Interpretation of Wavelet Analysis and Its Application in Partial Discharge Detection

X. Ma, C. Zhou and I. J. Kemp

School of Engineering, Science and Design
Glasgow Caledonian University,
Cowcadden Road, Glasgow, G4 0BA, UK

ABSTRACT

The objective of the paper is to discuss a tool which is proving extremely efficient in partial discharge measurement studies. Though the technique itself is not new, its application to partial discharge studies is. It will be demonstrated in this paper that it has tremendous power and this accounts for its rapid growth as an application in this field. The paper begins with the description of the fundamentals of wavelet analysis, wavelet categories and the properties of the associated wavelet transforms. PD pulses as acquired from detectors composed of different detection circuits are investigated and numerically simulated, and a method on how to select optimally the wavelet corresponding to the representative forms of PD pulse is then presented. Finally, applications of wavelet analysis to partial discharge studies are explored. The paper demonstrates that the wavelet based denoising method proposed in the paper can be employed in separating PD pulses from electrical noise successfully and can be used in pulse propagation studies of partial discharge in distributed impedance plant to provide enhanced information and further infer the original site of the PD pulse.

1 INTRODUCTION

FAILURE of the solid insulating systems of HV equipment (rotating machines, transformers, bushings etc.) can lead to catastrophic failure of the equipment with serious health and safety, environmental and economic consequences. It is therefore imperative that degradation in such systems be detected and quantified at an early stage in its development so that appropriate replacement/refurbishment can be arranged at a convenient time. When degradation occurs in such systems, irrespective of the causative mechanism, it generally results in partial discharges being generated (small electrical sparks within the insulation). Once prevalent, partial discharges then represent the dominant mechanism of degradation. For this reason, it can be appreciated why the detection and characterization of partial discharge activity is a key approach in insulation system condition monitoring [1].

Partial discharge (PD) detection involves the capture, storage and processing of PD signals, which occur in the form of individual or series of electrical pulses. By acquiring the data at a convenient, non-intrusive location, usually the phase or neutral terminals of the power equipment under test, statistical quantities such as the PD-pulse magnitude, phase distribution (over the ac power cycle) and their variation with time can be obtained [2]. Furthermore, based on these quantities, the mechanisms of PD

activity can be categorized, and the nature, form and the extent of degradation can be inferred.

Traditionally, the techniques used for signal processing are realized in either the time or frequency domain to analyze and extract PD events. In the case of the frequency domain, for instance, the Fourier transform assumes that any signal could be decomposed into a series of sine and cosine waveforms and then summarized by these basis functions. Although the signal under analysis can be localized arbitrarily throughout the frequency domain following the Fourier transform, information in time is, however, lost. With regard to the PD pulse structure, there always exist non-periodic and fast transient features in the PD signals detected, which tend to be ignored and cannot be revealed efficiently and explicitly by this kind of conventional transform. For these reasons, when the Fourier transform is applied to partial discharge analysis, it has serious limitations.

Despite the advances achieved during past decades in measuring instrumentation, partial discharge detection circuits, e.g., the ultra wide-band detectors coupled to real-time oscilloscopes of very wide bandwidth (up to 4 GHz), a sophisticated analytical tool of similar power to pre-process PD data has yet to be realized.

Since its introduction to practical applications in the mid-1980s as a powerful mathematical means of signal

analysis and signal processing, the wavelet and its associated transforms [3–5] have been increasingly recognized by researchers and engineers and utilized to solve many engineering problems. This paper describes the use of the wavelet transform to obtain two-dimensional information on PD pulses in both the time and frequency domain, and to extract features of PD pulses in the measurement data. The paper centers on the following questions:

(1) What is the wavelet, its categories and the associated transforms?

(2) How is the wavelet, most suitable for analysis of PD measurement data, acquired in different detectors?

(3) How can the specific wavelet "pattern" following the wavelet transform (WT) be used to differentiate PD from other signals such as noise?

(4) How successfully can the WT be utilized in practice?

2 WAVELET METHODOLOGY

A wavelet, as its name implies, can be interpreted as a small wave that has a limited duration and a zero mean value. It oscillates in amplitude and decays to zero quickly on both sides of the central position of the waveform. Compared to sine and cosine waves, the basis functions of the Fourier transform, which extend from minus to plus infinity, wavelets usually tend to be irregular and asymmetric in terms of wave shape. An example wavelet $\psi(t)$ is shown in Figure 1a.

From a mathematical viewpoint, if a wavelet is defined as $\psi(t)$, it has the property of equation (1) and has to satisfy the admissibility condition $C_\psi (0 < C_\psi < \infty)$ in equa-

tion (2) so that a signal can be decomposed and then reconstructed perfectly through the wavelet transform.

$$\int_R \psi(t)\,dt = 0 \tag{1}$$

$$C_\psi = 2\pi \int_R \frac{|\hat{\psi}(\omega)|^2}{\omega}\,d\omega \tag{2}$$

where $\hat{\psi}$ is the form of the Fourier transform of $\psi(t)$, and R the set of real numbers. In wavelet terms, $\psi(t)$ is referred to as the mother wavelet. The family of scaled wavelets associated with $\psi(t)$ can be described as

$$\psi(at) = a^{-1/2}\psi\left(\frac{t}{a}\right), \quad (a = 1,2,3,\dots) \tag{3}$$

where the scaling variable a determines the amplitude and the duration of $\psi(at)$. More precisely, the amplitude of $\psi(at)$ is inversely proportional to the value of the square root of a, whereas its duration is linearly proportional to a, thus ensuring that each scaled wavelet has the equivalent energy of the mother wavelet, as can be seen from Figure 1a.

Also as can be observed in Figure 1b, the scaled wavelets of $\psi(at)$ have different central frequencies ω and bandwidth ranges $\Delta\omega$, high values of a corresponding to low values of ω and $\Delta\omega$, whereas low values of a yielding high values of ω and $\Delta\omega$. If t in expression (3) is replaced by $t - b$, the wavelet is shifted by b on the time axis. Therefore, localization properties in both the time and frequency can be achieved simultaneously when the signals under analysis are examined using such wavelets.

There are several families of wavelets already proven to be especially useful in practice [6]. In terms of wavelet expressions, some are given explicitly by their time definition, including the Gaussian wavelet, the Mexican hat wavelet and the Morlet wavelet, and some by their frequency definition such as the Meyer wavelet. Others are constructed by their finite impulse response (FIR) filter pair that exist a special mirror relationship between filters, also known as quadrature mirror filters (QMFs), typical examples of which are Daubechies wavelets, Symlet wavelets, Coiflet wavelets and Lemarie wavelets. The Daubechies wavelets constituted from 4 and 14 coefficients in QMFs are shown in Figures 2a, and 2b respectively; the former is defined as db2, the latter as db7. Among the wavelets available, the Daubechies wavelet family has almost all of the required properties such as compactness, limited duration, orthogonality and asymmetry for analysis of fast transient, irregular pulses. It is for this reason that the Daubechies wavelet can be applied effectively in partial discharge studies.

