

REMOÇÃO DO RÚIDO GAUSSIANO E OUTLIERS PARA MEDIÇÃO DE DESCARGA PARCIAL EM EQUIPAMENTOS DE ALTA TENSÃO UTILIZANDO A ANÁLISE DE COMPONENTES INDEPENDENTES

WALTER M. H. CUENCA, ALAIN F. S. LEVY

*Departamento de Linhas e Estações (DLE)/ Divisão de Laboratórios de Adrianópolis (DVLA), CEPEL
Caixa Postal 68007- 21944-970, Rio de Janeiro, RJ, BRASIL
martin@cepel.br, alain@cepel.br*

SANDOVAL C. JUNIOR, JOSÉ M. SEIXAS

*Programa de Engenharia Elétrica, COPPE/UFRJ
Caixa Postal 68504 - 21945-970, Rio de Janeiro, RJ, Brasil
sandoval@dee.ufrj.br, seixas@lps.ufrj.br*

Resumo— A análise de sinais de descargas parciais (PD) tem sido utilizada como uma das principais técnicas para detectar defeitos nos sistemas de isolamento, especialmente em sistemas elétricos em alta tensão, tais como, geradores, hidrogeradores e transformadores de potência. Porém, a detecção e a medição de PD são ambas afetadas por ruídos externos e/ou internos que deformam e atenuam os sinais de PD. Os sinais de ruído são sinais indesejados e suas derivações são conhecidas como interferências, *artifacts*, *outliers* e *crosstalks*. No presente trabalho, é apresentado uma aproximação para o problema de remoção de ruído gaussiano e *outliers* em sinais de fontes de descargas parciais. Para isso, utilizam-se os conceitos de separação cega de fontes (BSS) junto à análise de componentes independentes (ICA) e do teste de Grubbs respectivamente. Neste contexto, os componentes independentes, obtidos por funções de cumulantes estatísticos de quarta ordem, conduzem para a supressão do ruído e a separação ótima das fontes úteis dos sinais interferentes. Com a ajuda do teste de Grubbs e a inspeção visual, por meio de gráficos de controle, é possível a detecção e remoção de *outliers* das observações. Os melhores resultados foram obtidos utilizando fontes de sinais de PD provenientes de oito modelos experimentais, cujas características representam alguns defeitos típicos similares aos encontrados em enrolamentos estáticos de hidrogeradores.

Palavras-chave— Remoção de Ruídos, Outliers, Descargas Parciais, Hidrogeradores, ICA/BSS, Cumulantes, teste de Grubbs.

1 Introdução

Descargas corona ou descargas parciais é um termo cuja base consolidou-se no começo dos anos de 1920, nessa época, já era considerado como um campo de pesquisa desenvolvido. Sua preeminente importância ferramental tem crescido para assistir na qualidade e no desempenho característico dos equipamentos de alta tensão. Dessa maneira, as descargas parciais têm se tornado alvo de estudos relacionados a fenômenos físicos e químicos dos materiais isolantes nos equipamentos. Com o passar dos anos, o estudo das descargas parciais tem-se diversificado substancialmente, aparecendo estudos que relacionam o tipo de equipamento de alta tensão com o tipo de descarga. Assim, outras tendências apareceram enquanto estudos relacionados à natureza e forma de descarga. Consequentemente, outras condições surgiram, como os casos de degradação da isolação exposta aos efeitos de PD, transferência de carga aparente, taxa de repetição de pulsos de descarga, distribuição de pulsos de descarga, épocas de descarga relativa a uma fase da tensão referencial, perdas de energia, distribuição de pulsos de descarga, intervalo de tempo entre pulsos assim como o efeito de memória [1],[2],[3]. Recentemente, o conceito de descargas parciais entrou no campo gerencial das empresas concessionárias de energia elétrica, onde se fala de indicadores de falha e de indicadores de qualidade gerencial. Paralelo a is-

