the SVR to predict the classes of the disturbances. The results of this new approach will be categorized into three categories: 1) The first category is the classes of disturbances related to the fundamental components (sags, swells and interruption), 2) the second category is the classes for other disturbances (harmonics, notches, transients etc) and lastly, 3) the third category is called incipient fault. In this study, only the results related to the first and second categories will be presented.

Figure 2: Process flow for the prediction of classes of power quality disturbances



The arrangement of this paper is as follows. In section 3, the description on the application of the S-transform in the detection of power quality disturbances is presented. The theory of the S-transform and selection of features for the detection of the disturbances will be explained in detailed. In section 4, the theory of SVR and its application in predicting the disturbances will be presented. And lastly the results and discussion of the application of the new approach in the prediction of single and multiple power quality disturbances are presented in section 5.

3. Application of S-transform for detecting power quality disturbances

The S-transform is considered to be one of the most recent techniques developed for performing signal processing. It produces a time-frequency representation of a time series signal. The S-transform is also a generalization of the Short-time Fourier transform (STFT), an extension of the continuous wavelet transforms (CWT), and it overcomes some of the disadvantages of the wavelet transforms (Stockwell et al, 1996). The S-transform will perform multiresolution analysis (MRA) on a time varying power signal, as its window width varies inversely with the frequency. The basis function for the S-transform is a Gaussian modulation cosinusoid. The cosinusoid frequencies are used for the interpretation of a signal that will result in the time frequency spectrum. The output of the S-transform is an N x M matrix called the S-matrix whose rows pertain to the frequency and columns to time. Each element of the S-matrix is complex valued and can be used as features to classify the non-stationary single and multiple power quality disturbances. In the latest development in power quality analysis, the S-transform was reported to be the most superior signal processing technique because it is based on dynamic time-frequency dependent resolutions, which allows for the detection of high frequency bursts (Pinnegar and Mansinha, 2003: Pinnegar and Mansinha, 2004). High frequency burst is a common signature for the phase current during incident of partial discharges which will generate incipient fault (Weeks and Steiner, 1982).

The S-transform for a function *h(t)* can be calculated by defining a CWT with a specific mother wavelet function multiplied with a phase factor as shown accordingly,

$$S(\tau, f) = e^{i2\pi f\tau} w(t, f) \tag{1}$$

where the mother wavelet function is defined as

$$w(t,f)=\frac{|f|}{\sqrt{2\pi}}e^{-\frac{t^2 f^2}{2}}e^{-i2\pi ft} \tag{2}$$

Explicitly, the S-transform can be written as

$$S(\tau,f)=\int_{-\infty}^{\infty}h(t)\frac{|f|}{\sqrt{2\pi}}e^{-\frac{(\tau-t)^2 f^2}{2}}e^{-i2\pi ft}dt \tag{3}$$

Equation (3) is further simplified as follows,

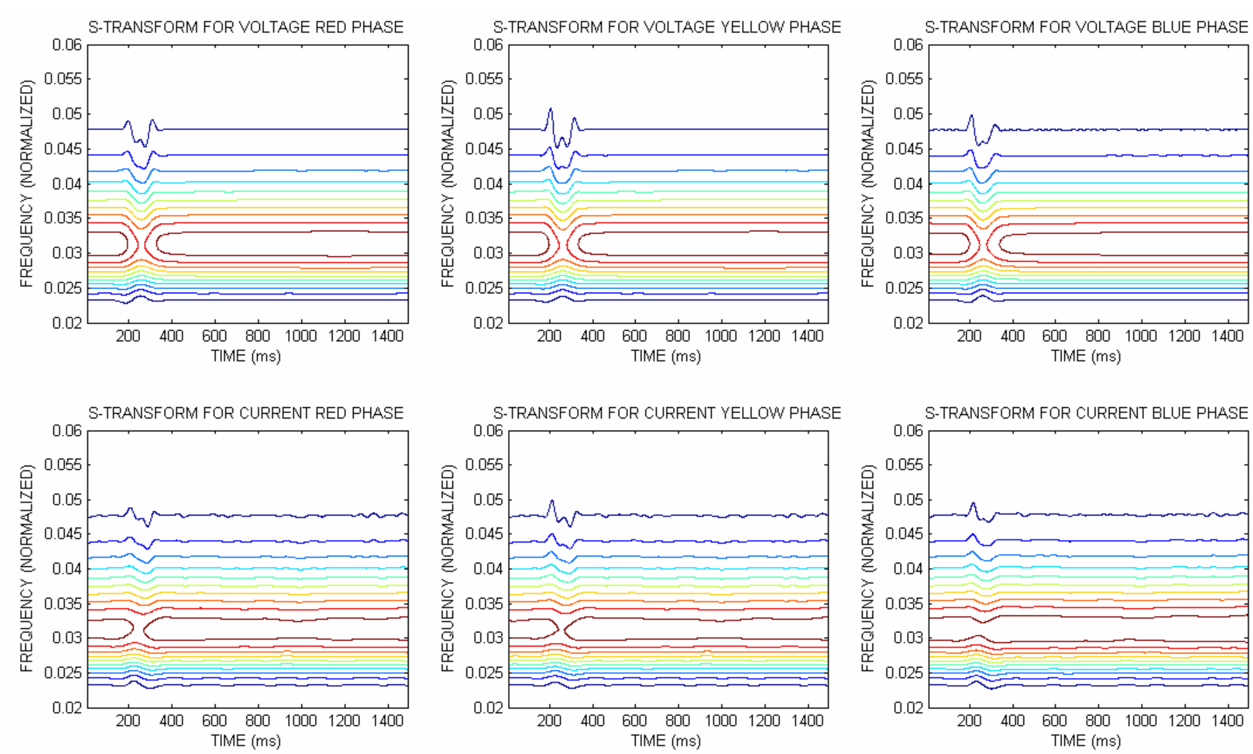
$$S(\tau,f)=\int_{-\infty}^{\infty}h(t)g(\tau-t,f)e^{-2\pi ft} \tag{4}$$

