



**UNIVERSIDADE FEDERAL DO PARÁ.
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS.
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA E ESTATÍSTICA.**

Eder Sarah de Lucena

**ANÁLISE DA INFLUÊNCIA DO CICLO HIDROLÓGICO NO
COMPORTAMENTO DAS DESCARGAS PARCIAIS EM
HIDROGERADORES**

Belém-PA
2017

**ANÁLISE DA INFLUÊNCIA DO CICLO HIDROLÓGICO NO
COMPORTAMENTO DAS DESCARGAS PARCIAIS EM
HIDROGERADORES**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Matemática e Estatística na Universidade Federal do Pará (PPGME-UFPA) como requisito parcial à obtenção do título de Mestre em Estatística.

Orientadora:

Dra. Terezinha Ferreira de Oliveira

Co-orientadora:

Dra. Marinalva Cardoso Maciel

Belém-PA
2017

Dados Internacionais de Catalogação-na-Publicação (CIP)
Biblioteca de Pós-Graduação do ICEN/UFPA

Lucena, Eder Sarah de

Análise da influência do ciclo hidrológico no comportamento das descargas parciais em hidrogeradores/ Eder Sarah de Lucena; orientador, Terezinha Ferreira de Oliveira .-2017.

114f. il. 29 cm

Inclui bibliografias

Dissertação (Mestrado) – Universidade Federal do Pará, Instituto de Ciências Exatas e Naturais, Programa de Pós-Graduação em Estatística, Belém, 2017.

1. Modelos lineares (Estatística). 2. Geradores hidrelétricos-Vida útil. 3. Descargas elétricas. 4. Ciclo hidrológico. 5. Umidade. I. Oliveira, Terezinha Ferreira de, orient. II. Título.

CDD – 22 ed. 519.72

Eder Sarah de Lucena

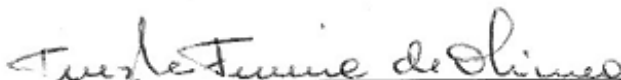
**ANÁLISE DA INFLUÊNCIA DO CICLO HIDROLÓGICO NO
COMPORTAMENTO DAS DESCARGAS PARCIAIS EM
HIDROGERADORES**

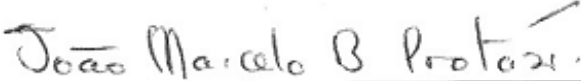
Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Matemática e Estatística na Universidade Federal do Pará (PPGME-UFPA) como requisito parcial à obtenção do título de Mestre em Estatística.


Belém, 7 de dezembro de 2017.

Resultado: aprovado

BANCA EXAMINADORA


Prof. Dra. Terezinha Ferreira de Oliveira (Orientadora)
PPGME/ UFPA


Prof. Dr João Marcelo Protázio (Membro Interno)
PPGME/ UFPA


Prof. Dr. Fernando de Souza Brasil (Membro Externo)
Devry/ FACI/ Eletronorte

AGRADECIMENTOS

A Deus, por sempre me abençoar e me dar forças para sobrepujar as dificuldades advindas do cotidiano.

À minha esposa que sempre me incentivou a buscar novos conhecimentos e acreditou que a conclusão de mais essa etapa de minha vida seria provida de sucesso.

Às minhas orientadoras Terezinha Ferreira de Oliveira e Marinalva Cardoso Maciel pela paciência, pelas valiosas orientações que nortearam esta pesquisa, pela credibilidade que me foi concedida.

À Centrais Elétricas do Norte do Brasil, à equipe do Dr. Fernando de Souza Brasil pela disponibilização dos dados e informações técnicas que contribuíram sobremaneira para o êxito deste trabalho.

Dedico este trabalho à minha amada esposa Cibele Lucena que sempre me encorajou a vencer os diversos obstáculos que enfrentei e sempre esteve ao meu lado em todos os momentos, abdicando de sua individualidade em prol da minha pos-graduação, ao meu pai Lucena e minha mãe Suely, que muito me ajudaram e nunca mediram esforços para me proporcionar as melhores condições de estudos e de vida.

Às minhas filhas Clara Inês e Cecília Lucena que fizeram parte desta conquista, iluminando de maneira especial os meus pensamentos, me motivando a buscar novos conhecimentos.

Eder Lucena

RESUMO

Este trabalho propõe um método para a análise de descargas parciais em hidrogeradores por meio dos modelos lineares generalizados considerando o ciclo hidrológico da Usina Hidrelétrica de Tucuruí com a variável resposta, ocorrência de descarga, obtida por dois acopladores capacitivos conectados à fase V de quatro medições. Os dados foram organizados por classes de 90° referentes aos ângulos e em classes de amplitude: menores de -100 mV, entre -100mV e 100mV e acima de 100mV e foram criadas três variáveis dummies. A primeira para tentar captar se a ocorrência aconteceu na classe de -100mV a 100mV. A segunda para captar se a ocorrência da descarga parcial aconteceu no período de cheia ou de seca. E a terceira variável dummy para captar a ocorrência de descargas no período de cheia depois da intervenção dos analistas. Nos três modelos avaliados, na cheia, a ocorrência de DP na classe de amplitude de -100 a 100 mV aumenta e no período de seca ocorre uma redução na quantidade de descargas e mesmo não considerando as ocorrências de DP nessa classe de amplitude, o efeito da cheia é significativo.

Palavras-Chave: descargas parciais, umidade, Modelos Lineares Generalizados, modelo Binomial Negativo.

ABSTRACT

This paper proposes a method for the analysis of partial discharges in hydrogenerators through generalized linear models considering the hydrological cycle (full and dry) of the Tucuruí Hydroelectric Power Plant with the response variable, discharge occurrence, obtained by two capacitive couplers connected to phase V of four measurements. The data were organized by classes of 90° referring to the angles and in classes of amplitude: less than -100 mV, between -100mV and 100mV and above 100mV and three dummy variables were created. The first, to capture occurrence in the class from -100 mV to 100mV. The second, to capture the occurrence of the partial discharge occurred during the full or dry season. In addition, the third dummy variable to capture the occurrence of discharges in the full season after analysts' intervention. In the three models evaluated, in the full season, the occurrence of partial discharges (PD) in the amplitude class of -100 to 100mV increases and in the dry season there is a reduction in the amount of discharges and even though the occurrence of PD in this class of amplitude does not occur, the effect of is significant.

Keywords: partial discharges, humidity, Generalized Linear Model, Negative Binomial model.

LISTA DE FIGURAS

Figura 3.1- Série temporal da vazão afluente no período de janeiro de 2007 a abril de 2015 da UHE de Tucuruí.....	8
Figura 4.1 – Gráfico do kernel de densidade das amplitudes e dos ângulos das descargas do acoplador E.....	17
Figura 4.2 – Gráfico do kernel de densidade das amplitudes e dos ângulos das descargas do acoplador D.....	18
Figura 4.3 – Histograma das ocorrências de descargas parciais na 4ª medição no intervalo de amplitude de -800mV a 800mV.....	19
Figura 4.4 - Gráficos dos resíduos padronizados, os pontos de alavanca e a distância de Cook do modelo (4.2) com função de ligação log.....	23
Figura 4.5 - Gráficos dos resíduos padronizados, os pontos de alavanca e a distância de Cook do modelo (4.3) com função de ligação log.....	25
Figura 4.6 - Envelope dos resíduos studentizados sem as observações #37, #190 e #212.....	26

LISTA DE TABELAS

Tabela 4.1 – Estatísticas das amplitudes por medição do acoplador E do hidrogerador por medição considerando o intervalo de -800mV a 800mV.....	17
Tabela 4.2 – Estatísticas das amplitudes por medição do acoplador do lado direito (D) do hidrogerador por medição considerando o intervalo de -800mV a 800mV.....	18
Tabela 4.3. Estatísticas descritivas dos dois acopladores dos sensores do lado Esquerdo (E) e lado Direito (D).....	19
Tabela 4.4. Resultado do teste t de <i>Student</i> considerando o teste de igualdade de variância por meio do teste de Levene para a diferença de médias da ocorrências de DP nos dois sensores do lado Esquerdo (E) e lado Direito (D) por medição.....	20
Tabela 4.5. Ocorrência de DP por ângulo e intervalo de amplitude do sensor do lado esquerdo (E).....	21
Tabela 4.6 . Ocorrência de DP por ângulo e intervalo de amplitude do sensor do lado direito (D).....	21
Tabela 4.7 – Estimativa dos parâmetros do modelo Binomial Negativo (4.2).....	22
Tabela 4.8 – Valores da <i>deviance</i> residual, graus de liberdade, Razão, AIC e valor de p para o modelo (4.2) e variações.....	23
Tabela 4.9 – Estimativa dos parâmetros do modelo Binomial Negativo.....	24
Tabela 4.10 – Valores da <i>deviance</i> residual, graus de liberdade, Razão, AIC e valor de p para o modelo (4.3) e variações	24
Tabela 4.11 – Estimativa dos parâmetros do modelo Binomial Negativo (4.4).....	25
Tabela 4.12 – Valores da <i>deviance</i> residual, graus de liberdade, Razão, AIC e valor de p para o modelo (4.4) e variações.....	26