Consider now the types of wavelet transform (more theoretical details can be found in [3–5]). The key idea underlying the wavelet transform (WT) strategy is that a given
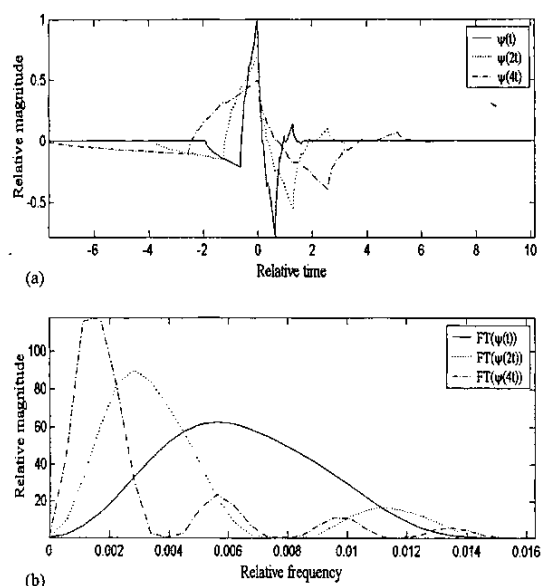

(a)

(b)

Figure 1. Comparisons between scaled wavelets in the time domain and in the frequency domain. a, scaled wavelets in the time domain; b, their results following the Fourier transform.
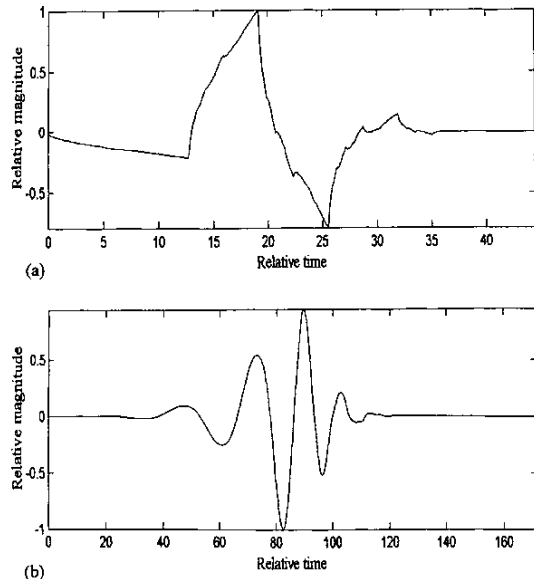
(a)

(b)

Figure 2. Examples of the shape of wavelets. a, db2 wavelet; b, db7 wavelet.

signal can be disassembled into a series of scaled and time shifted forms of mother wavelet producing a time-scale view of a signal from which the original signal can be recovered. The continuous wavelet transform (CWT) of a

time-dependent signal $f(t)$ is described as

$$W_\psi f(a,b) = |a|^{-1/2} \int f(t) \psi\left(\frac{t-b}{a}\right) dt \qquad (4)$$

The CWT calculates the wavelet coefficients at every possible scale and along every time instant. The value of $W_\psi f(a,b)$ represents the similarity extent between the examined section of $f(t)$ and the scaled and shifted wavelets. The greater the $W_\psi f(a,b)$, the more energetic the $W_\psi f(a,b)$, and the greater the similarity between the wavelets and the original signal.

The discrete wavelet transform (DWT) is realized by characterizing the scale variable a in $a = 2^m$ and assuming the time variable b as $b = n2^m$ at a given scale m (m,n being integer values), thus avoiding time consumed in coefficient calculation and eliminating wavelet coefficient redundancy. The practical fast discrete wavelet transform (FDWT) is implemented based on the filtering scheme of QMF pairs g, h at different scales as illustrated in Figure 3. When a signal is fed through a pair of low-pass filter g and high-pass filter h respectively, the WT yields the low frequency content and the high frequency content of the signal respectively, the former being known as the signal's "approximations" and the latter known as its "details".

The signal can be perfectly reconstructed using the inverse discrete wavelet transform (IDWT) by starting from
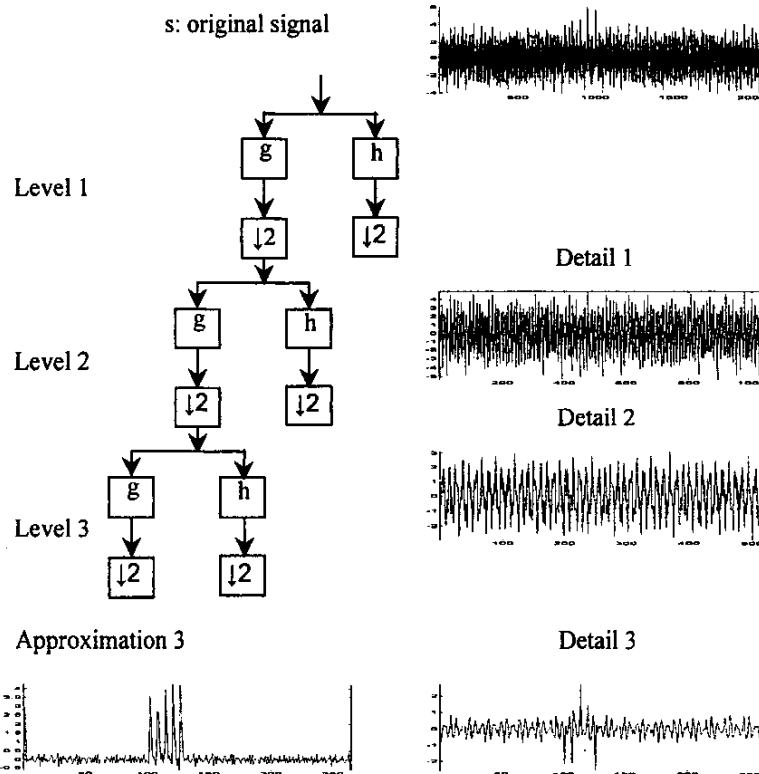

Figure 3. Tree structure of three level FDWT decomposition of a signal.

the bottom components and moving to the root component rather than employing the top to bottom approach of the FDWT. At each synthesis level, both the coarser approximations and details are firstly up-sampled by the factor of 2 and then convoluted with synthesis filters $\bar{g}$, $\bar{h}$, producing a finer approximation signal. The coefficients in synthesis filters $\bar{g}$, $\bar{h}$ are identical with those in decomposition filters g, h, but are inverse in the time arrangement.

3 PULSE SHAPE OF PARTIAL DISCHARGES

The selection of a suitable wavelet and its associated analysis algorithm is application dependent. For example, time-defined wavelets such as the Gaussian wavelet, the Mexican hat wavelet and their CWT have often been used for pattern recognition [7], and the Morlet wavelet for high-impedance fault detection [8] and the analysis of disturbances within power systems [9]. Dyadic orthogonal wavelets such as the Daubechies wavelet have found applications in data compression and signal denoising [10–13].

Prior to discussing wavelet selection with respect to partial discharge studies, it is important to discuss PD pulse shape. For solid or fluid dielectrics, partial discharge phenomena are mainly caused by voids or bubbles within or at the surface of insulating materials. The charge displacement in the system produces current flowing in the leads to the object, which can be detected by a coupling capacitor in parallel, or by an impedance in series, with the test sample. Normally discharge pulses are recognized as having a very steep wave-front and short duration. The rise time of the discharge pulse is as low as a few ns (perhaps even faster) whereas its duration is in the order of 10 ns. The pulse shape of the discharge current is determined by the cavity dimension, such as its radius and height, and the external measuring circuit, whereas the number of almost simultaneously occurring discharges is determined by the cavity shape. For instance, spherical cavities cause one individual discharge at a time, whereas flat cavities yield many discharges at a time [14].