where $g(\tau,f)$ is the Gaussian modulation function which is given by,

$$g(\tau,f)=\frac{|f|}{\sqrt{2\pi}}e^{-\frac{\tau^2 f^2}{2}}e \tag{5}$$

The S-transform will generate time frequency contours, which will display the disturbance patterns for visual identification for the single and multiple power quality disturbances. These contours can provide excellent features, which can be used by a pattern recognition system for classifying the power quality disturbances. Examples on the time frequency contours for voltages and currents for a power quality disturbance are shown in Figure 3. The data in the figure were recorded in one substation in Malaysia. In the figure, the first row showed the time frequency contours for three phase voltage sags and in the second row are the respective time frequency contours for the phase currents. The cause of the three phase voltage sag was due to lightning activities at the transmission networks. The S-transform clearly showed the existence of voltage sags by the sudden changes in the time frequency contours. The resolutions of the contours showed brief reduction during the voltage sag events. The same condition was also reflected for time frequency contours for the currents which also showed brief reduction during the voltage sags events.

Figure 3: Plots of the time frequency contours for voltages and currents



In this study, nine features were extracted from the time frequency contours of the S-transform. The first set of features was based on the maximum values in the S-matrix. In Figure 4, comparison between the original disturbance waveforms (1st row) and the maximum value plots (2nd row) is shown. The maximum value plots for the red, yellow and blue phases showed the existence of voltage sags which coincided with the disturbances seen in the original waveforms. Based on this observation, it was shown that the maximum value plots are very suitable for the detection of classes of disturbances related to the fundamental components (sags, swells and interruption). Four features (F1, F2, F3 and F4) were selected from the maximum value plots. The details of these features are explained in Table 1. In Table 1, the parameters of 0.90 and 1.10 were selected based on the parameters used to define voltage sag and swell as stated in IEEE 1159:1992 standard. Voltage sag is detected when the root mean square (rms) voltage reduce below 90% of the nominal line to neutral voltage for duration between 10 ms to 60 second. And voltage swell is defined when the rms voltage increase above 110% of the nominal line to neutral voltage for the same duration. In this study, the same parameters were used to evaluate the maximum value plots. If the minimum value of the plot is less than 0.90 of the normalized value, then voltage sag is detected. The same methodology is applied for the detection of voltage swell.