so, o campo da manutenção de equipamentos elétricos de alta tensão incorpora a detecção, a medição e a monitoração de sinais de PD [4]. Nestes últimos anos, PD tem ganhado mais um aliado nos altos conceitos da manutenção centrada em confiabilidade (RCM) [5], focalizando a manutenção preditiva e a condição de monitoramento (CM) de isolamentos em enrolamentos estáticos em máquinas girantes e em sistemas de isolamento sólido/líquido em transformadores de potência [2]. Nestes novos cenários de manutenção, diagnosticam-se possíveis falhas com base nos indicadores de defeitos que o próprio equipamento fornece durante seu funcionamento. Além disso, diversos métodos, técnicas e práticas têm sido utilizados com o objetivo de determinar os indicadores mais precisos de defeito. Para os casos de equipamentos de alta tensão, por exemplo, em hidrogeradores, a avaliação de PD tem se mostrado um excelente indicador de falhas futuras, se é detectado no seu estado inicial [3]. Por outro lado, os sistemas de medição e detecção de PD são afetados por ruídos, atenuando e/ou deformando os sinais de PD. O ruído, cronologicamente, tem estado presente inerentemente em sinais de PD. Para isso, diversos estudos para remoção e/ou eliminação de ruído têm sido propostos, muitos deles baseados na filtragem analógica e digital [1],[6]. As normas IEEE P1434 e IEC 60270 contemplam circuitos para eliminar algumas fontes de ruído provenientes do circuito de medição [7], [8]. Novos

instrumentos de medição de PD, baseado na aquisição digital dos sinais de PD, apresentam uma taxa de amostragem suficientemente alta para evitar efeitos de *aliasing* e assim podem coletar um número suficiente de pulsos, os quais imediatamente são tratados por módulo de análise estocástica [2],[3]. Além disso, o instrumento digital, dotado de um *trigger* partilhado e memória-buffer on-line, torna possível um melhor armazenamento das amostras de um determinado bloco de memória, antes que estas sejam substituídas por outros novos pulsos de DP. Neste tipo de instrumento é possível fazer a separação de formas de pulsos no tempo e na frequência, permitindo a separação de um conjunto de pulsos de interesse. Segundo [2], é possível assumir diferentes tipos de fontes de PD que podem produzir uma rajada de pulsos com diferentes características estocásticas, assim é possível identificar as fontes de PD.

Outros estudos, encontrados na literatura, são baseados nas técnicas de inteligência artificial. Redes neurais, lógica *fuzzy*, algoritmos genéticos, *simulated annealing*, componentes principais, curvas principais e cadeias markovianas têm sido utilizadas para fins de separação e classificação de fontes de DP [4]. O ruído não é de vez eliminado e sim é reduzido para uma condição ótima, pois pode ser introduzido nas fontes pelos sistemas de medição ou, em alguns casos, pela semelhança com os sinais de PD.

Neste artigo, apresentamos uma nova aproximação para a remoção de ruído gaussiano em fontes de sinais de PD. Para tal efeito, utilizamos a separação cega de fontes (BSS, no inglês) [10]. Um número de algoritmos baseados na teoria ICA, Análise de Componentes Independentes, vem fornecendo soluções aos problemas BSS, sendo aplicados com sucesso em diversos campos, tais como em sistemas de comunicação, processamento de sinais biomédicos e processamento de sinais de PD [9],[10], além de começaram a ser usados em técnicas de filtragem de imagens como, por exemplo, em *wavelets shrinkage* e *denoise* [6].

2 Modelos Experimentais

As técnicas de detecção dos sinais de PD são baseadas em vários fenômenos produzidos pela descarga como luminescência, calor, ruído audível, transformações químicas, pulsos de descargas elétricas, perdas de energia e campo eletromagnético, os quais são descritos detalhadamente em [1]. Neste trabalho, utilizamos, para ensaios práticos, um instrumento de medida de PD ICM, que emprega a técnica de medição da carga aparente, baseada no método elétrico, segundo as normas IEC-60270 e IEEE-P1434. Este equipamento tem a particularidade de relacionar o ângulo de fase da tensão de teste, a amplitude de des-