In relation to detection methods, both narrow-band and ultra wide-band (UWB) detecting circuits have been adopted. Narrow-band detection can give an erroneous quantification and information loss due to quasi-integration adopted or when multiple overlapped pulses exist. The characteristic parameters, which describe the shape of an individual discharge pulse, cannot be described explicitly in narrow-band detection. With ultra wide-band (UWB) detection techniques [14,15] coupled to fast oscilloscope acquisition, the true PD pulse shape can be recorded and the shape of the individual PD pulse can be observed. Compared to the phase-resolved measurement primarily related to narrow-band detection, UWB detec-
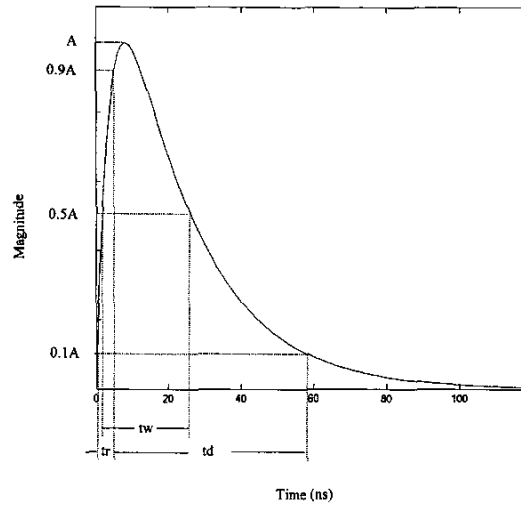

Time (ns)

Figure 4. Typical time parameters describing the shape of a PD current pulse.

tion is primarily associated with time-resolved measurements.

Figure 4 demonstrates a typical PD current pulse, which could be characterized by a series of discharge parameters including A (pulse peak value), t_r (pulse rise time from 10% to 90% levels), t_w (pulse width between 50% levels) and t_d (pulse decay time from 90% to 10% levels) [11,14]. In practical measurements, discharge voltage signals are captured by feeding the discharge current through a detection circuit. On this basis, detected voltage signals are likely to have different pulse shapes, depending on the configuration of the detection circuit. As far as the detection circuit [16] is concerned, it is realized in either a RC impedance circuit or a RLC impedance circuit, as shown in Figures 5(a) and 5(b) respectively. The transfer functions of RC and RLC impedance circuits can be expressed in the following Laplace forms

$$G_{RC}(s) = \frac{1}{C} \frac{1}{s + 1/\tau} \tag{5}$$

$$G_{RLC}(s) = \frac{1}{C} \frac{s}{s^2 + s/\tau + \omega_0^2} \tag{6}$$

where $\tau = RC$, and $\omega_0 = 1/\sqrt{LC}$. For the input of a Dirac current pulse $i(t)$ (ideal impulse, no time duration), the output voltage pulse $v(t)$ is represented as a damped exponential pulse in the RC impedance circuit and a damped
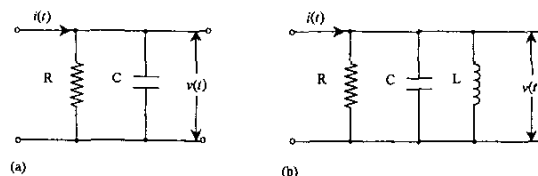

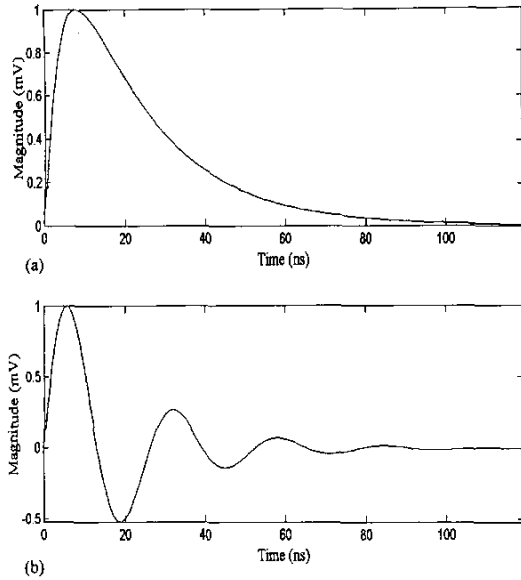**Figure 5.** Detection impedance. a, RC circuit; b, RLC circuit.

Figure 6. Two typical PD pulses. a, simulated DEP-type pulse; b, simulated DOP-type pulse.

oscillatory pulse in the RLC impedance circuit. The value of τ controls the decay time of both representative forms of pulse and ω_0 determines the oscillation frequency of the latter output pulse.

In practice, the current pulse $i(t)$ caused by a PD is not an ideal Dirac current, because it has duration in time, as shown in Figure 4. When a practical PD pulse passes $i(t)$ through a detection circuit, the output pulse produced always has a finite rise time. Considering the shape of a PD current pulse and the characteristics of detection circuits, the damped exponential pulse (DEP) and the damped oscillatory pulse (DOP) can be numerically simulated as the mathematical models of (7) and (8) respectively, and further displayed in Figures 6a and 6b.

$$DEP(t) = A(e^{-t/t_1} - e^{-t/t_2}) \qquad (7)$$

$$DOP(t) = A\sin(2\pi f_c t)(e^{-t/t_1} - e^{-t/t_2}) \qquad (8)$$

where A gives the pulse peak value, t_1, t_2 the damping coefficients that determine those typical characteristic parameters t_r, t_w, t_d describing the pulse shape, whilst f_c is the oscillatory frequency of the DOP-type pulse.

4   OPTIMAL WAVELET SELECTION

The WT of a signal produces a wavelet detail coefficient distribution throughout the entire time-scale (frequency) view which, referred to as the "wavelet pattern" of a signal, depends on the wavelet chosen in the WT. For a better understanding of this, an example of a discharge signal with 5 successive DEP-type pulses is given,
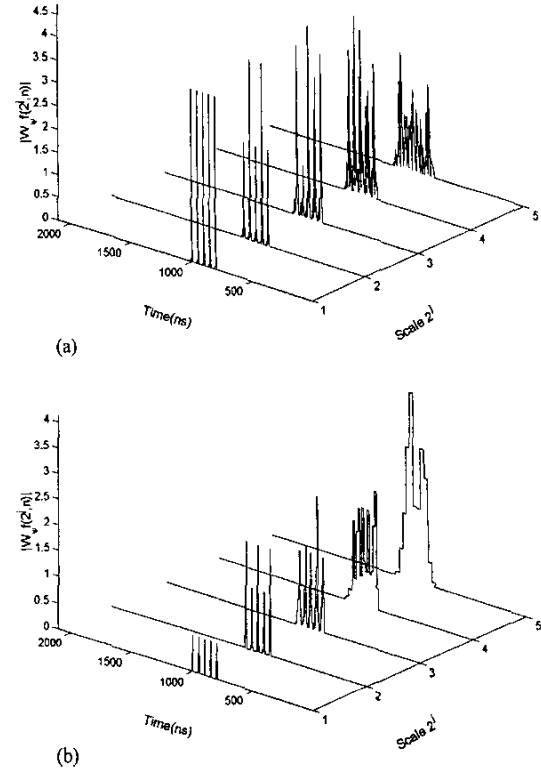

Figure 7. Wavelet patterns of a DEP-type pulse. Sampling frequency, 1GHz; data record, 2100 points. a, With the db2 wavelet; b, With the db7 wavelet.

where $t_r = 5ns$, $t_w = 10ns$ and $t_d = 20ns$ in each pulse. Its wavelet pattern after applying FDWT with 5-level (scale) using the db2 and db7 wavelet is displayed in Figures 7a and 7b respectively. Here, FDWT with low level is chosen as the PD pattern can be characterized within such a level range. The same reasoning applies for the level determination later in characterizing the patterns of noise. Despite the similarity in both patterns, the values of their coefficient are significantly different, as seen from all levels. The "optimal" wavelet suitable for analyzing a given signal is the one which is capable of generating as many coefficients with maximal values as possible within the time-scale domain. In this case, db2 is a more appropriate wavelet than db7 for analyzing this type of PD pulse.