SUMÁRIO

1. Introdução.....	1
1.1 Aspectos Gerais.....	1
1.2 Justificativa.....	2
1.3. Objetivos.....	3
1.4 Estrutura do Trabalho.....	3
2. Descargas de Parciais.....	4
2.1 Estado da Arte.....	4
2.2. Relação entre a umidade e as descargas.....	5
3. Materiais e métodos	8
3.1. Área de estudo e coleta de dados.....	8
3.2. Análise exploratória dos dados.....	9
3.3. Modelos Lineares Generalizados.....	10
3.3.1 Modelo de regressão Binomial Negativa.....	11
3.4. Adequabilidade do Modelo.....	12
3.4.1. <i>Deviance</i>	12
3.4.2. Teste Wald.....	12
3.4.3. Método de Akaike.....	13
3.4.4. Superdispersão.....	13
3.4.5. Análises gráficas.....	13
3.5. Modelos de dispersão.....	14
3.5.1. Modelos próprios de dispersão.....	14
3.5.2. Modelos exponenciais de dispersão.....	15
3.5.3. Modelo de Tweedie.....	15
4. Aplicação em descargas parciais.....	16
4.1. Organização dos dados.....	16
4.2. Análise exploratória dos dados.....	16
4.3. Ajuste do Modelo.....	20
4.3.1. Modelo para o lado esquerdo (E) e lado direito (D).....	21
4.3.2. Modelo para o lado (E).....	23
4.3.3. Modelo para o lado (D).....	25
5. Conclusão.....	27
6. Referências Bibliográficas.....	29

Capítulo 1

1.INTRODUÇÃO

1.1.Aspectos gerais

A energia hidroelétrica é a principal fonte renovável em mais de 30 países e representa cerca de 20% de toda a eletricidade gerada no mundo. Segundo o Operador Nacional do Sistema Elétrico (ONS), atualmente 69,2% da energia elétrica consumida no Brasil advém de usinas hidrelétricas, 17,8% das usinas termelétricas, 9,5% Eólicas e 3,5% Nucleares. Os hidrogeradores são o cerne do circuito brasileiro de energia, isso implica dizer que caso ocorra uma indisponibilidade ou uma falha destes equipamentos, poderão ser atribuídas multas com valores elevados às concessionárias aplicadas pela Agência Nacional de Energia Elétrica (ANEEL), por não atendimento das demandas e até mesmo por paradas impróprias para efetivar manutenções, além de comprometer a confiabilidade do serviço de fornecimento de energia elétrica aos consumidores finais.

Diante do exposto, foi observado por diversos pesquisadores a importância da otimização da vida útil destes hidrogeradores nas usinas hidroelétricas, foram realizados vários estudos que almejavam elaborar estratégias para se priorizar a manutenção preditiva em detrimento das manutenções preventivas.

Um estudo internacional do CIGRE (2009) referente à causas de falhas em hidrogeradores, examinou minuciosamente 69 incidentes e mostrou que as principais causas de falhas nestes equipamentos podem ser categorizadas na seguinte ordem: 56% foram falhas no sistema de isolamento, 24% foram defeitos mecânicos, 17% problemas térmicos e por último, 3% foram falhas nos mancais de rolamento.

De acordo com Brasil (2016, p.5), no decorrer das pesquisas foi constatado que uma das principais causas de falhas em máquinas rotativas elétricas de grande porte são as Descargas Parciais (DPs) que são entendidas como defeitos no isolamento elétrico de forma parcial, que com o passar do tempo, podem evoluir para um defeito total causando um curto-circuito que poderá danificar o equipamento elétrico.

1.2. Justificativa

Numerosas pesquisas têm sido desenvolvidas utilizando técnicas estatísticas e da inteligência computacional para detecção, localização (GUILLEN *et al*, 2016), identificação de padrões de descargas parciais que possibilitem analisá-las corretamente assegurando o funcionamento do sistema elétrico (FRUTH, 1995; GROSS, 1998; SUWARNO; MA, 2013; RAYMOND, 2015).

Entre as principais técnicas estão os modelos de cadeia de Markov (GURURAJ e SATISH 1993), análise de componente principal, kernel análise de componentes principais, incorporação estocásticas de vizinhos, transformada wavelet discreta e as técnicas de inteligência computacional tais como redes neurais artificiais, algoritmos genéticos e técnicas da *matching learning*, sistemas fuzzy, sistema de inferência fuzzy baseado em rede adaptativa (MA *et al* , 2013; MAS'UD, 2014; SU *et al*, 2014; ALVES, 2015; RAYMOND, 2015) e modelos lineares generalizados (GOMES, 2016).

No entanto, ainda existem grandes desafios na aplicação com sucesso de técnicas estatísticas e da inteligência computacional no reconhecimento dos padrões e nas causas da ocorrência das descargas parciais em hidrogeradores. Assim, a pesquisa que gerou esta dissertação de mestrado, utilizou a análise exploratória de dados e os modelos lineares generalizados e propôs um método de investigação da influência do ciclo hidrológico e da umidade na ocorrência de descargas parciais, a partir de dois acopladores capacitivos conectados em ambos os lados do estator, com o estudo direcionado a fase V em períodos de seca e cheia de operacionalização da Usina Hidrelétrica de Tucuruí.

1.3. Objetivos

1.3.1 Geral

O objetivo desta pesquisa é propor um método para analisar a influência do ciclo hidrológico e da umidade no comportamento das descargas parciais em um hidrogerador.

1.3.2 Específicos

- Utilizar a análise exploratória de dados para avaliar as ocorrências das descargas por período de medição, ângulo e amplitude.
- Avaliar a ocorrência das descargas utilizando os modelos lineares generalizados.

1.4. Estrutura do trabalho

O presente trabalho está estruturado em cinco capítulos, considerando este introdutório. O segundo capítulo é constituído de uma revisão literária sobre descargas parciais e a influência da umidade nas ocorrências destas. No capítulo três tem-se uma explanação teórica atinente à análise exploratória dos dados e Modelos Lineares Generalizados, no tocante a estrutura, parâmetros que possibilitam verificar a qualidade do ajuste, além de abordar o Modelo Binomial Negativo, o qual descreve o comportamento das DPs no Hidrogerador de interesse. O quarto capítulo trata sobre os resultados e discussões da análise exploratória dos dados colhidos, especificamente de dois sensores localizados um do lado esquerdo e outro do lado direito do estator, bem como sobre o ajuste do modelo aos dados das descargas parciais. Por fim, no capítulo cinco se faz a conclusão desta pesquisa apresentando considerações referentes aos resultados obtidos.

Capítulo 2

2. DESCARGAS PARCIAIS

2.1. Estado da arte

De acordo com a norma internacional IEC 60270, a descarga parcial é definida como uma descarga elétrica localizada, que apenas parcialmente abre caminho na isolação entre os condutores e que pode ou não ocorrer adjacente a um condutor. Geralmente, tais descargas aparecem como pulsos com duração média de 1 nanossegundo.

Para Gururaj e Satish (1993), é um processo estocástico em que a posição da fase, quantidade e amplitude dos pulsos variam aleatoriamente com o tempo, devido à tensão e o envelhecimento da isolação. O processo de descarga parcial pode ser descrito matematicamente como um fenômeno que varia no tempo.

Os estudos com o objetivo de melhor compreender e medir as descargas parciais em máquinas rotativas tem sido realizados há mais de 60 anos. O primeiro trabalho sobre medições de descargas parciais foi realizado por Johnson em 1951, que investiu um método de detecção de DPs utilizando equipamentos auxiliares de teste de alta tensão, sem que os hidrogeradores fossem desligados. Além disso, o autor observou o aparecimento de sinais de alta frequência durante a operação da máquina. Este estudo associou a origem desses sinais com um tipo de descargas parciais denominada descarga de ranhura (BRASIL, 2016).

O padrão da descarga é representado por um conjunto de atributos (ou características), que podem ser extraídos dos dados de medição. É bem conhecido que diferentes tipos de fontes de descargas parciais podem gerar padrões de descarga diferentes (GURURAJ e SATISH, 1993; CAMPBELL, 1994; SATHIYASEKAR *et al*, 2011; MA *et al*, 2013).

Mais atualmente, Guillen *et al*. (2014) propuseram um método inovador para a localização de descargas parciais ao longo dos enrolamentos dos transformadores, no qual os sinais das DPs são processados por meio da transformada wavelet e a função de Laplace. Ao final, os sinais processados são comparados com os iniciais utilizando a distância de Hellinger para a determinação da posição destas descargas ao longo do enrolamento.

Alves (2015) apresenta técnicas de Inteligência computacional para caracterização de descargas parciais em hidrogeradores. Inicialmente, o autor utilizou as técnicas de binarização aos sinais para adequá-los ao uso das técnicas de inteligência computacional. E utilizou KNN (K-Nearest Neighbors.), Floresta Randômica e MVS (Máquinas de Vetores de Suporte) e constatou que os melhores resultados foram obtidos com a técnica MVS com a qual obteve 96,07% de exatidão em decorrência desta técnica possuir mecanismos de seleção das principais variáveis durante o processo de treinamento.

Raymond *et. al.* (2015) realizou uma revisão teórica sobre medições de descargas parciais utilizando testes *off-line* e *on-line*, ressaltando a necessidade de se quantificar os pulsos com base no ângulo de fase e nas amplitudes em relação a tensão aplicada na amostra. Enfatizou a vantagem em se conhecer as técnicas de classificação das DPs para que se possa identificar o defeito no início do ciclo e auxiliar a tomada de decisão no tocante a uma manutenção agendada e até mesmo na substituição completa da isolação elétrica, caso seja necessário.