carga e taxa de ocorrência ou repetição na forma de um mapa bidimensional das assinaturas dos sinais PD. A Figura 1 mostra um mapa de PD $\Gamma(\phi, q, N)$, onde os pulsos de descarga $H_N(\phi)$ são referenciados angularmente sob a tensão aplicada, as amplitudes $H_{q_{max}}(\phi)$ são os picos máximos dos pulsos de descarga detectados em janelas ajustáveis de registro (5 a 600 μ s) e a taxa de ocorrência codificada em 16 cores normalizadas. Tais mapas digitais são definidos como PRPHA – *Phase Resolved Pulse Height Analyzer* [3].

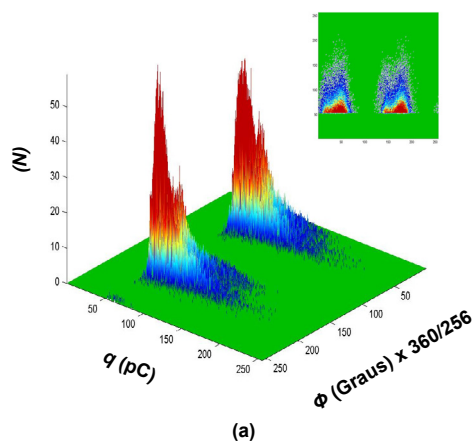


Figura 1. Histograma com distribuição bivariada. Mapa típico de PD com informações associadas à fase (a).

2.1 Modelos baseados em barras de prova

Neste trabalho, os dados de PD, para o estudo em hidrogeradores, são procedentes de ensaios rotineiros praticados na Usina de Itaipu. Na maioria dos casos, utilizam-se barras estatóricas envelhecidas para efeito de prova e detecção de defeitos. Cada barra foi preparada com dimensões de aproximadamente um metro de comprimento, na qual foram provocados defeitos característicos e típicos, similares aos encontrados no campo (ver Figura 2). A finalidade deste procedimento é reproduzir fontes de PD localizadas, de modo que cada barra de prova percorra exclusivamente as espiras do enrolamento estatórico, cujos terminais se encontrem disponíveis para medição [4].

Defeitos em barras como descargas de ranhura, corona (corona in *end-winding*) e delaminação são as fontes de defeitos mais comumente encontradas na literatura [1].



Figura 2: Modelo experimental e montagem típica da barra de prova sob o enrolamento estatórico de hidrogeradores [11].

Algumas variantes foram realizadas para conseguir representar até oito fontes diferentes testes. Destas, quatro são fontes características (*) e o restante é praticado nas outras fases do estator, com influência de tiristores e corona na cabeça de bobina, segundo a Tabela 1.

Tabela 1: Amostras de Fontes de Sinais de PD como em [4].

Fontes	Defeitos em Barras (10,5kV) Sistema: capacitor de acoplamento/ barra.
S1	Barra 0 (corona - barra pivô)
S2	Barra 1 (slot)*
S3	Barra 2 (corona)*
S4	Barra 2 (corona na cabeça da bobina)
S5	Barra 3 (delaminação)*
S6	tiristor fase B*
S7	tiristor fase C
S8	tiristor neutro

3 Representação de DP em ICA

Os ruídos introduzidos nos sistemas de medição e a elevada dimensão do espaço de dados coletados são dois grandes obstáculos, não só para caracterizar sinais-fonte de PD, mas também para desenvolver técnicas eficientes de reconhecimento de padrões em PD. Os dados provenientes das medições de PD no campo, na maioria dos casos, contêm ruído e/ou *outliers*. Técnicas de filtragem buscam remover ou minimizar algumas fontes de ruído. Geralmente estas técnicas atuam na fase de detecção e aquisição dos sinais, podendo rejeitar ou aceitar o sinal. Neste contexto, também há sinais PD de pequena amplitude e que, muitas vezes, são rejeitados e confundidos com ruído de fundo de escala do instrumento de medição. Além disso, existem outros fatores que influem diretamente na medição, por exemplo, a sensibilidade do instrumento, dos sensores e até a influência dos conversores A/D e D/A [4],[6].