Figure 4: Comparison between plots of the (a) original waveforms, (b) maximum values in the S-matrix

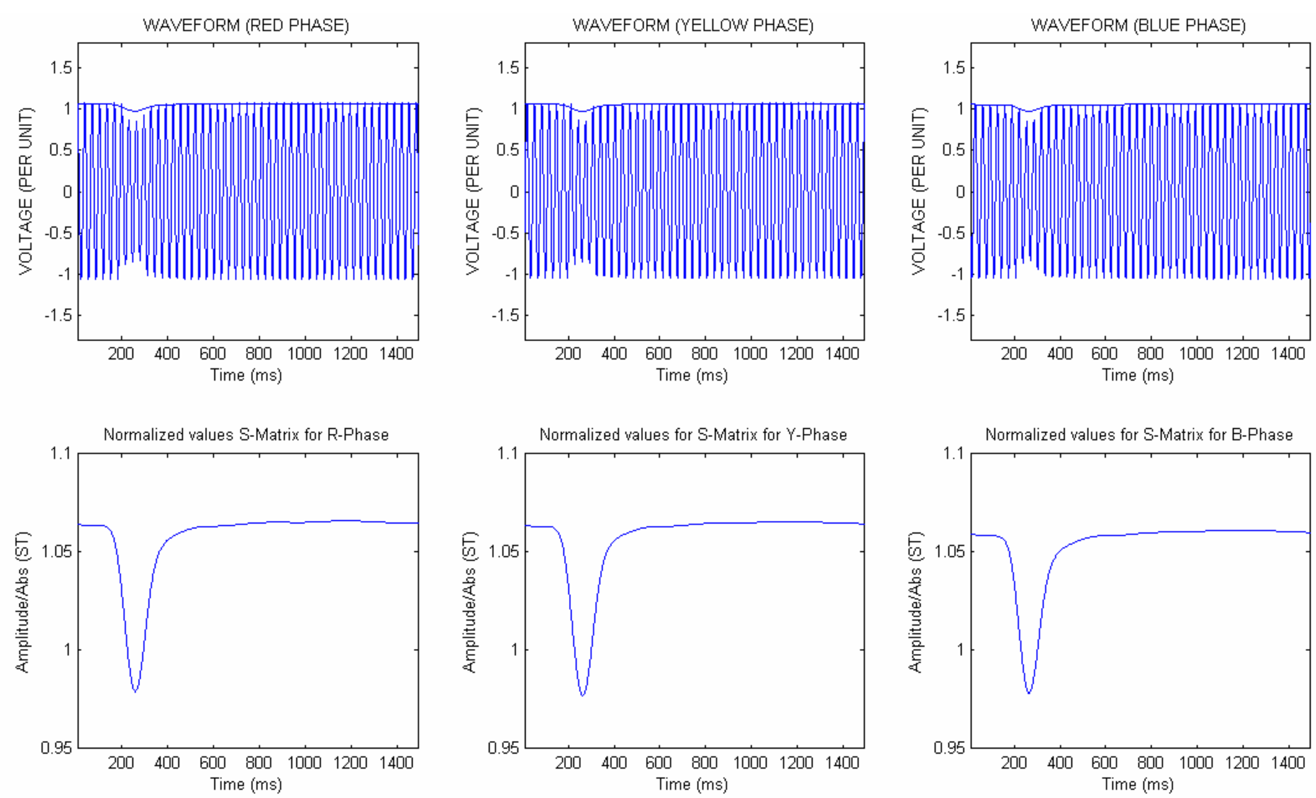


Table 1: Descriptions of features based on the maximum values in the S-Matrix

Features	Description
F1	Values of time resolution (ms) for the data below the absolute value of 0.90 in the maximum value plots.
F2	Values of time resolution (ms) for the data above the absolute value of 1.10 in the maximum value plots.
F3	The minimum value below the absolute value of 0.90 in the maximum value plots.
F4	The maximum value above the absolute value of 1.10 in the maximum value plots.

The second set of features was selected from the values of the frequency resolutions in the S-matrix. In a study on a set of 124 power quality disturbances, it was observed that most of the voltage disturbances could be characterized by the values of the frequency resolutions, except for voltage sags and swells. The results of the analysis showed that both voltage sags and swells have the same frequency resolution ranging from 0.000 to 0.0061. Harmonics can be detected between the frequency resolutions of 0.0061 and 0.022, and notches are detected between 0.022 and 0.080. Oscillatory and impulsive transients can be detected between the frequency resolutions of (0.080 to 0.4) and (0.4 to 0.5), respectively. The summary of the second set of features selected based on frequency resolutions are explained in Figure 5 and Table 2. The performance of these new features in detecting power quality disturbances will be presented in other section of this paper.