Gomes (2016) propôs um método de análise de descargas parciais de um hidrogerador da Usina Hidráulica de Tucuruí por meio de Modelos Lineares Generalizados. O autor utilizou os modelos de Poisson e Binomial Negativo e observou que o Binomial Negativo foi o modelo que melhor se ajustou dado a superdispersão destes.

2.2. A relação entre a umidade e as descargas parciais

De acordo com a norma internacional IEC 60270, Corona é uma forma de descarga parcial que ocorre em meios gasosos em torno de condutores que estão distantes do isolamento sólido ou líquido.

Com a umidade de ar relativa do ar muito elevada, a absorção de água na superfície de uma partícula terá um impacto significativo no campo elétrico de corona, no aumento da condutividade elétrica e na mobilidade elétrica de íons.

As descargas corona na cabeça da bobina (end-winding) podem ocorrer em vários locais onde haja uma alta concentração do campo elétrico e geralmente ocorrem na interfaces entre diferentes elementos da cabeça da bobina dos enrolamentos do estator. Estas descargas de superfície geralmente se iniciam quando o campo elétrico ao longo da superfície excede o valor da rigidez dielétrica do gás ou líquido que está ao redor, tornando o revestimento ineficiente devido à contaminação por sujeira ou umidade (BRASIL, 2013).

Yawootti *et. al.*(2015), fizeram um estudo experimental para avaliar os efeitos da umidade relativa do ar, de entrada e da taxa de fluxo de ar nas ocorrência de descargas corona.

Esses autores apresentam uma revisão da literatura desde 1920, quando Peek propôs uma equação empírica para avaliar o campo elétrico propícios à descarga corona sem levar em consideração o efeito da umidade relativa do ar e as modificações da equação considerando este efeito por meio de estudos teóricos e experimentais. Os resultados experimentais mostraram que a ocorrência dessa descarga é fortemente afetada pelo nível de umidade relativa do ar de entrada.

Yawootti *et. al.*(2015) estudaram os efeitos da umidade relativa do ar (UR) de entrada e da taxa de fluxo de ar em descargas corona positivas e negativas. Eles utilizaram tensões entre 0 e 3,1 kV, taxas de fluxo de ar entre 5 e 15 L / min, uma umidade relativa entre 20 e 90%, e uma pressão de operação de cerca de 101,3 kPa. Os resultados experimentais mostraram que a corrente de descarga é fortemente afetada pelo nível da umidade relativa do ar de entrada. Eles verificaram que a corrente de descarga positiva diminuiu com o aumento da UR, abaixo de 60% e, aumentou com o aumento do valor de UR, acima de 60%, na mesma tensão.

Os autores *op.citem* verificaram que a corrente de descarga negativa permaneceu estável com o aumento UR (UR<40%) mas aumentou com o aumento do valor de UR (UR>40%) na mesma tensão de corona. Para os efeitos da taxa de fluxo de ar, a corrente de descarga positiva diminuiu um pouco com o aumento da taxa de fluxo (UR<90%) e aumentou com UR(UR> 90%). Para a corona negativa, a corrente de descarga diminuiu monotonicamente quando o aumento da taxa de fluxo de ar.

Wang *et. al.* (2016) investigaram de que forma a umidade relativa do ar afeta as atividades das descargas parciais no espaço compreendido entre o eletrodo e a isolação elétrica constituída de mica, resina epoxi e fibra de vidro de hidrogeradores, com aplicação de ondas de tensão trapezoidal alternada. Fizeram uma abordagem teórica que se iniciou reiterando que as principais causas de falhas nos geradores são o envelhecimento do material e as descargas parciais, que são definidas basicamente como descargas internas, de superfície, de delaminação e de ranhura, sendo esta, uma das mais prejudiciais para o sistema de isolação elétrica, ressaltando que, a ocorrência de uma de suas espécies, descarga de ranhura elétrica, se dá na região supracitada do estator. Fato

causado por uma condutividade inadequada na superfície do revestimento. Os autores destacaram que vários estudos mostraram que a umidade relativa do ar apresenta uma correlação negativa com as atividades das DPs, de forma que as descargas de ranhura diminuem ao passo que aumenta a umidade, pelo fato da natureza eletronegativa das moléculas de águas oriundas do ambiente úmido aumentarem o “coeficiente efetivo de ligação”, reduzindo assim a eficiência de ionização na região, o que leva à redução de elétrons livres disponíveis causadores das descargas, pois formam “íons de água” mais estáveis com uma probabilidade muito baixa de liberar seus elétrons durante colisões de moléculas. No entanto, enfatizaram que as atividades das DPs mais baixas com as umidades mais elevadas não reduzem os danos à isolamento elétrica, pelo contrário, a degradação se faz de maneira mais severa. Os autores simularam em laboratório ambientes com 8%, 29% e 77% de umidade relativa do ar e aferiram os pulsos temporais depois de cada tensão não senoidal ter sido aplicada na amostra durante 3 minutos. Concluíram que no ambiente mais úmido, houve uma diminuição na amplitude da tensão de inicialização das descargas, no entanto, ocorreu um aumento de um outro fator relevante para as atividades das DPs: o campo elétrico em torno da superfície isolante, pois com a condensação de vapor de água, a condutividade desta superfície aumentou significativamente, fato que influencia diretamente na mobilidade das cargas depositadas, espalhando-as em área maior neste local. Sendo assim, apesar dos eventos das descargas parciais ocorrerem mais tardiamente em relação ao ambiente seco, principalmente devido à falta de elétrons livres causados pela existência de moléculas de água no ar, o número médio de pulsos das descargas de ranhuras aumenta proporcionalmente ao aumento da umidade relativa do ar, podendo degradar mais o equipamento em detrimento do ambiente menos úmido.

Capítulo 3

Materiais e métodos

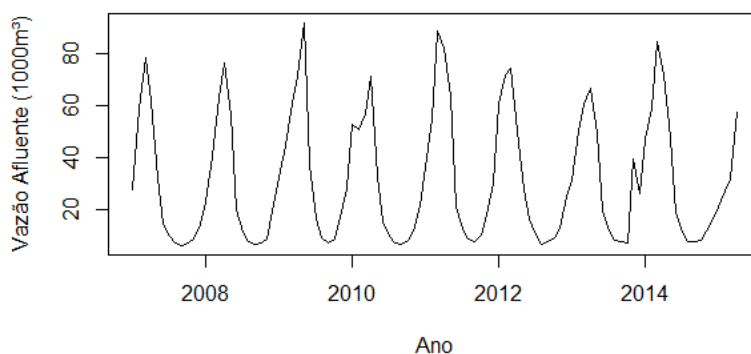
3.1 Área de estudo e coleta de dados

As hidrelétricas são responsáveis por 69,2% da produção de energia no Brasil, e tem como uma das principais vantagens utilizarem na geração de energia a água, que é uma fonte renovável, desde que haja uma constante preservação nas nascentes dos rios.

A Usina Hidrelétrica de Energia(UHE), opera com três tipos de vazões: afluyente, turbinada e vertida. A vazão afluyente é a vazão que seria observada no posto de medição considerando o rio na sua condição natural. A vazão turbinada é a vazão que passa pelas turbinas da hidrelétrica gerando energia. E a vazão vertida que é toda água escoada pelo vertedouro a fim de manter o reservatório no nível operacional estabelecido pela usina. O estudo dessas vazões é importante pois se pode minimizar o impacto das sazonalidades no nível operacional do reservatório, uma vez que toda UHE precisa estocar água durante o período das chuvas e estabelecer um planejamento eficaz deste recurso para manter a produção de energia de forma equilibrada.

A Figura 3.1 apresenta a série temporal da vazão afluyente (m^3/s) que foram coletados diariamente na UHE de Tucuruí no período de janeiro de 2007 a abril de 2015.

Figura 3.1- Série temporal da vazão afluyente no período de janeiro de 2007 a abril de 2015 da UHE de Tucuruí (Extraído de RAMOS, 2017)



Pode-se perceber que a vazão afluyente tem período sazonal bem definido e possui um volume maior concentrado nos meses de janeiro a julho (cheia) e menor de julho a

dezembro (seca). O período de cheia é o que mais gera energia e também o período que mais se verte água permitindo assim estabelecer o nível operacional da usina.

Em março de 2014, período de cheia, foram gerados em média 6552 MW de potência, com a utilização de 20 máquinas em média, (o valor mínimo 5370 MW e 16 máquinas; valor máximo 7195 MW e 22 máquinas), em setembro de 2014, período de seca, foram gerados 3898 MW em média, com a utilização de 12 máquinas em média, (o valor mínimo 2339 MW e 8 máquinas; valor máximo 5166 MW e 16 máquinas).

Valores aproximados em 2015, período de cheia, foram gerados 6853 MW em média, com a utilização de 20 máquinas em média, (o valor mínimo 6182 MW e 18 máquinas; valor máximo 7166 MW e 22 máquinas).

Para esta pesquisa foram utilizados os dados processados, coletados e armazenados pelo software Instrumento de Monitoração e Análise de Descargas Parciais Autônomo (IMA - DP autônomo), desenvolvido em parceria entre CEPEL (Centro de Pesquisas de Energia Elétrica) e ELETRONORTE, obtidos de dois acopladores capacitivos conectados no lado esquerdo (E) e direito (D) do estator da fase V, nos seguintes períodos: março de 2014 (cheia) setembro de 2014 (seca), maio de 2015 (cheia) e março de 2016 (cheia com intervenção), período este que foi feita a manutenção no hidrogerador.