Contudo, novas técnicas, como ICA e filtros digitais sofisticados, estão sendo introduzidas por meio de processadores digitais de sinais (DSP) para assistir nas medições de sinais de PD [2],[4],[6].

3.1 Modelo básico - ICA

A análise de componentes independentes busca mapear os dados de forma a encontrar fontes de sinais estatisticamente independentes [10]. Essa técnica tem a capacidade de extrair sinais-fonte os mais independentes possíveis, a partir de um conjunto de sinais misturados, denominados observações, sem o conhecimento a priori das fontes originais. No caso as fontes recuperadas não têm uma ordem definida. Por

outro lado, não é possível determinar uma escala para as fontes recuperadas.

Cada fonte de observação de PD pode ser disposta como uma matriz reduzida ou vista na forma de vetor de amostras $\mathbf{x}(k)$ de tal forma que

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \nu(k) \quad (1)$$

onde k indica a k -ésima amostra. Neste caso, \mathbf{A} é chamada de “matriz de mistura”. É uma matriz $n \times m$ dimensional, não singular, desconhecida e invariante à transformações. O vetor de ruído $\nu(k)$ é assumido ser independente com $\mathbf{s}(k)$, de média nula e variância ν_{ij} e $\mathbf{s}(k)$ é o vetor de m fontes de sinais estatisticamente independentes e desconhecidos.

O objetivo de ICA é estimar a matriz \mathbf{A} ou sua inversa \mathbf{W}_k , também chamada de “matriz de separação”, expressa por,

$$\mathbf{y}_{(k+1)} = \mathbf{W}_k \mathbf{x}_{(k+1)}. \quad (2)$$

onde $\mathbf{W}_k = \mathbf{A}_k^{-1} + \varepsilon \mathbf{A}^{-1} \mathbf{A}^{-T} \mathbf{A}^{-1}$, ε é um valor real. As equações (1),(2) constituem o modelo básico da BSS [11], sendo $\mathbf{s}(k)$, \mathbf{A} , \mathbf{W}_k e $\nu(k)$ desconhecidos, ver a Figura 3.

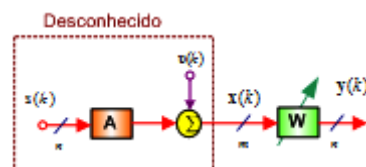


Figura 3: Modelo básico do BSS com componentes de ruído.

4 ICA para remoção de ruídos

O modelo ICA utilizado neste trabalho foi desenvolvido por Akuzawa [11] para resolver o problema BSS. Este modelo de Akuzawa também utiliza o método do gradiente natural - desenvolvido por Amari [9] para calcular a direção ótima quando é aplicada uma função de contraste - para resolver o problema de minimização dos cumulantes de quarta ordem em espaços curvos ou Reimanianos. O alvo deste modelo é extinguir os cumulantes de quarta ordem a fim de que as fontes estimadas sejam estatisticamente independentes entre si, mesmo contendo as fontes de ruído. Por outro lado, este método ICA é robusto sob ruídos gaussianos, uma vez que não utiliza estatísticas de segunda ordem. Outra característica deste método é sua estabilidade global quando são introduzidas deformações homotópicas no espaço de matrizes não singulares GL [11]. Essas deformações são introduzidas a cada atualização da “matriz de separação” e das variáveis (Equação 3), expressas por funções cumulantes, sob o gradiente natural. A “matriz de mistura” será sempre não singular, já que é desenvolvida no espaço homogêneo das matrizes, ou seja, numa curva

manifold (um espaço topológico) chamado *Riemannian space* [11]. A métrica, do espaço paramétrico é não Euclidiana e não ortogonal, neste caso, esta característica faz com que o método seja matematicamente livre de singularidades.

Outra qualidade deste método ICA é que não requer pré-branqueamento, ou seja, preserva a metade dos graus de liberdade que são utilizados nos métodos que precisam de branqueamento.