Figure 5: Features from the S-transform frequency resolutions

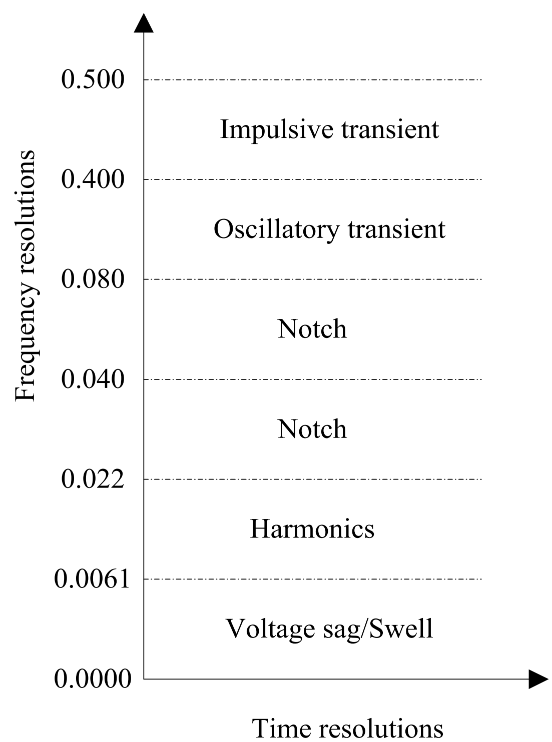


Table 2: Descriptions of features based on the values of the frequency resolutions in the S-Matrix

Features	Description
F5	Values of frequency resolution from 0.0061 to 0.022
F6	Values of frequency resolution from 0.022 to 0.04
F7	Values of frequency resolution from 0.04 to 0.08
F8	Values of frequency resolution from 0.08 to 0.40
F9	Values of frequency resolution from 0.40 to 0.50

4. Support Vector Regression for prediction of power quality disturbances

The foundations of Support Vector Machines (SVMs) have been developed by Vapnik and Cortez (1995), and they are gaining popularity due to many attractive features and promising empirical performance. Their formulation embodies the structural risk minimization (SRM) principle, which has been shown to be superior to the traditional empirical risk minimization (ERM) principle, employed by conventional neural networks (Vapnik, 1995). SRM minimizes an upper bound on the expected risk, as opposed to ERM, which minimizes the error on the training data. It is this difference that equips SVMs with a greater ability to generalize, which is the goal in statistical learning. Initially, SVMs were developed to solve classification problems, but recently, they have been extended to the regression problem domain (Vapnik et al, 1996). The term SVM is typically used to describe classification with support vector methods, and support vector regression (SVR) is used to describe regression with support vector methods.

Originally, the SVM was introduced within the context of statistical learning theory and SRM. Viewing input data as two sets of vectors in an n-dimensional space, an SVM will construct a separating hyperplane in that space, one that maximizes the margin between the two data sets. To calculate the margin, two parallel hyperplanes are constructed, one on each side of the separating hyperplane, which are "pushed up against" the two data sets. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the neighboring data points of both classes, since, in general, the larger the margin is, the better the generalization error of the classifier. Consider the sample data below, i.e., training data, which are a set of points of the form:

$$D = \{(x_i, c_i) \mid x_i \in R^p, c_i \in \{-1,1\}\}_{i=1}^{n} \qquad (6)$$

where c_i is either 1 or −1, indicating the class to which the point x_i belongs. Each x_i is a *p*-dimensional real vector. To classify the data set, a maximum-margin hyperplane is required that can divide the points having $c_i = 1$ from those having $c_i = -1$. The classification is done by means of a dividing hyperplane, which takes the form of the equation:

$$f(x) = w^T x - b \qquad (7)$$

where w : orthogonal weight vector $[w_1, w_2, ...w_n]^T$ and b : a scalar used to increase the margin

The vector w is a normal vector: it is perpendicular to the hyperplane. The parameter b determines the offset of the hyperplane from the origin along the normal vector w. The dividing hyperplane is achieved when $w.x - b = 0$. Without this parameter, the hyperplane created will be restricted to pass through the origin only. Thus, the position of the dividing hyperplane is determined by the vector *w* and scalar *b*. In order to distinguish between the two classes, label y_i is used as:

$$y_i = +1 \text{ if } x \quad \text{belong to Class 1,} \qquad (8)$$

$$y_i = -1 \text{ if } x \quad \text{belong to Class 2,} \qquad (9)$$

In other words, the dividing hyperplane has to follow the following constraints:

$$f(x_i) \geq 0, \quad \text{if } y_i = +1 \qquad (10)$$

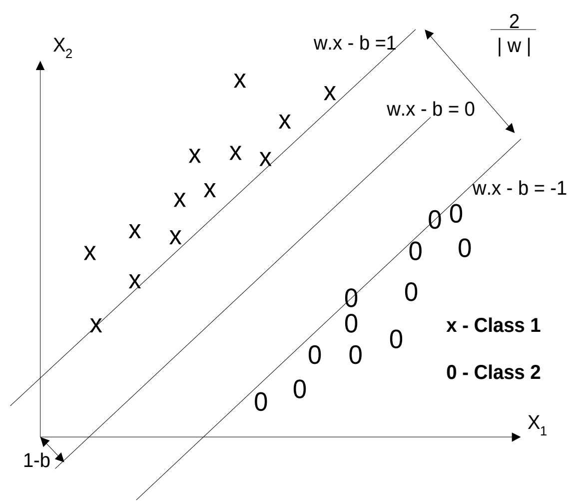
$f(x_i) \le 0$, if $y_i = -1$ (11)

Thus, two parallel hyperplanes are created on each side of the dividing hyperplane, satisfying the above constraints. The hyperplanes can be described by these equations:

$w.x - b = +1$ and $w.x - b = -1$ (12)

The samples on the margin in Figure 3 are called the support vectors. By using geometry, the distance between these two hyperplanes is $\frac{2}{|w|}$. The optimal dividing hyperplane is obtained by maximizing the margin $\frac{2}{|w|}$, such that there are no points between the above parallel hyperplanes and thus minimizing $|w|$.

Figure 3: Maximum-margin hyperplane and margins for an SVM trained with samples from two classes



In statistics, regression analysis is a collective name for techniques used in the modeling and analysis of numerical data consisting of values of a dependent variable (also called response variable or measurement) and of one or more independent variables (also known as explanatory variables or predictors) (Rodríguez, 1996). The dependent variable in the regression equation is modeled as a function of the independent variables, corresponding parameters ("constants"), and an error term. The error term is treated as a random variable. It represents unexplained variation in the dependent variable. The parameters are estimated so as to give a "best fit" of the data. Most commonly, the best fit is evaluated by using the least squares method, but other criteria have also been used. The uses of regression rely heavily on the underlying assumptions being satisfied. The support vector method can also be applied to the case of regression while maintaining all the main features that characterize the maximal margin algorithm: a non-linear function is learned by a linear learning machine in a kernel-induced feature space, while the capacity of the system is controlled by a parameter that does not depend on the dimensionality of the space. As explained earlier, the roots of SVM lie in statistical learning theory, which describes properties of learning machines that enable them to generalize well to unseen data. In the case of SVR, the goal is to find a function that predicts the target values of the training data with a deviation of at most ε while requiring this function to be as flat as possible (Drucker et al, 1998). According to Smola and Scholkpof (1998), the core of the support vector algorithm does this for linear functions $f(x) = (w,x) + b$, where (w,x) denotes the dot product of vectors w and x, thereby enforcing flatness by minimizing |w| (|w| denotes the Euclidian norm of vector w). By using a dual representation of the minimization problem, the algorithm requires only dot products of the input patterns. This allows the application of non-linear regression by using a kernel function that represents the dot product of the two transformed vectors. The support vector algorithm will now fit a flat-as-possible function by searching for a suitable separating hyperplane for the SVR.

In this paper, the SVR performed regression to predict the existence of single and multiple power quality disturbances. Two types of SVRs were developed based on two different kernel functions for performance comparison. The use of the kernel trick provides a powerful way to obtain non-linear algorithms capable of handling non-separable data sets in the original input space. The idea of the kernel function is to enable operations to be performed in the input space rather than the potentially high-dimensional feature space. The basic concept is to construct a mapping into a higher dimensional feature space by the use of reproducing kernels. However, the computation is still heavily dependent upon the number of training patterns, and generating a good data distribution for a high-dimensional problem will generally require a large training set. The first SVR uses the radial basis function (RBF) kernel and is shown in equation 13. The parameter σ is associated with the RBF function and will be tuned in order to get the targeted results.

$$k(x, y) = e^{-(\frac{|x-y|^2}{2\sigma^2})} \tag{13}$$

The second SVR was developed based on the multi-layer perceptron (MLP) kernel, and the equation for the MLP kernel is shown in equation 14. The value of the kernel will depend on certain values of the scale, ρ, and offset, ∂, parameters. Here the MLP SVR corresponds to the first layer and the Lagrange multipliers to the weights.

$$k(x, y) = \tanh(\rho(x, y) + \partial) \tag{14}$$

5. Experimental tests and results

5.1. Preparation of training database and testing data

The performance of the SVR based S-transform is dependent on the training database. The first part of the study involved the development of the training database. The training database was developed based on analyses performed on 525 number of disturbance data with known causes. The measurement data included a short pre-fault waveform (approximately 6 cycles long) followed by the actual disturbance and a post fault waveform (approximately 10 cycles). The description of the classes of the power quality disturbances to be predicted by the SVR based S-transform is shown in table 3. Overall there are 21 numbers of classes to be predicted by the SVR.