E as análises dos dados foram feitas utilizando o programa computacional R (R program, 2014).

3.2. Análise Exploratória de Dados

A análise descritiva também denominada de análise exploratória de dados é o método que envolve a coleta, caracterização e apresentação, de um conjunto de dados de modo a descrever de maneira simples e objetiva as características deste por meio de gráficos e tabelas (BUSSAB e MORETTIN, 2005).

Neste trabalho foi utilizado além das tabelas e gráficos testes de hipóteses e análises das estatísticas descritivas além das estimativas da densidade do kernel tanto para amplitudes como para os ângulos.

O kernel é função simétrica de ponderação $K(x)$ que possui valores reais, com a integral definida sobre seu conjunto suporte é igual a 1, $\int_{\mathbb{R}} K(x)dx = 1$. Alguns tipos de função kernels são definidos sendo os mais comuns o Gaussiano, Retangular, Triangular, Biweight e Epanechnikov (GOMES, 2016).

A estimativa de densidade de kernel é um método não-paramétrico pois não assume qualquer distribuição subjacente para a variável para cada dado é criada uma função de kernel com o ponto de referência no seu centro. A função densidade de probabilidade é então estimada pela adição de todas estas funções do kernel e dividindo o resultado pelo número de dados. Assim, a estimativa pode ser dada por (3.1)

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (3.1)$$

onde $K(\cdot)$ é a função de kernel devidamente escolhida, n é a quantidade de dados, h é a largura de banda (*bandwidth*) ou parâmetro de suavização e x_i representa cada dado. Neste trabalho foi utilizado o programa computacional R (R program, 2014) que e foi escolhido o kernel gaussiano e automaticamente o melhor valor para h de acordo com os dados analisados.

3.3. Modelos Lineares Generalizados

O modelo de regressão normal linear, por muitas décadas atendeu satisfatoriamente aos objetivos pretendidos nessas análises. Entretanto uma das condições impostas por este modelo era que o conjunto de valores observados segue uma distribuição normal. Naturalmente surgiram outras situações que não puderam ser explicadas adequadamente pelo modelo, pelo fato de sua variável resposta seguir a uma distribuição não normal, ou ainda, por possuir um caráter não linear.

Desta forma outros modelos surgiram na tentativa de explicar melhor o fenômeno estudado tais como o modelo *probit*, os modelos *log-lineares*, o modelo complementar *log-log* e outros. Esses modelos apresentam uma estrutura de regressão linear e sua variável resposta segue uma distribuição pertencente a um conjunto de distribuições denominado família exponencial (CORDEIRO, 2013).

Os Modelos Lineares Generalizados (MLG) correspondem a uma síntese dos modelos citados e de outros modelos, unificando a teoria da modelação estatística da época. Assim, a variável resposta passa a seguir uma distribuição, membro da família exponencial de distribuições, sendo a sua estrutura linear descrita por um conjunto de variáveis independentes. A relação entre a média de todas as observações e a estrutura linear é estabelecida por uma função de ligação adequadamente escolhida para este fim (CORDEIRO, 2013).

As distribuições de probabilidade Normal, Gamma, Poisson, Binomial, Binomial Negativa, Normal Inversa, e outras, fazem parte da família exponencial e os modelos *logístico*, *log-linear*, *probit* e *normal linear* passam a ser particularidades suas.

3.3.1 Modelo de regressão Binomial Negativa

Os dados na forma de contagem são eventos que ocorrem numa mesma unidade de observação. A contagem pode ser realizada num intervalo espacial definido (número por unidade de área, volume e outros) ou no intervalo temporal (MONTROYA, 2009; CORDEIRO, 2010; SANTOS, 2013; GOMES, 2016).

O modelo de regressão de Poisson é gerado a partir da distribuição de Poisson. Nesta distribuição a variância é igual a sua média mas quando há indícios de superdispersão o modelo Binomial Negativo é mais adequado.

Para o modelo Binomial Negativo assume-se que a variável aleatória Y_i segue uma distribuição Binomial Negativa de parâmetros t e π , ou seja, $Y_i \sim \text{BN}(t, \pi)$. Assim a função de probabilidade de Y_i com $i = 1, \dots, n$, é dada por (3.2)

$$f(y_i; t, \pi) = \frac{\Gamma(y_i+t)}{\Gamma(y_i+1)\Gamma(t)} \pi^t (1-\pi)^{y_i} \quad (3.2)$$

onde $0 < \pi < 1$ e $t > 0$.

A função de probabilidade (3.2) pode ser reparametrizada considerando $t = \frac{1}{\alpha}$ e $\pi = \frac{1}{1+\alpha\mu_i}$ e daí se obtêm (3.3).

$$f(y_i; \mu_i, \alpha) = \frac{\Gamma(y_i+\frac{1}{\alpha})}{\Gamma(y_i+1)\Gamma(\frac{1}{\alpha})} \left(\frac{1}{1+\alpha\mu_i}\right)^{\frac{1}{\alpha}} \left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right)^{y_i} \quad (3.3)$$

com $y_i = 0, 1, \dots$, $0 < \mu_i < 1$ e α representando o parâmetro de heterogeneidade.

A média $E(Y)$ e a variância $\text{VAR}(Y)$ são dadas por (3.4)

$$E(Y) = \mu \quad \text{e} \quad \text{Var}(Y) = \mu + \alpha\mu^2 \quad (3.4)$$

Se $\alpha > 0$ então a $\text{Var}(Y) > E(Y)$, de modo que a distribuição binomial negativa pode ser utilizada para modelar dados com superdispersão.

Considerando Y uma variável aleatória que representa o número de ocorrências de determinado evento para um número finito de n observações, com o vetor X de variáveis explicativas dado por $X = (X_1, \dots, X_p)^T$ com $x_i^T = (x_{i1}, \dots, x_{ip})^T$ representando o vetor que corresponde a uma determinada observação do indivíduo i . Então se a observação x_i^T segue uma distribuição Binomial Negativa, esta pode se definida com (3.5).

$$Y|X = x_i^T \sim \text{BN}(\mu_i, \alpha) \quad (3.5)$$

onde $\mu_i = f(x_i^T, \beta)$ representa o número médio de ocorrências de um evento dada a observação x_i^T .

Se $\mu_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}$, $0 < \mu_i < 1$, utilizando a função log como função de ligação se obtém (3.6)

$$\log(\mu_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} \quad (3.6)$$

onde β_0 e β_j , com $j = 1, 2, \dots, p$, são os vetores dos coeficientes de regressão que indicam a variação esperada da média por unidade de variação na covariável X_j . Os parâmetros do modelo Binomial Negativo também podem ser estimados pelo método de máxima verossimilhança (PAULA, 2010; CORDEIRO, 2013).

3.4. Adequabilidade do Modelo

3.4.1 Deviance

A *deviance* avalia a discrepância entre os valores ajustados pelo modelo saturado e os valores ajustados pelo modelo proposto e é dada por (3.7)

$$D(y; \hat{\mu}) = 2(l_s - l_a) \quad (3.7)$$

onde $l_a = l(\hat{\mu}; y)$ representa a função log-verossimilhança do modelo proposto e $l_s = l(y; y)$, a função log-verossimilhança do modelo saturado.

A *deviance* assume sempre valores maiores ou iguais a zero e seu valor diminui à medida que se acrescenta covariáveis ao modelo, chegando ser igual a zero quando $l_a = l_s$. Para avaliar se o modelo proposto se ajusta bem aos dados, recorre-se ao teste de hipótese, com H_0 : o ajuste do modelo é igual ao do modelo saturado.

Para amostras consideravelmente grandes, e sob H_0 , que a *deviance* apresenta distribuição assintótica qui-quadrado com $j-(p+1)$ graus de liberdade, com j sendo o número de covariáveis e $(p + 1)$, o número de parâmetros do modelo proposto.

Para o Modelo Binomial Negativo a *deviance* é obtida por (3.8)

$$D(y; \hat{\mu}) = 2 \sum_{i=1}^n \left[y_i \log \left(\frac{y_i}{\mu_i} \right) - \left(y_i + \frac{1}{\alpha} \right) \log \left(\frac{1 + \alpha y_i}{1 + \alpha \mu_i} \right) \right] \quad (3.8)$$

3.4.2. Teste Wald

O teste de Wald segue uma distribuição normal assintótica de $\hat{\beta}$ e é dado por (3.9)

$$W = (\hat{\beta} - \beta_0)^T K_0 (\hat{\beta} - \beta_0) \quad (3.9)$$

onde K_0 representa a matriz de informação de Fisher dada por $K_0 = (X^T W_0 X)$. Para grandes amostras rejeita-se $H_0: \beta = \beta_0$ se $W > \chi_{p,1-\alpha}^2$, onde $(1 - \alpha)$ é o número de graus de liberdade e p é a dimensão do vetor β .