Os cumulantes de quarta ordem computados por este método são expressos para as variáveis estocásticas dos sinais de PD. Denotando a média amostral por $\langle \rangle$ e os cumulantes $\langle \rangle_c$ são estimados a partir dos dados observados $\mathbf{x}(k)$. Assim, funções contraste são representadas por $Q_p(y_i, y_j) = \langle y_i^p y_j^{4-p} \rangle_c$, onde p é o número de critérios assumidos. A matriz gradiente, deformada por ξ , e representada por

$$V = \begin{bmatrix} K_i & (3-\xi)R_{ij} \\ (3-\xi)R_{ij} & K_j \\ 2Q_{ij} & 2Q_{ji} \end{bmatrix} \quad (3)$$

onde, $K_i = \langle y_i^4 \rangle_c$, $K_j = \langle y_j^4 \rangle_c$ e $R_{ij} = \langle y_i^2 y_j^2 \rangle_c$,

4.1 Métricas de Avaliação

Dois medições são utilizadas para avaliar o desempenho do modelo ICA, a relação sinal-ruído (SNR) e os *crossstalks*. A qualidade da informação e do sinal reconstruído e é avaliada pelas SNRs. Já os *crossstalks*, são interferências ou acoplamentos mútuos das observações em aquisição de multicanal, por exemplo, em sinais de PD [4]. Uma medição do *crossstalk* remanescente SNR, indica o resíduo.

Estes indicadores são obtidos da equação $SNR = -10\log(v_v/v_x)$, onde v_v e v_x são as variâncias de v e x , respectivamente. E

$$CT_{1 \leq N \leq 8} = SNR(\mathbf{W} \times \mathbf{A}_k), \quad (4)$$

onde CT é o *crossstalk* do produto interno de $\mathbf{W} \times \mathbf{A}_k$.

5 Remoção de Outliers

Frequentemente, os dados das observações de PD estão sujeitos a vários tipos de observações enganosas conhecidas como *outliers* ou observações discordantes. As observações podem ser *outliers* por causa de um único valor grande ou pequeno de uma variável ou por causa de uma combinação incomum de valores de duas ou mais variáveis [12].

O algoritmo baseado em estatísticas de ordem elevada é tipicamente sensível a *outliers*. Mais especificamente, o impacto de *outliers* no desempenho de separação de sinais de um algoritmo ICA deve ser amortecido. Para isso, as observações devem passar

previamente por testes estatísticos a fim de detectar *outliers*. Entre os testes mais usuais, se destacam o teste de Dixon, que usa relações de faixas do dado, e o teste de Grubbs, que aplica relações de duas somas de quadrados [13]. Ambos os testes são baseados em uma suposição de erros normalmente distribuídos. Outro modo prático, para identificar *outliers*, é a inspeção visual das observações via *scatter plots* e *box plots*. O modelo ICA pode ainda utilizar funções de contraste que sejam mais imunes a *outliers*, como a função *Beta-Divergence*, ($0 \leq \beta \leq 1$) [12], funções não lineares, como a tangente hiperbólica combinada com a função gaussiana, função de potência três (*pow3*) e *skewness*. Por outro lado, é possível minimizar as influências de *outliers*, utilizando matrizes de rotação entre zero e 90 graus, fazendo rotação ótima que minimize o índice de desempenho de separação de fontes de Amari (API) [12].

6 Resultados Experimentais

Os sinais coletados na medição de PD representam os dados observados, naturalmente compostos pelo sinal-fonte contaminado por ruído. Esta coleção de dados foi obtida a partir das medições rotineiras em alguns hidrogeradores da Usina de Itaipu [4]. Um pré-processamento, baseado no teste de Grubbs [13], foi adotado com a finalidade de detectar e remover *outliers* (ver Figura 4).