In statistical analysis, the correlation coefficient γ [17] is used to detect one particular relationship between variables. The greater the value of γ, the more approximate in wave shape between two variables. Uncorrelated data sets result in an γ of 0, whereas equivalent data sets have an γ of 1. For this reason, the correlation coefficient γ can also be used as a criterion in identifying an appropriate wavelet choice for PD pulse examination as described below:

1. the discharge pulse data set is first obtained from its mathematical expression (7) or (8), and

Table 1. Values of correlation coefficient γ between wavelets and the given PD pulses.

Wavelets	DEP-type	DOP-type
db2	0.5272	0.2410
db3	0.4557	0.2042
db4	0.3761	0.3281
db5	0.2213	0.3526
db6	0.1443	0.7747
db7	0.0234	0.7841
db8	0.0447	0.7257
db9	0.0632	0.6590
db10	0.0151	0.5607
db11	0.0196	0.5894
db12	0.0633	0.6108
db13	0.0708	0.6055
db14	0.0622	0.6075
db15	0.0516	0.6057
db16	0.0231	0.5949
db17	0.0212	0.5755
db18	0.0418	0.5648
db19	0.0527	0.5517
db20	0.0484	0.5351
sym2	0.5237	0.2410
sym3	0.4557	0.3486
lem2	0.5126	0.2800
lem3	0.2462	0.5972

Figure 8. Comparisons between the PD pulse and the determined optimal wavelet. a, optimal wavelet for DEP-type PD pulse; b, optimal wavelet for DOP-type PD pulse.
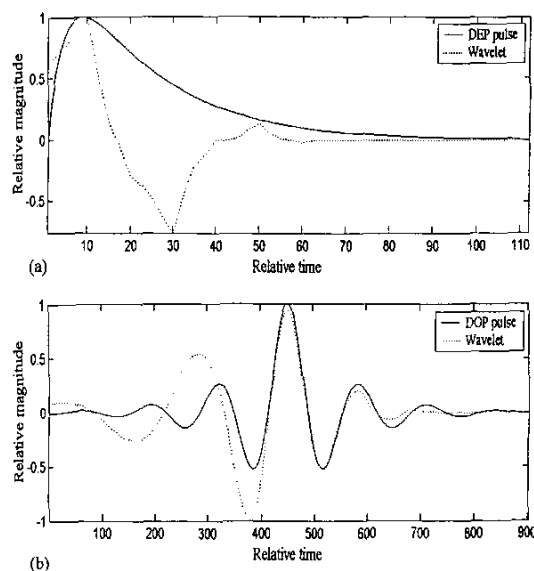
2. the wavelet data set is acquired from the available wavelet families by use of the aforementioned methods in the section of wavelet methodology,

3. a set of correlation coefficients γ is then calculated and finally

4. after comparing all of the correlation coefficients, the optimal wavelet for the defined discharge pulse is the one with the maximal value of γ.

For example, let us consider a DEP-type pulse with parameters $t_r = 5ns$, $t_w = 10ns$ and $t_d = 20ns$ and a DOP-type pulse with the same time parameters as above with an oscillation frequency $f_c = 20MHz$. Table 1 lists all of the calculated values of the correlation coefficient γ between the PD pulses and the available wavelets investigated, where all of wavelets are named using their abbreviations and the number behind indicates the order of the wavelet in its corresponding family. The results show that the db2 is the best wavelet for the DEP-type discharge pulse and the db7 for the DOP-type pulse. It can be seen from Figures 8a and 8b that the best wavelet based on the correlation coefficients appear remarkably similar to the pulse under analysis.

Does this technique apply to all other discharge pulses with different characteristic parameters? The authors believe that during the WT, the mother wavelet is scaled both in amplitude and in duration at different scales producing the best match result with the detected discharge pulse at a certain scale as long as a reasonable mother wavelet is chosen. Normally, the wavelets with lower order such as db2, db3, sym2, sym3, lem2 and lem3 are appropriate to the PD data measured with RC detection circuit, whilst for the PD data measured with RLC detection circuit, higher order wavelets of db6 ~ db10 are appropriate.

In practical use, the potentially candidate wavelets can be applied to the same problem under investigation simultaneously to determine the one which yields the best result.

5 PD PULSE EXTRACTION

The denoising of PD measurement data is closely associated with the feature extraction from the discharge pulse. Prior to an investigation of the PD activity of high voltage components in service, a reduction in the effect of noise may be a prerequisite to increase the signal-to-noise ratio. The noise appearing on-site includes corona, communicator operation, thyristor firing etc. generated inside electrical machines and/or externally induced interference from communication systems, e.g., radio signals from broadcasting stations etc. In terms of their nature, noise types prevalent on-site can be classified as sinusoidal noise such as that found from communication systems, pulse-type noise from thyristors etc. and stochastic noise, random both in amplitude and time occurrence, such as that associated with corona and electrical noise induced by the detection circuit itself. Over the years, methods employing filtering techniques [18], neural networks [19] etc., have been designed to suppress noise with some limited success. To date, only a few research papers [11,13] have been published using the wavelet analysis technique to deal with noise problems in PD studies. Angrisani et al. [11] dealt with the distortion in the PD pulse by external interference due to poor electrical contacts and random noise in an ultra wide-band measurement system to improve the recognition of the pulse's time parameters, where measurement data have a very high signal to noise ratio. Shim et al. [13] gave a general introduction on the relationship between wavelets and filters and wavelet based denoising.
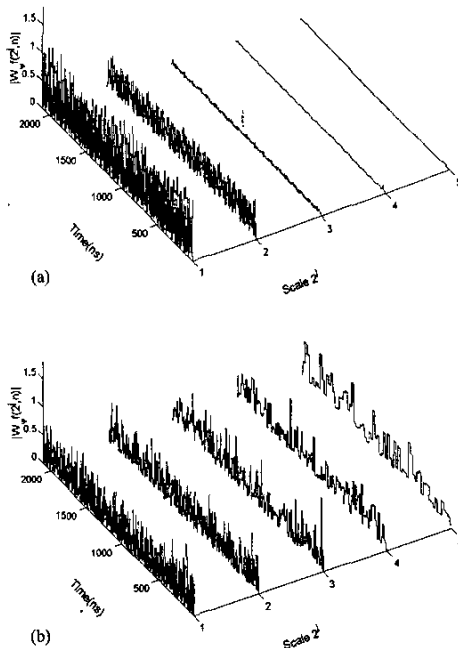
Figure 9. Wavelet patterns of the noise with the db2 wavelet. Sampling frequency, 1GHz; data record, 2100 points. a, sinusoidal noise by broad frequency range between 100 kHz and 200 MHz; b, random noise by white noise.

The present paper addresses two important and novel issues, which are missing from previous publications, related to wavelet based denoising methods, i.e., optimal wavelet selection and automated thresholding to accommodate the typical PD pulses and representative forms of noise with varying signal to noise ratio. The former has been addressed in the previous section, whereas this section presents an automatic level-dependent thresholding criterion which can be used to separate PDs from "noise" by means of characteristic patterns following the application of the DWT with the selected wavelets.