Para testar a hipótese nula $H_0: \beta_j = 0$ contra a hipótese alternativa $H_1: \beta_j \neq 0$, com $j = 1, 2, \dots, p$, a estatística de Wald é dada por (3.10).

$$W = \frac{\hat{\beta}_j}{\sigma(\hat{\beta}_j)} \sim N(0,1) \quad (3.10)$$

3.4.3. Método de Akaike

A informação de Akaike (AIC) se baseia na função log-verossimilhança e fornece uma medida relativa da informação perdida na adoção de um determinado modelo. Desse modo a escolha do modelo é feita a partir do menor valor apresentado por este critério. A estatística correspondente é dada por (3.11)

$$AIC = -2l(\hat{\theta}) + 2p \quad (3.11)$$

onde $\hat{\theta}$ é o estimador de máxima verossimilhança, l é a função de log-verossimilhança e p representa o número de parâmetros do modelo

3.4.4. Superdispersão

Se espera que no ajuste de um modelo a *deviance* residual seja aproximadamente igual ao seu número de graus de liberdade. A não ocorrência disso pode estar relacionada ao componente aleatório do modelo, pela presença de uma variabilidade não percebida pelo modelo. Pode ainda estar ligada ao seu preditor linear, pela ausência de uma estrutura ótima, ou ainda, a uma função de ligação inadequada (CORDEIRO, 2013).

Para avaliar a superdispersão se sugere obter a razão entre a deviance residual e os graus de liberdade residual. Se esta razão for igual a 1, então conclui-se que a variância é igual a média da variável resposta dos dados e alguns autores recomendam verificar a presença de superdispersão caso esta razão seja maior que 2 (GOMES, 2016).

3.4.5. Análises gráficas

A análise gráfica corresponde a técnicas usadas para auxiliar na verificação do ajuste de um modelo. Estas são de âmbito informal, pois se baseia em análise visual de gráficos para detectar padrões ou pontos atípicos. Alguns tipos de gráficos são importantes para se verificar a presença de pontos aberrantes ou influentes no modelo proposto entre os quais o gráfico de probabilidade normal e os gráficos de índice.

No eixo horizontal dos gráficos de índices estão as ordens das observações. Estes permitem verificar a presença de pontos atípicos que poderão ser classificados como influentes, não influentes ou como pontos aberrantes (*outlier*), para os propósitos deste trabalho foram utilizados os gráficos:

- i. *deviance* residual padronizada para detectar pontos aberrantes;
- ii. pontos de alavanca h_{ii} (*leverage*) para observar pontos influentes;
- iii. distância de Cook usado para confirmar se os pontos tidos como *leverage* são realmente influentes.

3.5. Modelos de dispersão

Os Modelos Lineares Generalizados foram, originalmente, desenvolvidos para famílias exponenciais de distribuição, mas a ideia principal pode ser estendida para uma classe mais ampla de modelos probabilísticos denominada modelos de dispersão (MD).

Um modelo de dispersão, MD (μ, σ^2) , com parâmetro de locação μ e parâmetro de dispersão σ^2 pertence à família de distribuições com função densidade de probabilidade dada por (3.12):

$$f(y; \mu, \sigma^2) = a(y; \sigma^2) \exp \left\{ -\frac{1}{2\sigma^2} d(y; \mu) \right\}, y \in C \quad (3.12)$$

em que $\Omega \subseteq C \subseteq \mathbb{R}$, $a(\cdot) \geq 0$ é uma função apropriada, $\mu \in \Omega$ e $\sigma^2 > 0$ e d é o componente do desvio, satisfazendo $d(\mu; \mu) = 0$ para $\mu \in \Omega$ e $d(y; \mu) > 0$ para $y \neq \mu$. A unidade do desvio é dita regular se for diferenciável duas vezes respectivamente por y , μ em $\Omega \times \Omega$ satisfazendo:

$$\frac{\partial^2 d}{\partial \mu^2}(\mu, \mu) > 0, \forall \mu \in \Omega$$

A função de variância $V: \Omega \rightarrow \mathbb{R}_+$ do desvio regular é definida por (3.13):

$$V(\mu) = \frac{2}{\frac{\partial^2 d}{\partial \mu^2}(\mu, \mu)} \quad (3.13)$$

Esses modelos são classificados em dois tipos: modelos próprios de dispersão PD (μ, σ^2) e modelos exponenciais de dispersão ED (μ, σ^2) os quais são definidos a seguir.

3.5.1. Modelos próprios de dispersão

Um modelo próprio de dispersão denominado por PD (μ, σ^2) , com parâmetro de locação μ e parâmetro de dispersão σ^2 tem a função densidade de probabilidade dada por (3.14):

$$f(y; \mu, \sigma^2) = \frac{a(y; \sigma^2)}{\sqrt{V(\mu)}} \exp \left\{ -\frac{1}{2\sigma^2} d(y; \mu) \right\}, y \in C \quad (3.14)$$

em d é o desvio regular com função de variância $V(\mu)$.

Esses modelos são caracterizados pelo desvio d , pois a função de variância (3.13) está em função de d , com a $a(\sigma^2)$ em torno da constante normalizada. E são utilizados para análise de dados limitados em um intervalo, como os dados direcionais (ou cíclicos) e as proporções (SOUSA, 2010).

3.5.2. Modelos exponenciais de dispersão

Um modelo exponencial de dispersão denominado por $ED(\mu, \sigma^2)$, com parâmetro de locação μ e parâmetro de dispersão σ^2 tem a função densidade de probabilidade dada por (3.14):

$$f(y; \mu, \sigma^2) = c(y) \exp\{\theta y - k(\theta)\}, \quad y \in \mathbb{R} \quad (3.14)$$

em que $c(y)$ é uma função e $k(\theta)$ é a correspondente função de cumulante com parâmetro canônico θ pertencente ao conjunto $\Theta = \{\theta \in \mathbb{R} | k(\theta) < \infty\}$. A média de y é $\mu = \tau(\theta)$, $\tau(\theta) = k'(\theta)$ e $\mu \in \Omega = \tau(\text{int}\Theta)$ e função de variância, $V(\mu) = \tau''(\tau^{-1}(\mu))$.

3.5.2.1. Modelo de Tweedie

O modelo de Tweedie é uma classe do modelo de dispersão exponencial denominado por $T_{Wp}(\mu, \sigma^2)$ com função de variância definida por (3.15)

$$V(\mu) = \mu^p, \quad \mu \in \Omega \quad (3.15)$$

Em que p é o parâmetro com domínio $(-\infty, 0] \cup [1, \infty)$. Este modelo foi utilizado por Jorgensen & Sousa (1992) para resolver problemas de precificação de seguros. O modelo tem média μ e variância $\text{Var}(Y) = \sigma^2 \mu^p$.

O modelo Tweedie é invariante para as transformações de escala, isto é,:

Se $Y \sim T_{Wp}(\mu, \sigma^2)$, então, $cY \sim T_{Wp}(c\mu, c^{2-p}\sigma^2)$, já que $E(cY) = c\mu$, e $\text{Var}(cY) = c^2\sigma^2\mu^p = \sigma^2 c^{2-p}(c\mu)^p$, para $c > 0$. Esta propriedade é característica do modelo de Tweedie (FERREIRA, 2005; SOUSA, 2010).

Alguns exemplos do Modelo de Tweedie são encontrados como distribuição Normal ($p = 0$), distribuição Poisson ($p = 1$); distribuição Gama ($p = 2$) e distribuição Poisson Composta ($1 < p < 2$).

Capítulo 4

Aplicação em Descargas Parciais

Neste trabalho foram utilizados os registros de descargas parciais da base de dados do centro de pesquisa da Eletronorte para a realização desta pesquisa. Os dados processados são oriundos de um dos hidrogeradores da Usina Hidroelétrica de Tucuruí e foram coletados a partir de dois acopladores capacitivos conectados no lado esquerdo (E) e direito (D) do estator. Neste, será considerado os dados oriundos da fase V nos seguintes períodos: março de 2014 (cheia), setembro de 2014 (seca), maio de 2015 (cheia) e março de 2016 (cheia com intervenção). Isto é, devido ao aumento de DP no hidrogerador os analistas fizeram uma manutenção em março de 2016.

4.1. Organização dos dados

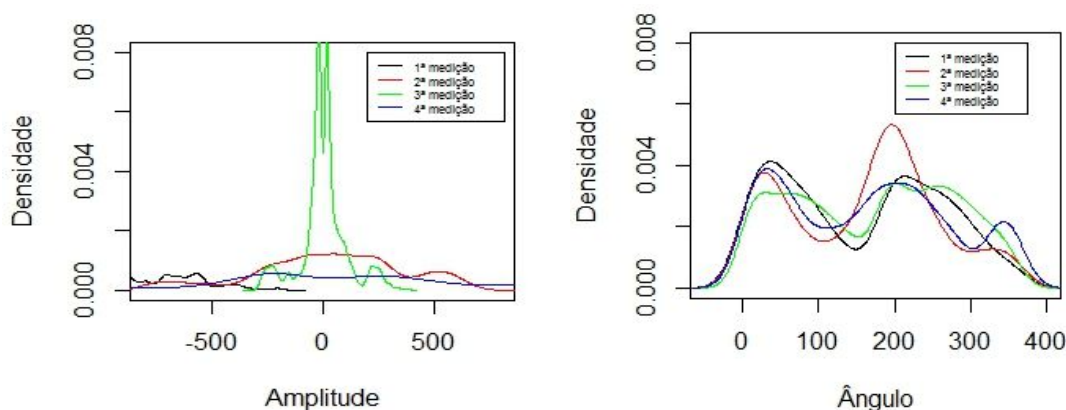
Os dados obtidos a partir de cada acoplador foram organizados em tabelas considerando o número de ocorrências de descargas parciais com sua respectiva amplitude e ângulo de fase. Foi considerada a ocorrência de descargas em intervalos de classe de 100 mV para amplitude e para os ângulos em intervalos de classe de 90°. As Tabelas a seguir, exibem as ocorrências de descargas parciais nos intervalos de amplitude e ângulo para os sensores, nas quatro medições realizadas nos períodos supracitados, bem como os seus respectivos lados.