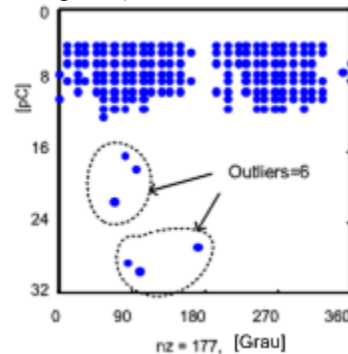


Figura 4: *Outliers* via *scatter plot* e teste de Grubbs.

A detecção foi realizada por teste de probabilidade P . Se maior do que 5% de significância, o item testado é aceito como não disperso; entre 1% e 5% de significância é declarado como estranho ou suspeito e P menor do que 1%, o item é declarado *outlier*.

A Figura 5 mostra oito sinais fonte de acordo com o que foi apresentado na Tabela 1. Uma fonte de ruído gaussiano foi introduzida em cada sinal-fonte para verificar a eficácia do ICA, resultando numa relação sinal-ruído SNR igual a 4,34 dB.

O desempenho do ICA foi medido pelos *crossstalks*. Foi realizada uma bateria de testes, dos quais sete eventos são resumidos na Tabela 2. O modelo de Akusawa [11] é aplicado na mistura ruidosa. Os sete

melhores resultados, qualificados como ótimos, são os valores dos *crostalks* médios CT tal como mostra a primeira coluna da Tabela 2 e os ganhos obtidos são melhores que o estado original das observações.

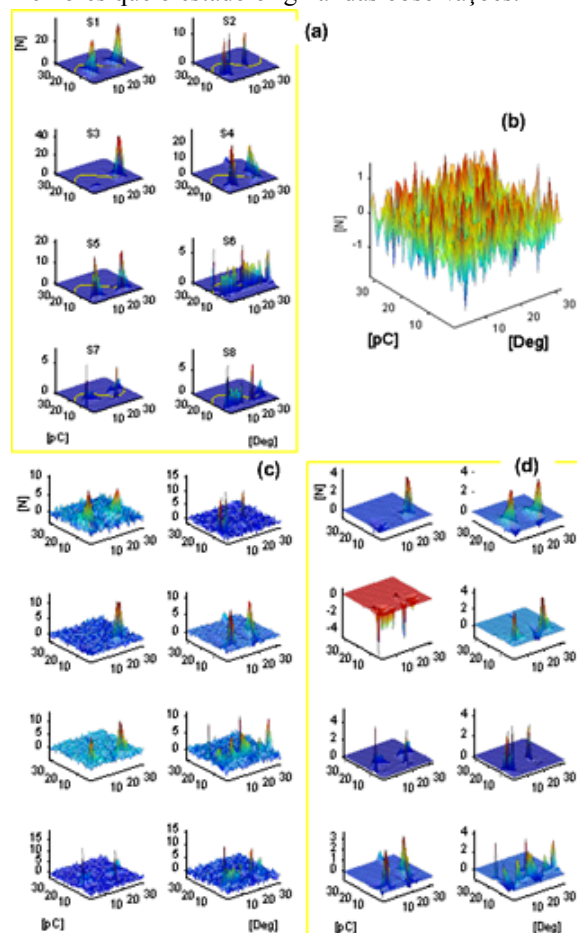


Figura 5: Resultado do modelo ICA. Sinais fonte (SNR:22,20 dB) (a); ruído gaussiano (b); observações (SRN: 4,34 dB) (c) e fontes estimadas (SRN:28,15 dB) (d) .

Tabela 2: Medida de desempenho por *crostalks*

CT	Max. $\langle \rangle_c$	SNR (dB)	Noise (%)	$ Gain $ (dB)
2,15	1,03 {3}	Inf	--	5,03
0,29	1,86 {6}	41,13	1	8,55
0,30	7,87 {7}	20,12	10	2,25
0,33	5,15 {8}	16,84	15	22,08
0,51	2,67 {7}	7,33	30	24,44
0,43	2,67 {7}	6,29	40	25,43
0,31	0,40 {7}	4,34	67	28,15

Numa segunda etapa foi realizada uma comparação do método de Akusawa com outros dois métodos clássicos de ICA, FastICA [10] e JADE [9]. Os resultados desta comparação são mostrados na Tabela 3, onde o método de Akusawa consegue os melhores resultados.