Table 3: Descriptions of classes of power quality events

Types of power quality disturbances	Classes
Pure waveform/Normal voltage	C1
Voltage sag	C2
Voltage swell	C3
Harmonics	C4
Notches	C5
Oscillatory transient	C6
Impulsive transient	C7
Sag & harmonic	C8
Sag & notch	C9
Sag & Oscillatory transient	C10
Sag, harmonic & notch	C11
Sag, notch & oscillatory transient	C12
Swell & harmonic	C13
Swell & notch	C14
Swell & oscillatory transient	C15
Swell, harmonic & notch	C16
Swell, notch & oscillatory transient	C17
Harmonic & notch	C18
Notch & oscillatory transient	C19
Sag & impulsive transient	C20
Sag, harmonic & impulsive transient	C21

In table 4 are the statistics for both the training and testing data for the experiments. The testing data are the data to be analyzed and predicted by the SVR for their respective classes. Lastly, in order to make the SVR training database, two extra parameters are needed which the details and values are explained in table 5.

Table 4: Number of voltage disturbances for training and testing of SVR

Disturbance class	Number of training data of each disturbance class	Number of data for testing for each class
C1	25	31
C2	25	31
C3	25	31
C4	25	31
C5	25	31
C6	25	31
C7	25	31
C8	25	31
C9	25	31
C10	25	31
C11	25	31
C12	25	31
C13	25	31
C14	25	31
C15	25	31
C16	25	31
C17	25	31
C18	25	31
C19	25	31
C20	25	31
C21	25	31
Total	**525**	**651**

Table 5: Description and values of the parameters for the SVR

Parameters	Descriptions	Values
γ (gam)	This is the regularization parameter, determining the trade-off between the fitting error minimization and smoothness	10
σ^2 (sig2)	This is the bandwidth of the RBF Kernel	0.2

The first step in the experiments was to extract all the nine features using the S-transform for the 525 numbers of training data and to rearrange the features based on the format in table 6 in order to classify the disturbances. This data will be termed as training database for the SVR. Next, the testing data will be analyzed using the SVR based S-transform and new set of features will be extracted for making prediction of the disturbances. Next, two experiments using both the SVR techniques, the non-linear radial basis function (RBF) SVR and the multi-layer perceptron (MLP) SVR, were conducted. In these experiments, both the SVRs were trained with the training data base. Once the trainings were completed, both the RBF SVR and MLP SVR were tested for their abilities to perform prediction for the 651 numbers of testing data.

Table 6: Features arrangement for data prediction by the SVR

F1	F2	F3	F4	F5	F6	F7	F8	F9	Class
0	0	0.99836	1.00340	0	0	0	0	0	C1
11	0	0.80734	1.00067	0	0	0	0	0	C2
0	348	0.97844	1.26610	0	0	0	0	0	C3
0	0	0.99384	1.00616	5	0	0	0	0	C4
0	0	0.98136	1.01888	0	84	69	0	0	C5
0	0	0.95790	1.00510	0	0	123	346	0	C6
0	0	0.99794	1.00484	0	0	0	1272	1085	C7
133	0	0.54167	1.00517	46	0	0	0	0	C8
1400	0	0.87040	0.91010	12	0	77	0	0	C9
850	0	0.11020	1.00540	0	0	0	88	0	C10
117	0	0.80827	1.00000	111	0	0	92	0	C11
653	0	0.45020	1.08100	0	100	100	100	0	C12
0	806	0.99862	1.69500	25	0	0	0.00	0	C13
0	541	0.95210	1.14020	23	500	500	500	0	C14
0	1543	0.91110	1.45120	0	0	0	11	0	C15
0	850	1.00500	1.25110	27	0	0	88	0	C16
0	200	0.96010	1.76120	0	98	75	123	0	C17
0	0	0.99467	1.06154	2302	0	298	0	0	C18
0	0	0.99896	1.00616	0	0	0	251	0	C19
28	0	0.87870	1.00011	277	1521	0	2425	128	C20
478	0	0.86054	1.00002	12270	34734	0	54899	20466	C21

5.2. Prediction results

The results of the conducted experiments are shown in table 7 through table 12. These tables contained the prediction results in terms of correct prediction (diagonal elements) and mis-prediction (the numbers outside the diagonal elements). Furthermore the overall prediction rate (i.e., the number of correctly predicted disturbances divided by the total number of disturbances) is also given in the tables.