4.2. Análise exploratória dos dados

A análise exploratória se houve por meio do estudo dos *kernels* considerando todas as ocorrências de descargas parciais por acoplador na fase de interesse V, nas quatro medições realizadas nos períodos já elencados.

A Figura 4.1 mostra os gráficos dos *kernels* das densidades das ocorrências de DP's nas amplitudes dos acopladores capacitivos do lado esquerdo (E) do hidrogerador. Desse modo, foi observado que a densidade da terceira medição foi muito maior do que a das outras medições. E quanto ao ângulo de ocorrências de DP não houve grandes modificações nos kernels.

Figura 4.1 – Gráfico do kernel de densidade das amplitudes e dos ângulos das descargas do acoplador E



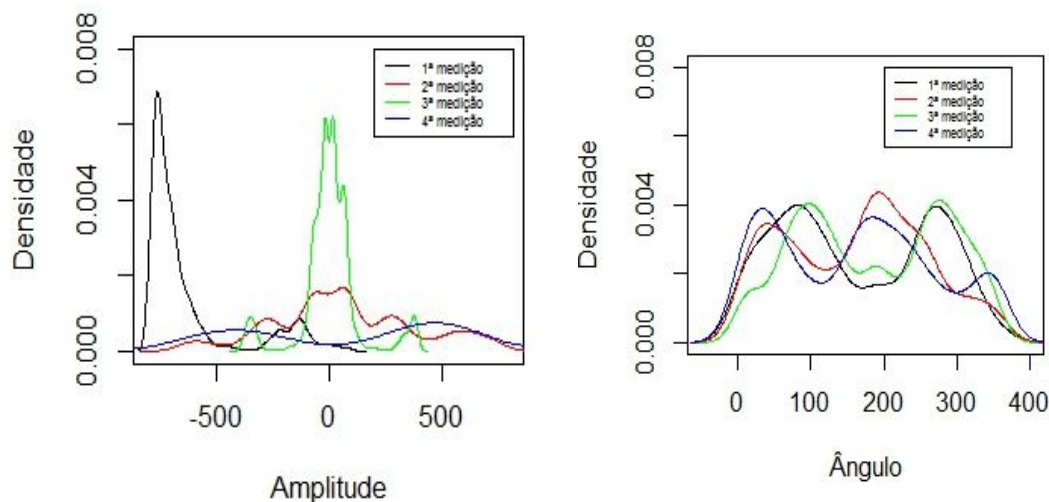
A Tabela 4.1. apresenta as estatísticas descritivas do acoplador E considerando o intervalo de amplitudes de -800mV a 800mV. Pode-se observar o grande aumento da média de ocorrência das descargas na terceira medição com 141,70, bem como a grande variação e todas assimetrias positivas (Figura 4.4.(a)).

Tabela 4.1 – Estatísticas das amplitudes por medição do acoplador E do hidrogerador por medição considerando o intervalo de -800mV a 800mV

Estatística	1ª.Medição	2ª.Medição	3ª.Medição	4ª.Medição
Média	55,23	18,10	141,70	30,05
Mediana	5	8	50	12
Desv. Pad.	118,04	24,64	215,39	40,49
Assimetria	2,28	2,12	1,87	1,90
Curtose	6,71	7,44	5,64	5,94

A Figura 4.2 mostra os gráficos dos *kernels* das densidades das ocorrências de DP's nas amplitudes do acoplador capacitativo do lado direito (D) do hidrogerador. Desse modo, foi observado que as densidades da primeira e terceira medição foram muito maiores quando comparadas às das outras medições e maiores variações nas densidades dos ângulos.

Figura 4.2 – Gráfico do kernel de densidade das amplitudes e dos ângulos das descargas do acoplador D



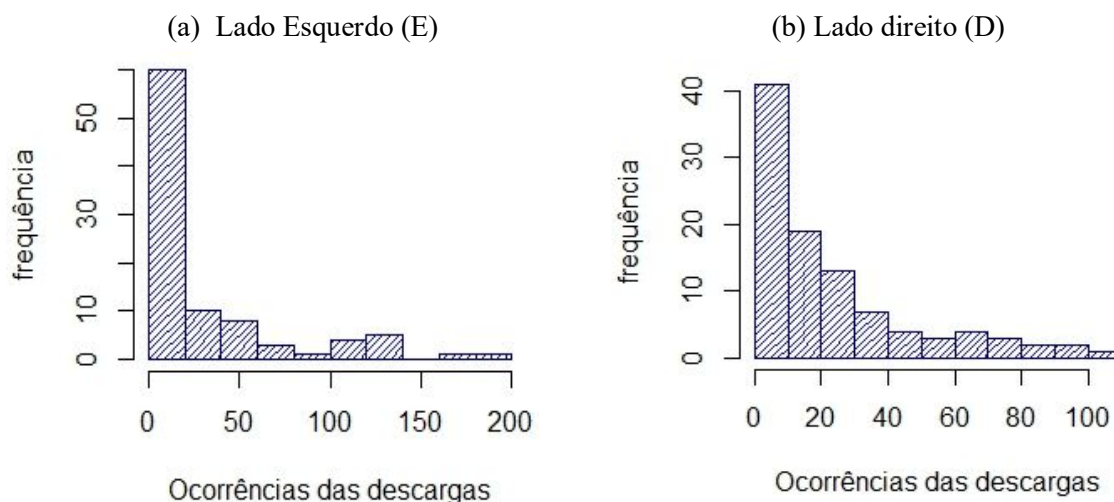
A Tabela 4.2. apresenta as estatísticas descritivas do acoplador D considerando o intervalo de amplitudes de -800mV a 800mV. Pode-se observar o grande aumento da média de ocorrência das descargas na terceira medição com 156,18, bem como a grande variação e com todas assimetrias positivas (Figura 4.4.(b)).

Tabela 4.2 – Estatísticas das amplitudes por medição do acoplador do lado direito (D) do hidrogerador por medição considerando o intervalo de -800mV a 800mV.

Estatística	1ª.Medição	2ª.Medição	3ª.Medição	4ª. Medição
Média (mV)	38,87	21,67	156,18	24,08
Mediana (mV)	12	13	38	15
Desv. Pad.(mV)	58,66	25,67	261,05	24,96
Assimetria	2,17	2,27	2,31	1,50
Curtose	7,24	8,97	7,76	4,63

As figuras 4.3 (a) e (b) exemplificam as assimetrias positivas na quarta medição a partir do histograma das ocorrências de DP nos dois acopladores.

Figura 4.3 – Histograma das ocorrências de descargas parciais na 4ª medição no intervalo de amplitude de -800mV a 800mV



Para os dois acopladores do lado esquerdo (E) e do direito (D) se têm as estatísticas da Tabela 4.3. Nesta, se pode observar a grande variabilidade das ocorrências das descargas nas quatro medições e que a quantidade de DP não muito diferentes entre os lados. Observa-se também que no período de cheia antes da intervenção (1ª e 3ª medições) há grande diferença na média da ocorrência de descarga parcial nos dois sensores.

Tabela 4.3. Estatísticas descritivas dos sensores do lado Esquerdo (E) e lado Direito (D)

Medição	Sensor	n	Média (mV)	Desv.Pad (mV)	CV(%)
1ª.	Esquerdo (E)	43	55,23	118,04	213,72
	Direito (D)	47	38,87	58,66	150,91
2ª.	Esquerdo (E)	75	18,11	24,64	136,10
	Direito (D)	80	21,68	25,68	118,48
3ª.	Esquerdo (E)	51	141,71	215,40	152,00
	Direito (D)	49	156,18	261,50	167,14
4ª.	Esquerdo (E)	93	30,05	40,50	134,75
	Direito (D)	99	24,08	24,97	103,69

Para testar se a ocorrência de DP parciais são similares nos dois acopladores do lado esquerdo (E) e do direito (D) foram conduzidos os testes de t de Student considerando as diferenças de variância pelo teste de Levene com os resultados na Tabela 4.4. As variâncias dos dois acopladores são significativamente diferentes na primeira e na quarta medições, mas as médias das ocorrências não são significativas.

Tabela 4.4. Resultado do teste t de *Student* considerando o teste de igualdade de variância por meio do teste de Levene para a diferença de médias da ocorrências de DP nos dois acopladores dos sensores do lado Esquerdo (E) e lado Direito (D) por medição.

Medição	Levene		t de <i>Student</i>		
	F	p	t	Gl	p
1ª.	6,83	0,010	0,82	60,31	0,415
2ª.	0,02	0,885	-0,88	153	0,379
3ª.	0,34	0,562	-0,30	98	0,763
4ª.	10,88	0,001	1,22	151,32	0,224

Valores em negrito são significativos considerando $\alpha=5\%$

4.3. Ajuste de modelos

Para facilitar a modelagem os dados foram sumarizados por classes de 90° referentes aos ângulos e em classes de amplitude: menores de -100 mV, entre -100mV e 100mV e acima de 100mV. E foi considerado que as ocorrências de descargas parciais obtidas pelos dois sensores são independentes nas quatro medições e que também são independentes entre eles.