The coefficient distributions of two types of noise have been examined against a PD pulse. The first was continuous sinusoidal noise with very broad frequencies between 100 kHz and 200 MHz. Its wavelet pattern, as demonstrated in Figure 9a, decays asymptotically, in a significant way, with the increment of decomposition level. Secondly, the wavelet signature of stochastic noise given by white noise has the form shown in Figure 9b. In this case, the average density of coefficients is inversely proportional to the dyadic scale $1/2^j$ (j is the decomposition level), or, half the number of coefficient extremes do not propagate from one scale $1/2^j$ to a coarser scale $1/2^{j+1}$ [20]. In singularity terms in signal processing, its Lipschitz exponent is negative.

As described above, PD signals and noise present different wavelet patterns, thus enabling PD feature extraction. For a given set of PD data in the presence of noise, the denoising can be performed through wavelet decom-

position, thresholding of wavelet coefficients and wavelet reconstruction. Thresholding is of vital importance during denoising, which is conducted by retaining the wavelet coefficients associated with discharge events and, on the other hand, discarding all other discharge-free coefficients which characterize noise.

An automatic level-dependent thresholding criterion and hard thresholding approach is recommended to be applied to both details across decomposition levels and approximations at the user-predefined maximal level of a signal following the FDWT. The former at level j is defined as, $m_j/0.6745 * \sqrt{2*\log(n_j)}$, where m_j is the estimated level-dependent median value and n_j the level-dependent length of the signal. As can be seen from this thresholding estimator which considers the coefficient fluctuation at each decomposition level, both the basic threshold estimate $\sqrt{2*\log(n_j)}$ and the rescaling factor $m_j/0.6745$ are level dependent, thus facilitating the most effective suppression of indeterminate noise.

Hard thresholding processes the data in such a way that those wavelet coefficients whose absolute values are greater than the threshold are kept and those less than the threshold are replaced by zero, as can be expressed mathematically as $y = x *(| x | > \lambda)$, where x, y are original and retained coefficients respectively and λ the threshold value. Unlike the hard thresholding, soft thresholding shrinks the kept coefficients as can be seen from the thresholding rule $y = sign(x)*(| x | - \lambda)$. Hard thresholding was preferred by the authors in PD denoising due to the higher coefficient values associated with discharge events being kept thus yielding an improved PD signal to noise ratio in comparison with software thresholding. More information on the thresholding rules can be referenced in [6,21].

Consider now the efficiency of wavelet analysis in the denoising of signals containing partial discharges from some examples. Two sets of PD pulse sequences, one with 5 individual DEP-type pulses and the other with 5 DOP-type pulses, were immersed in electrical noise constituting various forms of electromagnetic radiation (both sinusoidal and stochastic), and recovery implemented. As shown from Figure 10, the db2 wavelet yields an effective denoising result when adopted in the DEP-type PD extraction. The frequency spectra of such DEP-type pulses and the superimposed noise are illustrated in Figures 10b, and 10c, respectively. The correlation coefficient γ between the separated PD signal and the original one has a value of 0.890, indicating the great similarity between them in nature, and the relative error in amplitude between them is within only 2%. However, for the DOP-type PD extraction, db7 wavelet yields a better denoising result, as shown in Figure 11 as an example. By observing Figures 11b and 11c, it can be seen that the frequency spectra of DOP-type pulses and the superimposed noise partially overlap around the resonant frequency of such DOP-type
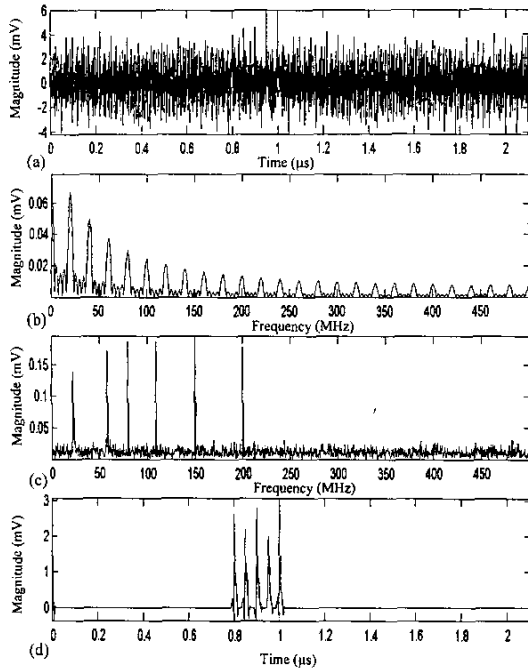
Figure 10. Denoising of a DEP-type PD event immersed in noise in the case of UWB measurement simulation. Sampling frequency, 1 GHz; data record, 2100 points. a, original noisy signal; b, frequency spectrum of PD pulses; c, frequency spectrum of the superimposed noise; d, PD pulses extracted.

PD pulses: $t_r = 5ns$, $t_w = 10ns$, $t_d = 20ns$, $50ns$ time separation; noise, combined by sinusoidal noise in broad frequency range between 100 kHz and 200 MHz and white noise.

Figure 11. Denoising of a DOP-type PD event immersed in noise in the case of narrow-band measurement simulation. Sampling frequency, 10 MHz; data record, 4200 points. a, original noisy signal; b, frequency spectrum of PD pulses; c, frequency spectrum of the superimposed noise; d, PD pulses extracted.

PD pulses: $t_r = 0.07 \mu s$, $t_w = 0.5 \mu s$, $t_d = 1 \mu s$, $20 \mu s$ time separation; noise, combined by sinusoidal noise in frequency range between 50 kHz and 1 MHz and white noise.

pulses. However, this kind of resonant PD pulse can still be separated from the combined electrical noise using the proposed denoising method. In this case, the denoising result shows the value of γ is high at approximately 0.95, and the relative error in amplitude is down to 4%.

Simulation studies showed that the denoising of both DEP-type and DOP-type PD pulses for narrowband sinusoidal noise is more efficient than that for wideband stochastic noise. As the DEP-type pulse has a wideband frequency spectrum, when the narrowband sinusoidal noise is superimposed on the DEP-type pulses, they can be identified from the coarsest approximations at maximal level and the details along decomposition levels which represent the specific frequency band involved in a signal. In the case of the narrowband DOP-type pulses superimposed by narrowband sinusoidal noise, the local irregularities of the signal are dominated by PD data when sampled with a suitable resolution, which can be detected from the details at lower levels following the FDWT. However, the denoising quality deteriorates gradually with the increment of the frequency spectrum range that the PD signals and noise overlap. This method is less efficient in suppressing stochastic noise from both types of PD pulses when the signal to noise ratio is less than −3dB due to
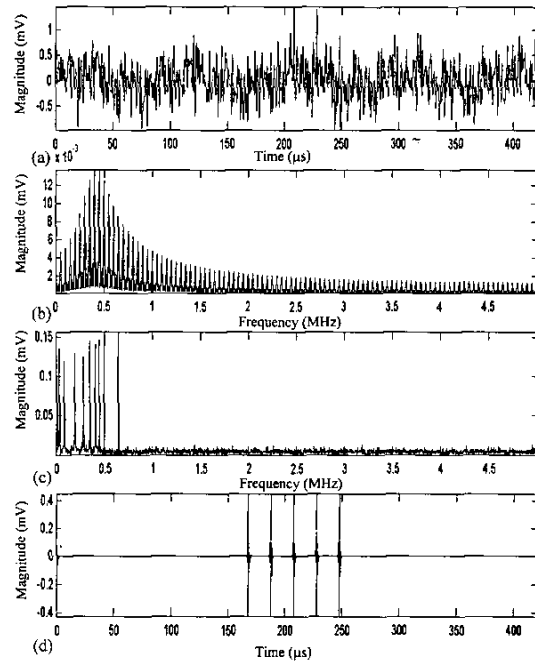
stochastic noise having a wideband frequency spectrum and a random occurrence in time.