Tabela 3: Quadro comparativo de três métodos ICA

	Akusawa	JADE	FastICA
Error (dB)	1,23	2,24	2,31
Sum $\langle y \rangle_c$	0,28	26,22	25,65
Sum $\langle s \rangle_c$	22,21	22,21	22,21

7 Conclusões

O resultado da pesquisa mostra a eficácia e excelência do método ICA (de Akuzawa) para lidar com sinais de PD, utilizado para a remoção de ruído gaussiano; associado como o teste de Grubs, empregado como detector e eliminador de *outliers*. Se comparado com os melhores métodos clássicos de ICA, O método de Akuzawa teve os melhores resultados. Embora existam métodos clássicos de filtragem, cujos resultados revelados em [6] são considerados satisfatórios, porém estes apresentam diversas restrições como, a necessidade de ter conhecimento da forma do sinal de PD, dificuldade de escolher limiares, dependência um comportamento estacionário do sinal de PD e entre outras. Enquanto o método de Akuzawa unicamente não conserva a escala para as fontes recuperadas devido à natureza do sistema BSS, mas não afeta a característica e forma de PD.

Agradecimentos

Os autores agradecem ao CEPEL, CNPq e FAPERG COPPE/UFRJ pelo apoio financeiro.

Referências Bibliográficas

- [1] R. Bartnikas, "Partial Discharge Their Mechanism, Detection and Measurement", *IEEE Trans. Elec. Insul.* Vol. 9, no. 5, pp. 763-808, Oct., 2002
- [2] A. Cavallini and G. C. Montanari, "A New Approach to the Diagnosis of Solid Insulation Systems Based on PD Signal Inference," *IEEE DEIS*, Vol. 19, no. 2, pp 23-31, March/April 2003.
- [3] S.Kumar, "Development of a phase Resolved Pulse Height Analyser and its Application to Partial Discharge Measurements", MSc. Dissertation, Bangalor, India, Feb. 1997.
- [4] W. M. Huaman, "Caracterização dos sinais de Descargas Parciais em Equipamentos de Alta Tensão a partir de Modelos Experimentais". DSc. Dissertation, Dept. Electric Engineering, COPPE/UFRJ, Brazil, 2005.
- [5] J. Moubray, *Reliability-centred Maintenance* (RCM II). -2 Ed., Lutterworth, Inglaterra: Aladon Ltd, 2000.
- [6] F.C.C.B. Diniz, "Supressão de Ruído, Detecção e Classificação de Sinais de Descargas Parciais em Transformadores de Potência". MSc. Dissertation, Dept. Electric Engineering, COPPE/UFRJ, Brazil, 2005.
- [7] *IEC Publication-60270, High-voltage Test Technique-Partial Discharge Measurement*, Third Edition, 2000.
- [8] *IEEE P1434, Guide to the Measurement of Partial Discharges in Rotating Machinery*, Electrical Machinery Committee of the IEEE Power Society, 1997.
- [9] A. Cichocki, & S. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*. John Wiley & Sons, Ltd. New york, 2002
- [10] A. Hyvärinen, J. Karhunen & E. Oja., *Independent Component Analysis*, John Wiley & Sons. Inc.2001.
- [11] T. Akuzawa, N. Murata, "Multiplicative Non-holonomic Newton like algorithm," *Chaos, Solutions & Fractals*, vol. 12, pp. 785-793, 2001.
- [12] N. Gadhok, W. Kinsner, "Rotation Sensitivity of Independent component Analysis to Outliers," 0-7803-8886-0/05/\$20.00, IEEE, CCECE/CCGEI, Saskatoon, May 2005.
- [13] J. A. C. Palma, "Estimação dos Parâmetros de Séries temporais em Presença de Outliers", Conferência de 10 anos da EST/Setúbal, 15 Abril 1999.