Table 7: RBF SVR prediction results for experiment 1 for classes C1 to C8

	C1	C2	C3	C4	C5	C6	C7	C8	Prediction rate %
C1	**31**	0	0	0	0	0	0	0	100.0%
C2	0	**31**	0	0	0	0	0	0	100.0%
C3	0	0	**31**	0	0	0	0	0	100.0%
C4	0	0	0	**28**	1	1	0	1	90.3%
C5	0	0	1	0	**28**	2	0	0	90.3%
C6	0	0	0	0	0	**27**	2	2	87.1%
C7	0	0	0	0	0	1	**28**	2	90.3%
C8	0	0	0	0	2	1	0	**28**	90.3%

Overall prediction accuracy: 93.5 %

Table 8: RBF SVR prediction results for experiment 1 for classes C9 to C15

	C9	C10	C11	C12	C13	C14	C15	Prediction rate %
C9	**29**	0	0	0	1	1	2	93.5%
C10	0	**29**	2	0	0	0	0	93.5%
C11	0	0	**29**	1	0	1	0	93.5%
C12	0	0	1	**30**	0	0	0	96.8%
C13	0	0	0	2	**29**	0	0	93.5%
C14	0	0	0	0	3	**28**	0	90.3%
C15	0	0	0	0	2	1	**28**	90.3%

Overall prediction accuracy: 93.1 %.

Table 9: RBF SVR prediction results for the experiment 1 for classes C16 to C21

	C16	C17	C18	C19	C20	C21	Prediction rate %
C16	**29**	1	0	0	1	0	93.5%
C17	1	**30**	0	2	1	0	96.8%
C18	0	1	**30**	0	0	0	96.8%
C19	0	0	0	**31**	0	0	100.0%
C20	0	0	2	1	**29**	1	93.5%
C21	0	0	0	2	1	**28**	90.3%

Overall prediction accuracy: 95.2 %.

Table 10: MLP SVR prediction results for experiment 2 for classes C1 to C8

	C1	C2	C3	C4	C5	C6	C7	C8	Prediction rate %
C1	**31**	0	0	0	0	0	0	0	100.0%
C2	0	**31**	0	0	0	0	0	0	100.0%
C3	0	0	**31**	0	0	0	0	0	100.0%
C4	0	0	0	**25**	2	1	0	3	80.6%
C5	0	0	1	1	**23**	3	0	3	74.2%
C6	0	0	1	1	0	**27**	2	0	87.1%
C7	0	0	1	1	1	1	**25**	2	80.6%
C8	0	0	1	1	1	0	1	**27**	87.1%

Overall prediction accuracy: 88.7%.

Table 11: MLP SVR prediction results for experiment 2 for classes C9 to C15

	C9	C10	C11	C12	C13	C14	C15	Prediction rate %
C9	**29**	1	0	1	0	0	0	93.5%
C10	0	**26**	0	1	1	0	3	83.9%
C11	0	1	**25**	2	1	0	2	80.6%
C12	2	1	1	**25**	0	0	2	80.6%
C13	0	0	0	2	**29**	0	0	93.5%
C14	1	1	0	0	1	**27**	1	87.1%
C15	1	1	1	1	1	0	**26**	83.9%

Overall prediction accuracy: 86.2 %.

Table 12: MLP SVR prediction results for experiment 2 for classes C16 to C21

	C16	C17	C18	C19	C20	C21	Prediction rate %
C16	**28**	2	0	0	0	1	90.3%
C17	1	**28**	0	1	0	1	90.3%
C18	2	1	**24**	0	2	2	77.4%
C19	2	2	0	**25**	2	0	80.6%
C20	3	2	2	1	**23**	0	74.2%
C21	0	1	1	2	0	**27**	87.1%

Overall prediction accuracy: 83.3 %.