It is worth underlining that the characteristic parameters of the PD pulse influence significantly the efficiency of PD pulse extraction from various forms of noise interference. The greater the damping coefficient and interval between pulses, the better the result in terms of PD event denoising. For the sampling frequency f_s utilized, the higher the f_s, the higher the wavelet decomposition level required in order to achieve the same efficiency for PD denoising, since a longer original sequence is available with the increment of f_s within a fixed time period.

Further evidence of the power of this technique is illustrated in Figure 12. In this case, partial discharge signals were acquired from a cable component energized at 5 kV in the presence of significant electronic noise and corona. The wavelet with higher order, db10, was chosen in this analysis, since an RLC detection circuit with a resonant frequency of 260 kHz was utilized in the measurements. By inspection of Figure 12b, three pulses appear at the time instants 0, 63 μs and 100 μs, respectively. By inspection of Figure 12a, the PD pulses at 0 and 100 μs might have been surmised, but clearly not that at 63 μs. Hence, the wavelet based denoising method is clearly successful in extracting the PD pulses from the noise.
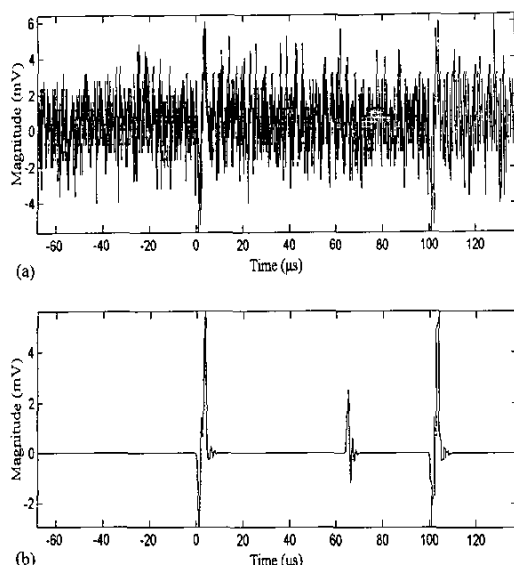
Figure 12. Denoising of PD pulses immersed in actual narrow-band electronic noise and corona. Sampling frequency, 10 MHz; data record, 2048 points. a, detected original PD data; b, separated PD pulses.

To further evaluate the performance of this denoising technique, the PD $H_n(\phi,q)$ patterns generated by using the original data without any processing, by using a threshold technique and the wavelet denoising technique are presented here for the test on the cable component. The sampling frequency utilized in this test was 2.5 MHz obtaining 5×10^4 points for each data record with respect to a complete ac power cycle of 50 Hz. Each PD pattern was generated with 100 data records as acquired by the above narrow-band detection system during a measurement duration of 10 min. The PD pattern shown in Figure 13a is dominated by the noise where the original data are directly used to produce the pattern. Figure 13b gives the PD pattern by using a threshold technique applied to the original data where the noise level is estimated as 2 times of the standard deviation value of the acquired data. Figure 13c gives the PD pattern based on the wavelet-based noise suppression. It is evident the wavelet denoising technique not only has eliminated the noise but also detected more PD activity with low apparent discharge magnitude, and is thus much more unambiguous and effective in characterizing the nature of the discharges in this particular PD geometry.

The time using a standard Pentium 3 to denoise a signal consisting of 2048 points illustrated in Figure 12 is 0.2 s. The time needed to denoise such a data stream of 5×10^4 in generating PD patterns is around 1.5 s, giving a good result potentially for on-line diagnostics or quality assurance of power equipment. Another attractive benefit coupled to this method is data compression of PD measurement data. Because only those coefficient data related to discharge events need to be used to precisely re-
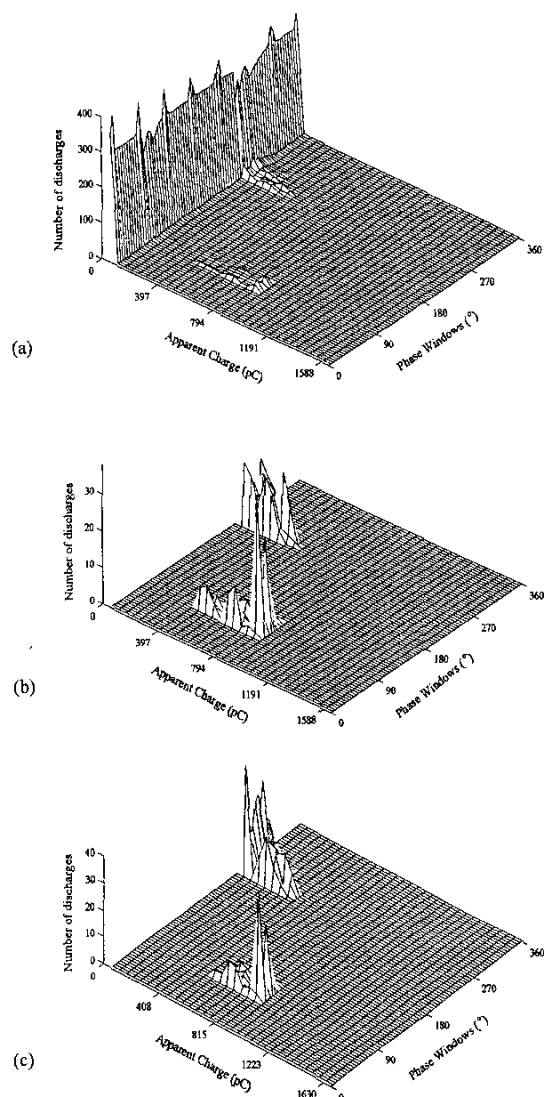

Figure 13. PD patterns of a cable component energized at 5kV. Sampling frequency, 2.5 MHz; measurement duration, 10 min. a, With original data without any processing; b, With a threshold method directly applied to original data; c, With wavelet-based noise suppression.

construct the actual PD signal extracted, the amount of data storage space can be greatly reduced. If the compression measure is evaluated in compression ratio C_r in percentage between the total retained coefficients across decomposition levels and the length of the original signal under analysis, C_r is found to be within 5%.

6 TIME-FREQUENCY ANALYSIS

An early paper by some of the present authors [22] reported the results of a pulse propagation mechanism along the winding in a 6.6kV star-connected induction motor which had 6 poles, 90 slots, and its rotor removed. A discharge simulating pulse of rise time 50 ns with a width of
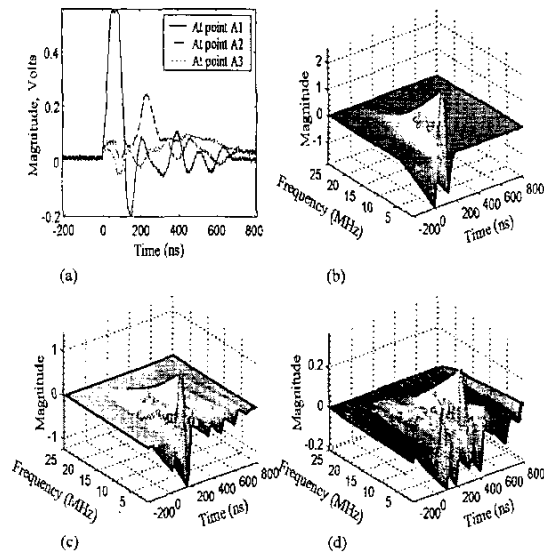
Figure 14. Time-frequency description of a discharge simulating pulse propagating along the winding. a, PD pulses measured at different points; b, wavelet pattern at point A_1 (a point on the end winding of the first turn of a coil); c, wavelet pattern at point A_2 (same point one turn away from A_1); d, wavelet pattern at point A_3 (same point two turns away from A_1).

twice the pulse rise time was injected at one point A_1 in the winding, yielding output responses at A_2, A_3 as shown in Figure 14a. The detailed experimental setup can be found in [22]. In this case, A_1 is a point on the end winding of the first turn of a coil, A_2 the same point one turn away from A_1, and A_3 two turns distant from A_1.