Nas Tabelas 4.5 e 4.6 pode-se observar que para os dois acopladores (E e D) um aumento da ocorrência de DP entre -100 a 100 mV na terceira medição e uma diminuição das ocorrências nessa classe após a intervenção dos analistas no hidrogerador.

Para o acoplador do lado esquerdo (E), na Tabela 4.5, considerando a segunda medição como base, houve um aumento nas ocorrências de descargas parciais de 910%, 1370%, 1248% e 3019% respectivamente para as classes de ângulos: 0° a 90°, 90° a 180°, 180° a 270° e 270° a 360° na classe de -100mV a 100mV. E após a intervenção dos analistas há uma diminuição de 4950%, 8608%, 6407% e 6688% para respectivamente estas mesmas classes de ângulos e classe de amplitude.

Tabela 4.5. Ocorrência de DP por ângulo e intervalo de amplitude do sensor do lado esquerdo (E)

Medição	Amplitude (mV)	Ângulo				Total
		0° - 90°	90 °-180 °	180° -270°	270° -360 °	
1a.	<-100	914	331	804	323	2372
2a.	<-100	230	72	5	66	373
	[-100, 100]	200	77	140	37	454
	>100	2	97	396	36	531
3a.	<-100	154	188	108	18	468
	[-100, 100]	2020	1132	1887	1154	6193
	>100	75	25	283	183	566
4a.	<-100	246	297	513	294	1350
	[-100, 100]	40	13	29	17	99
	>100	638	275	215	218	1346

Já para o acoplador do lado esquerdo (D), na Tabela 4.6, considerando a segunda medição como base, houve um aumento nas ocorrências de descargas parciais na classe de amplitude -100mV a 100mV de 347%, 1504%, 2343% e 738% respectivamente para as classes de ângulos: 0° a 90°, 90 ° a 180°, 180 ° a 270 ° e 270 ° a 360°. E após a intervenção dos analistas o sensor não captou DP nessa classe de amplitude.

Tabela 4.6 . Ocorrência de DP por ângulo e intervalo de amplitude do sensor do lado direito (D)

Medição	Amplitude (mV)	Ângulo				Total
		0° - 90°	90 °-180 °	180° -270°	270° -360 °	
1a.	<-100	557	391	427	393	1768
	[-100, 100]	23		36		59
2a.	<-100	240	49	8	50	347
	[-100, 100]	247	132	336	102	817
	>100	44	151	332	43	570
3a.	<-100	240	61	4	66	371
	[-100, 100]	1104	2117	1132	2492	6845
	>100		51	316	70	437
4a.	<-100	217	177	366	281	1041
	>100	601	341	242	159	1343

4.3.1. Modelo para o lado esquerdo (E) e lado direito (D)

A variável resposta obtida pelos dois acopladores capacitivos conectados à fase V, se refere a dados de contagem, pois diz respeito ao número de ocorrência de descargas parciais captadas. Como foi observado na Tabela 4.3, coeficientes de variação em todas as medições se houveram elevados, descartando a possibilidade que os dados seguirem

uma distribuição de Poisson. Foram estudados vários modelos para diferentes valores de p da equação (3.15), mas não apresentaram ganho no tocante a normalidade dos resíduos, desta forma, foram testados modelos da distribuição binomial negativa definido em (3.6). Inicialmente foram testados modelos considerando os dois acopladores e as interações entre os fatores, porém as estimativas de seus parâmetros não foram significativas, após outras especificações, foi obtido o modelo (4.2), onde X_1 é o fator que se refere ao lado esquerdo ou direito; X_2 é a variável dummy que indica se a ocorrência aconteceu na classe de -100mV a 100mV; X_3 é a variável dummy que indica se a ocorrência aconteceu no período de cheia ou de seca e, X_4 é a variável dummy que indica se a ocorrência aconteceu no período de cheia depois da intervenção dos analistas.

$$\log(\mu) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_2 X_3 \quad (4.2)$$

Foram retirados as observações #1, #8 e #18 e estudadas o efeito sobre a estimativa do modelo que não foram alteradas (Tabela 4.7). Na cheia, a ocorrência de DP na classe de amplitude de -100 a 100mV aumenta (11,55%) e no período de seca ocorre uma redução na quantidade de descargas. Mesmo não considerando as ocorrências de DP nessa classe de amplitude, o efeito da cheia (12%) é significativo. Observa-se também que no período depois da intervenção houve redução (0,70 %) nas descargas parciais. Houve um aumento de 712% das ocorrências da classe -100mV a 100mV e na cheia, quando comparado com a não ocorrência nesta classe de amplitude na seca. E que não existe diferenças significativas nas ocorrências dos lados esquerdo e direito como foi observado no teste t de Student na Tabela 4.4.

Tabela 4.7 – Estimativa dos parâmetros do modelo Binomial Negativo (4.2)

Coeficientes	Est.	Er. Pad.	z	Pr(> z)
β_0	4,06	0,27	14,94	0,000
β_1	-0,37	0,29	-1,31	0,190
β_2	2,53	0,73	3,44	0,000
β_3	2,56	0,39	6,56	0,000
β_4	-1,17	0,40	-2,90	0,000
β_5	6,57	1,49	4,40	0,000

Referências: β_1 : Lado Esquerdo; β_2 : não ocorrência na classe de -100mV a 100mV; β_3 ; Seca; β_4 : Não intervenção; β_5 : não ocorrência na classe de -100mV a 100mV e na seca

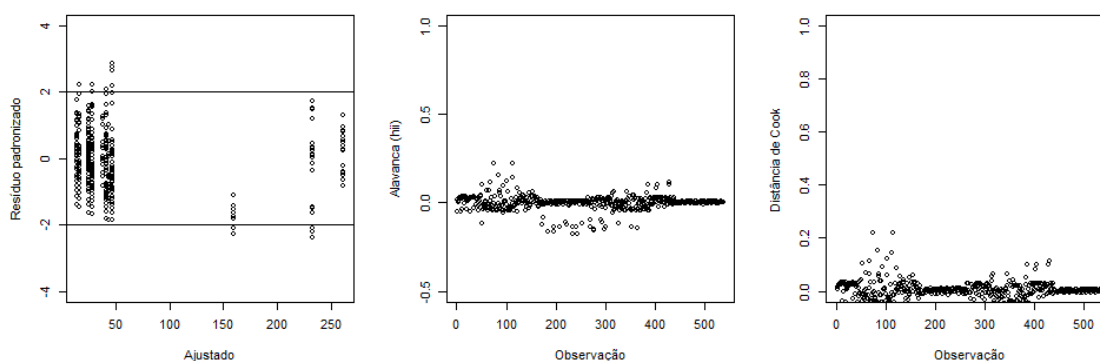
A *deviance* residual, seu grau de liberdade, a razão entre ambos, a medida de Akaike (AIC) e o valor de p para o teste de vários modelos estão dispostos na Tabela 4.8. É possível observar que razão da *deviance* residual dividida pelos graus de liberdade apresenta o valor muito próximo à unidade, indicando que os modelos captaram a superdispersão, embora os valores de p foram significativos decorrentes de dados

aberrantes como mostra a Figura 4.4 com os gráficos dos resíduos padronizados, os pontos de alavanca e a distância de Cook para o modelo (4.2).

Tabela 4.8 – Valores da *deviance* residual, graus de liberdade, Razão, AIC e valor de *p* para o modelo (4.2) e variações.

Modelo	Ligação	Deviance Residual	G. Liberdade	Razão	AIC	p
(4.2)	Log	626,79	531	1,18	4824,70	0,003
	Raiz quadrada	627,30	531	1,18	4829,10	0,002
	Identidade	627,52	531	1,18	4830,80	0,002
(4.2) Sem #1, #8, e #28	Log	620,25	528	1,17	4754,40	0,003

Figura 4.4. Gráficos dos resíduos padronizados, os pontos de alavanca e a distância de Cook do modelo (4.2) com função de ligação log.



4.3.2. Modelo para o lado (E)

Como pode ser observado na Tabela 4.5 só foram observadas as ocorrências de descargas parciais na segunda, terceira e quarta medição na classe de -100mV a 100mV. Assim, para modelagem só serão utilizados os dados destas medições para os modelo definido em (4.3).

$$\log(\mu) = \beta_0 + \beta_1 X_2 + \beta_2 X_3 + \beta_3 X_4 + \beta_4 X_2 X_3 \quad (4.3)$$

onde X_2 é a variável dummy que indica se a ocorrência aconteceu na classe de -100mV a 100mV; X_3 é a variável dummy que indica se a ocorrência aconteceu no período de cheia ou de seca e, X_4 é a variável dummy que se indica que a ocorrência se deu no período de cheia depois da intervenção dos analistas.

A Tabela 4.9 apresenta as estimativas dos parâmetros do modelo (4.3). Na cheia, a ocorrência de DP na classe de amplitude de -100 a 100mV aumenta ($1,41 \times 10^{13} \%$) e no período de seca ocorre uma redução na quantidade de descargas. Também é muito significativo o efeito da interação entre a ocorrência de DP na classe de amplitude de -100 a 100mV e o período da cheia ($1,33 \times 10^{63} \%$). Observa-se que no período depois da intervenção não é significativo.