5.3. Comments regarding the experiments

In this paper, two experiments were performed to evaluate the performance of the novel approach in the prediction of single and multiple power quality disturbances. Both the experiments applied all the nine features extracted from the S-transform. The first experiment was conducted using the RBF SVR, and the second experiment was done using the MLP SVR. The individual detection rates for experiment 1 range from 93.1% to 95.2%. These rates were sufficiently high to validate the high performance of the RBF SVR based S-transform technique. Based on these results, it was proven that the RBF SVR based S-transform technique was able to predict precisely all the classes of power quality disturbances in the voltage signals with an overall detection rate of 93.9%. In experiment 2, the individual detection rate range from 84.4% to 89.9%. The overall detection rate for the MLP SVR was at 86.1% which is lower than the RBF SVR technique. It is important to note that for the RBF SVR, the kernel selected was the Gaussian radial basis function and that the corresponding feature space is a Hilbert space of infinite dimension. As maximum margin classifiers are well regularized, the infinite dimension does not spoil the results but improves the efficiency of the RBF SVR (Gunn, 1998).

6. Conclusion

In this paper, a novel approach to perform data prediction using Support Vector Regression (SVR) based S-transform techniques is presented. The results of the two experiments performed in this study gave high prediction accuracy, implying that the SVR based S-transform prediction technique is an attractive choice for performing prediction for both single and multiple power quality disturbances. In performing prediction, generalized control of the prediction is obtained by maximizing the hyperplane margin, which corresponds to minimizing the weight vector in a canonical framework. The support vectors lie on the boundary and, as such, summarize the information required to separate the disturbance data. The choice of the kernel functions could also dictate the prediction accuracy of the SVR. Both the RBF and MLP kernel mappings provide a unifying framework for the SVR model architectures, enabling prediction to be performed. In this study, the RBF SVR was proven to be superior to the MLP SVR in the prediction of single and multiple power quality disturbances. This new prediction method will be useful to the PQMS for performing for real-time prediction of the classes of the recorded power quality disturbances.

7. Acknowledgement

This work was carried out with the financial support from the Universiti Kebangsaan Malaysia under the research grants UKM-GUP-BTT-07-25-151 and UKM-OUP-TK-15-68/2009.

References

[1] Dugan, R.C, McGranaghan, M.F, Betty, H.W, Electrical Power Systems Quality, McGraw Hill Companies. ISBN 0070180318, (1996). pp. 240-242.

[2] Stockwell, R.G, Mansinha, L, Lowe, R.P, Localization of the Complex Spectrum: The S-transform, IEEE Transactions on signal processing, vol.44, no.4 April 1996, S 1053-587X (96(02)790-0). DOI: 10.1109/78.492555, (1996).

[3] Pinnegar, C.R., Mansinha, L, The S-transform with windows of arbitrary and varying shape, Journal of Geophysics, DOI: 10.1190/1.1543223. Vol.68, No.1 (January-February 2003), (2003) pp. 381-385, Available Online from: http://www.cora.nwra.com/stockwell/rgspages/S-Transform/st_otherauthors.shtml

[4] Pinnegar, C.R., Mansinha, L, Time local Fourier analysis with a scalable, phase modulated analyzing function: the S-transform with a complex window, Signal Processing , (2004) pp.1167-1176, Available Online from: http://www.cora.nwra.com/~stockwel/rgspages/pdfdocs/other/Pinnegar_complex_window_ST.pdf

[5] Weeks, W.L, Steiner, J.P, Instrumentation for the detection and location of incipient fault on power cable, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-lOl, No. 7 July 1982.

[6] Vapnik, V, Cortez, C, Support-vector networks. Machine Learning (1995), pp.273-297. DOI=10.1.1.15.9362. Available online from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.15.9362

[7] Vapnik, V, The Nature of Statistical Learning Theory. Springer, N.Y, ISBN 0387945598, (1995)

[8] Vapnik, V, Golowich, S. and Smola, A. Support vector method for function approximation, regression estimation, and signal processing, (1996), DOI: 10.1.1.41.3139. Available online from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.41.3139

[9] Rodríguez, C.C. Definition of regression, 1996, http://omega.albany.edu:8008/

[10] Drucker, H, Burges, C.J.C, Kaufman, L, Smola, A, Vapnik, V, Support Vector Regression Machines, (1998), Available on-line: http://www.cs.cmu.edu/~pakyan/compbio/references/Drucker_NIPS_1996.pdf

[11] Smola, A.J. and Scholkopf, B, A Tutorial on Support Vector Regression, (1998), Available on-line: http://www.svms.org/regression/SmSc98.pdf

[12] Gunn, S.R. Support Vector Machines for Classification and Regression, University of Southampton, 10th May 1998, (1998), Available on-line: http://trevinca.ei.uvigo.es/~cernadas/tc03/mc/SVM.pdf