In the case of motors, generators and transformers, partial discharge detection and interpretation is often hampered by the distorting effect on the PD pulse due to the distributed impedance of the windings as the pulse propagates from its site of origin to the measurement terminals. However, its wavelet pattern of a signal created by the continuous wavelet transform (CWT) offers the possibility of improving the characterization of the pulse propagation and providing better insight into the mechanism of propagation since it contains unique and enhanced information compared with traditional methods in both the time and frequency domains [23].

As is well known, the factor a in the CWT, as described in equation (4), can be varied in either linear or dyadic scaling depending on the frequency component contained in the signal under examination and the required frequency resolution. Here, the scaling factor a is varied in linear increments so that the CWT can provide a finer wavelet pattern view through the entire time-frequency domain. In the CWT, the relationship between the scale and frequency can be understood in the following procedures. If a mother wavelet has a central frequency of $\omega_0/2\pi$ where ω_0 is defined as $\omega_0 = \int \omega |\hat{\psi}(\omega)|^2 d\omega$, the central frequency created by $\psi(at)$ as shown in equation

(3) becomes $\omega_0/(2\pi a)$. Furthermore, the frequency f_a at scale a corresponding to the sampling frequency f_s of the original signal is then given by $f_a = (f_s\omega_0)/(2\pi a)$.

Consider now the wavelet choice suitable for the time-frequency analysis of a PD pulse propagating along the winding. To detect the local edges or sharp transitions in a signal, which are defined as singularities, the wavelets with some vanishing moments are crucial. The wavelets with k vanishing moments satisfy the condition

$$\int_R t^j \psi(t)dt = 0, \tag{9}$$

where j is positive integer $j < k$. The modulus maxima of the CWT coefficients of $f(t)$ yielded with such $\psi(t)$ are given by

$$|W_\psi f(a,b)| \le A a^{\alpha(b)} \tag{10}$$

where A is a constant and α the Lipschitz exponent which is used to measure the singularity of a signal $f(t)$ at time instant b. The Lipschitz exponent a gives an indication of the differentiability of $f(t)$ more precisely, for example, the α of a Dirac signal is estimated as 0, and the α of a signal which is differentiable at time b is measured as 1. To estimate a signal with a Lipschitz exponent up to k, the wavelet selected should have k vanishing moments. However the number of CWT coefficient maxima at a given scale often increases linearly with the number of the vanishing moment of the wavelet [20]. Therefore, the wavelets suitable for the PD pulse propagation application are those with as few vanishing moments as possible but with enough moments to detect all sharp transitions that are of interest in this field.

The db2 wavelet with 2 vanishing moments is chosen here to illustrate the wavelet pattern of PD pulses detected at different points along the winding. Figures 14b to 14d display the wavelet patterns of the pulses shown in Figure 14a following the CWT with db2 wavelet. The energetic wavelet coefficients cover the active frequency components up to around 20 MHz and the time range of 0–600 ns. Clearly, the wavelet pattern can reveal explicitly not only the dominant frequencies involved in the signal under examination but also the time instants when they appear.

A "traveling wave" on the copper of the winding is dominantly present during the propagation of this pulse. The further the detection point from the injection point, the progressively wider in shape and lower in magnitude the pulse observed, since high frequency components are seriously attenuated when the pulse propagates from its origin site to the measurement point. Meanwhile, the "coupling waves" are found in the waveforms of A_2, A_3 points which appear before the "traveling wave" arrives due to the mutual capacitive interaction intensified by the extreme high frequency components involved in the PD pulses.
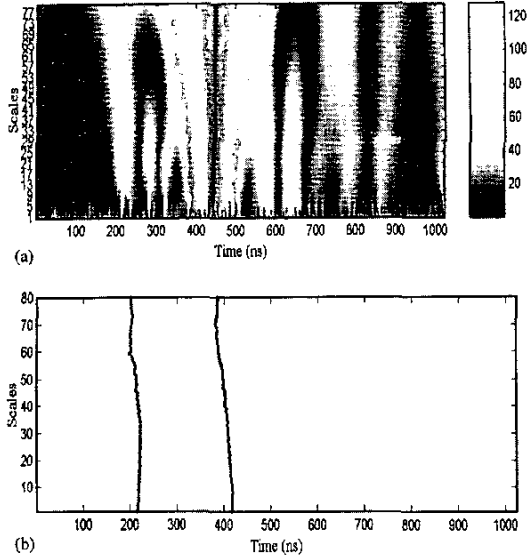
Figure 15. The coefficient distribution of the detected PD pulse and the associated local maxima lines used to infer the original site of PD pulse. Wavelet, db1; linear scaling, from 1 to 80; a, two-dimensional coefficient distribution of PD pulse at point A_2 in the wavelet domain; b, local maxima lines.

The combination of information contained in the "traveling wave" and "coupling wave" has been used for locating PD site in HV plant [24]. Does the wavelet pattern improve the time-flight method for inferring the original PD site? This will be addressed below in terms of the local maxima lines of the wavelet coefficients in the direction of the scale. The local maxima line is defined in this context as a continuous curve in the wavelet domain along which all points are modulus maxima [20]. The singularities in an irregular signal indicate the onset of the sharp transients, the timing of which is of interest for locating the transient onset especially resulted from the combination of the "traveling wave" and "coupling wave". In order to minimize the amount of computation, the db1 wavelet with only one vanishing moment is chosen as the analyzing wavelet for locating the singularities in time. Figure 15a illustrates a two-dimensional wavelet pattern of the PD pulse at point A_2 in the time-scale domain where the scales are selected within 80. Figure 15b shows the kept local maxima lines in the same wavelet domain by applying a predetermined threshold $\lambda = \rho \max(|W_{\psi}f(a,b)|)$ to truncate other local maxima lines whose amplitudes are lower than this threshold, where ρ is a turning factor with a value between 0 and 1. As can be seen from the local maxima lines, the "coupling wave" and the "traveling wave" appear at time instants 0.215 μs and 0.417 μs, respectively. The "coupling wave" arrives almost instantly due to a stray field formed in the dielectric media in which energy is coupled to the neighbouring conductors, the arrival time of which depends on the physical distance being divided by the approximate speed of light. The "traveling

Table 2. Arrival time of two propagation mechanisms and the inferred distance with respect to the pulse injection point.

Detection points	Arrival time of the coupling wave (μs)	Arrival time of the traveling wave (μs)	Calculated distance (m)	Electrical distances (m)
A2	0.215	0.417	2.424	2.42
A3	0.218	0.630	4.944	4.84

wave" arrives with a time delay depending on the electrical distance propagating along the winding divided by the velocity which is insulation material dependent. Therefore the arrival time difference between the "traveling wave" and the "coupling wave" can be used to calculate the distance between the detection point and the origin site of the PD pulse. Table 2 gives the arrival time of such two propagation modes in the waveforms acquired at points A_2, A_3 by which the original PD pulse site can be inferred. In this case, the velocity of the traveling wave in the induction motor winding is around 120m/μs, and the electrical distances (conductor length) of points A_2, A_3 with respect to pulse injection point A_1 are 2.42 m and 4.84 m respectively [22]. It demonstrates that the error in distance is within 3% but it increases with the distance of detection point, due to the amplitude decrement of the PD pulse along the winding.

7 CONCLUSIONS

WAVELET transforms can reveal enhanced and reliable information contained in partial discharge signals. The WT is capable of locating time and frequency components simultaneously allowing the analysis of signals with irregular and transition features, such as in the case of partial discharges. The scaling factor selected in the WT can be varied in either linear or dyadic scaling depending on the frequency component contained in the signal under examination and the required frequency resolution.

To use wavelet theory for PD analysis, it is vital to select the most appropriate wavelet and this is determined by the nature of the PD pulse. The correlation coefficient between the wavelet and the measured PD pulse can be used as an efficient criterion to evaluate the optimal wavelet for the measured PD pulse. For PD data acquired with a RC detector circuit, the optimal wavelets are those dyadic orthogonal wavelets of lower order, whereas for PD data obtained with a RLC detector circuit, those dyadic orthogonal wavelets of higher order are preferred.

Based on the patterns of the coefficient distribution of the PD pulses and noise following the WT, denoising problems can be solved. Further tests on practical PD data obtained on-site will confirm its power.

Because only a convolution calculation is involved in the implementation of the FDWT, the algorithm may be potentially applied for on-line PD detection in the presence of severe electrical disturbance. Furthermore, the possibil-

ities for data compression that accompanies the wavelet-based method alleviate the problems related to data storage and data communication during on-line PD detection.

The technique employing the local maxima lines of the PD pulse following the CWT provides the potential for locating the original site of a PD pulse propagating along the winding more accurately. However, studies have shown that this application depends on the wavelet choice, the scale range selected and the turning factor for thresholding, therefore it is application dependent and needs to be further justified.

The present effort of this paper has been devoted to finding wavelets from the extensive libraries of mother wavelets based on the clear understanding of the test specimen and the detection system utilized. Provided the behavior of the detection circuit being utilized is fully characterized and understood, as is the case in these studies, there should be no need to adopt an adaptive wavelet approach. However, if the PD data set is non-stationary in the measurement, an adaptive wavelet should be considered to automatically adjust the shape of the analyzing wavelet to the signal of interest. All of these show there is still a lot of work to be done in this area, but it holds very significant promise for the future.

ACKNOWLEDGMENTS

REFERENCES

[1] I. J. Kemp, "Partial Discharge Plant-monitoring Technology: Present and Future Developments", IEE Proc. Sci. Meas. Technol., Vol. 142, No. 1, pp. 4–10, 1995.

[2] F. H. Kreuger, E. Gulski and A. Krivda, "Classification of Partial Discharge", *IEEE Trans. EI*, Vol. 28, pp. 917–931, 1993.

[3] A. K. Louis, P. Maab, and A. Rieder, *"Wavelets: Theory and Applications"*, John Wiley & Sons, 1997.

[4] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis", IEEE Trans. Information Theory, Vol. 36, pp. 961–1005, 1990.

[5] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 11, pp. 674–693, 1989.

[6] M. Misiti, Y. Misiti, G. Oppenheim and J. Poggi, *"Wavelet Toolbox Manual - User's Guide"*, The Math Works Inc., USA, 1996.

[7] M. Mukai, T. Okano, S. Nishimoto, I. Kitani and K. Arii, "Study on Degradation Diagnosis of Partial Discharge in a Void by Wavelet Analysis", Proc. Symposium on Electrical Insulating Materials, Japan, pp. 463–466, 1995.

[8] S. Huang and C. Hsieh, "High-Impedance Fault Detection Utilizing a Morlet Wavelet Transform Approach", IEEE Trans. Power Delivery, Vol. 14, pp. 1401–1407, 1999.

[9] S. Huang and C. Hsieh, "Application of Morlet Wavelets to Supervise Power System Disturbances", IEEE Trans. Power Delivery, Vol. 14, pp. 235–243, 1999.

[10] T. B. Littler and D. J. Morrow, "Wavelets for the Analysis and Compression of Power System Disturbances", IEEE Trans. Power Delivery, Vol. 14, pp. 358–364, 1999.

[11] L. Angrisani, P. Daponte, G. Lupo, C. Petrarca and M. Vitelli, "Analysis of Ultrawide-band Detected Partial Discharges by means of a Multiresolution Digital Signal-processing Method", Measurement, Vol. 27, pp. 207–221, 2000.

[12] X. Ma, C. Zhou and I. J. Kemp, "Investigation into the Use of Wavelet Theory for Partial Discharge Pulse Extraction in Electrically Noisy Environments", Proc. 8th International Conference on Materials, Measurements and Applications, Edinburgh, UK, pp. 123–126, 2000.

[13] Shim, J. J. Soraghan and W. H. Siew, "Detection of PD Utilising Digital Signal Processing Methods, Part 3: Open Loop Noise Reduction", IEEE Electrical Insulation Magazine, Vol. 17, No. 1, pp. 6–13, 2001.

[14] P. Morshuis, "Assessment of Dielectric Degradation by Ultrawide-band PD Detection", IEEE Trans. DEI, Vol. 2, pp. 744–760, 1995.

[15] G. C. Stone, H. G. Sedding, N. Fujimoto and J. M. Braun, "Practical Implementation of Ultra wide band Partial Discharge Detectors", IEEE Trans. EI, Vol. 27, pp. 70–81, 1992.

[16] G. Zingales, "The Requirements of a PD Measuring System Analyzed in the Time Domain", IEEE Trans. DEI, Vol. 7, pp. 2–5, 2000.

[17] G. Keppel and S. Zedeck, *"Data Analysis for Research Designs — Analysis of Variance and Multiple Regression/Correlation Approaches"*, W. H. Freeman and Company, New York, 1989.

[18] U. Kopf and K. Feser, "Rejection of Narrow-band Noise and Repetitive Pulses in On-site PD Measurements", IEEE Trans. DEI, Vol. 2, pp. 1180–1191, 1995.

[19] H. Borsi, E. Gockenbach and D. Wenzel, "Separation of Partial Discharge from Pulse-shaped Noise Signals with the help of Neural Networks", IEE Proc. Sci. Meas. Technol., Vol. 142, pp. 69–74, 1995.

[20] S. Mallat and W. L. Hwang, "Singularity Detection and Processing with Wavelets", IEEE Trans. Information Theory, Vol. 38, pp. 617–643, 1992.

[21] D. L. Donoho, "De-noising by soft-thresholding", IEEE Trans. Information Theory, Vol. 141, pp. 613–627, 1995.

[22] C. Zhou, I. J. Kemp and M. Allan, "The PD Pulse Behavior in Rotating Machine Stator Windings", IEEE Conf. on Electrical Insulation and Dielectric Phenomena, Vol. 1, pp. 372–375, 1995.

[23] X. Ma, C. Zhou and I. J. Kemp, "PD Pulse Propagation Studies Using the Wavelet Transform", IEEE Conf. on Electrical Insulation and Dielectric Phenomena, Vol. 2, pp. 626–629, 2000.

[24] R. E. James, Q. Su, B. T. Phung, S. C. Foong and R. C. Tychsen, "Location of Partial Discharge on an 80MW/12.5kV Hydro-generator with the aid of Digital Filtering Techniques", J. Electrical and Electronics Engineering, Australia-IE Aust. and IREE Aust., Vol. 10, No. 4, pp. 338–343, 1990.