Tabela 4.9 – Estimativa dos parâmetros do modelo Binomial Negativo (4.3)

Coeficientes	Est.	Er. Pad.	z	Pr(> z)
β_0	15,32	2,67	5,74	0,000
β_1	13,05	9,80	1,33	0,182
β_2	30,28	7,32	4,14	0,000
β_3	-12,47	8,40	-1,49	0,138
β_4	145,35	48,73	2,98	0,003

Referências: β_1 : não ocorrência na classe de -100mV a 100mV; β_2 : Seca;

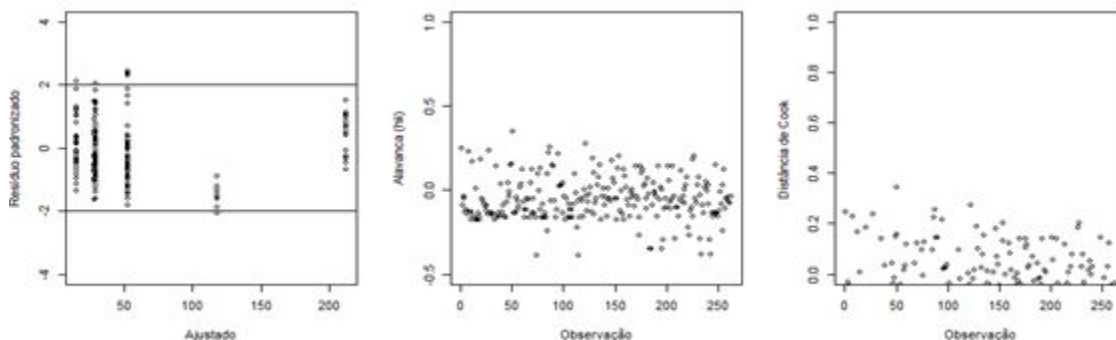
β_3 : Não intervenção; β_4 : Não ocorrência na classe de -100mV a 100mV e na seca

A *deviance* residual, seu grau de liberdade, a razão entre ambos, a medida de Akaike (AIC) e o valor de *p* para o teste de vários modelos estão dispostos na Tabela 4.10. É possível observar que razão da *deviance* residual dividida pelos graus de liberdade apresenta o valor muito próximo à unidade, indicando que os modelos captaram a superdispersão, embora os valores de *p* foram significativos, decorrentes de dados aberrantes como mostra a Figura 4.5 com os gráficos dos resíduos padronizados, os pontos de alavanca e a distância de Cook para o modelo (4.3).

Tabela 4.10 – Valores da *deviance* residual, graus de liberdade, Razão, AIC e valor de *p* para o modelo (4.3) e variações

Modelo	Ligação	Deviance Residual	G. Liberdade	Razão	AIC	p
(4.3)	Log	310,61	257	1,21	2376,60	0,012
	Raiz quadrada	311,15	257	1,21	2380,80	0,012
	Identidade	311,34	257	1,21	2382,20	0,011
(4.3) Sem #1, #8, e #28	Log	310,61	257	1,21	2376,60	0,012

Figura 4.5. Gráficos dos resíduos padronizados, os pontos de alavanca e a distância de Cook do modelo (4.3) com função de ligação log.



4.3.3. Modelo para o lado (D)

Como pode ser observado na Tabela 4.6 só foram observadas as ocorrências de descargas parciais na segunda, terceira medição na classe de -100mV a 100mV. Assim, para modelagem só serão utilizados os dados destas medições para os modelo definido em (4.4).

$$\log(\mu) = \beta_0 + \beta_1 X_2 + \beta_2 X_3 + \beta_3 X_2 X_3 \quad (4.4)$$

onde X_2 é a variável dummy que se indica que a ocorrência aconteceu na classe de -100mV a 100mV e X_3 é a variável dummy que se indica que a ocorrência aconteceu no período de cheia ou de seca.

A Tabela 4.11 apresenta as estimativas dos parâmetros do modelo (4.4). Na cheia, a ocorrência de DP na classe de amplitude de -100 a 100mV aumenta e no período de seca ocorre uma redução na quantidade de descargas. Também é muito significativo o efeito da interação entre a ocorrência de DP na classe de amplitude de -100 a 100mV e o período da cheia. E o efeito da cheia nas ocorrências de DP quando comparado ao período da seca.

Pode-se observar que por meio dos coeficientes estimados quando comparados às referências que: houve um aumento de 2,52% das ocorrências da classe -100mV a 100mV e um aumento de 0,52% na cheia. Houve um aumento de 2,78 % de ocorrências das descargas na classe -100mV a 100mV e na cheia.

Tabela 4.11 – Estimativa dos parâmetros do modelo Binomial Negativo (4.4)

Coeficientes	Est.	Er. Pad.	z	Pr(> z)
β_0	3,02	0,09	32,32	0,000
β_1	1,26	0,18	6,68	0,000
β_2	0,42	0,15	3,46	0,000
β_3	1,13	0,27	3,87	0,000

Referências: β_1 : Não ocorrência na classe de -100mV a 100mV; β_2 : Seca; β_3 : Não ocorrência na classe de -100mV a 100mV e na Seca

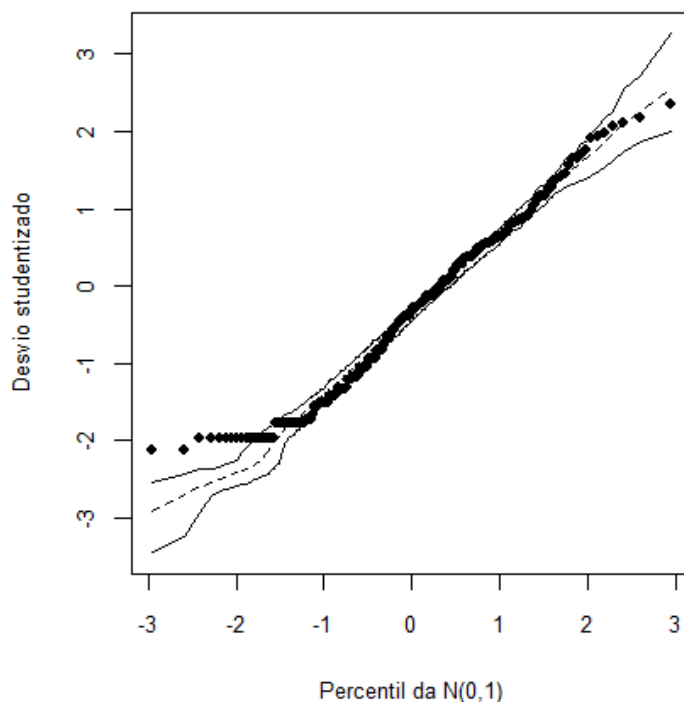
A *deviance* residual, seu grau de liberdade, a razão entre ambos, a medida de Akaike (AIC) e o valor de p para o teste de vários modelos estão dispostos na Tabela 4.12. É possível observar que razão da *deviance* residual dividida pelos graus de liberdade apresenta o valor muito próximo à unidade, indicando que os modelos captaram a superdispersão. A Figura 4.6 exhibe o envelope dos desvios studentizados para o modelo (4.4) sem as observações #37, #190 e #212, sem as quais as estimativas do modelo (4.4) não foram alteradas.

Tabela 4.12 – Valores da *deviance* residual, graus de liberdade, Razão, AIC e valor de p para o modelo (4.4) e variações

Modelo	Ligação	Deviance Residual	G. Liberdade	Razão	AIC	p
(4.4)	Log ^(*)	366,57	316	1,16	3093,10	0,026
(4.4) Sem #37, #190 e #212	Log	361,41	313	1,15	3033,60	0,031

^(*) Mesmos valores para as funções de ligação raiz quadrada e identidade.

Figura 4.6. Envelope dos resíduos studentizados sem as observações #37, #190 e #212.



Capítulo 5

5. Conclusão

Neste trabalho foi apresentado um método de organização e de análise de descargas parciais considerando o ciclo hidrológico da Usina Hidrelétrica de Tucuruí.

A primeira etapa deste método foi a estruturação dos dados brutos, que foram agrupados exibindo o número de ocorrências de descargas parciais por amplitude e ângulo de fase. Em seguida, foi realizada a análise exploratória dos dados que consistiu no estudo dos *kernels* de dois acopladores do hidrogerador: um do lado esquerdo e o outro do lado direito.

A variável resposta obtida pelos dois acopladores capacitivos conectados à fase V, se refere a dados de contagem, mas como os coeficientes de variação foram muito altos foi descartada a possibilidade que os dados seguirem uma distribuição de Poisson e testados modelos da distribuição binomial negativa, considerando que as ocorrências de descargas parciais obtidas pelos dois sensores são independentes nas quatro medições e que também são independentes entre eles.

Para facilitar a modelagem os dados foram sumarizados por classes de 90° referentes aos ângulos e em classes de amplitude: menores de -100 mV, entre -100mV e 100mV e acima de 100mV e foram criadas três variáveis dummies. A primeira para tentar captar se a ocorrência aconteceu na classe de -100mV a 100mV. A segunda para captar se a ocorrência da descarga parcial aconteceu no período de cheia ou de seca. E a terceira variável dummy para captar a ocorrência de descargas no período de cheia depois da intervenção dos analistas.

Na etapa da aplicação do modelo inicialmente foi ajustado um modelo considerando os acopladores dos dois lados e posteriormente para cada lado individualmente, considerando vários modelos Tweedie para diferentes valores de p mas não apresentaram ganho no tocante a normalidade dos resíduos, desta forma, foram testados três modelos considerando que as ocorrências de descargas parciais seguem a distribuição binomial negativa.

Nos três modelos avaliados, na cheia, a ocorrência de DP na classe de amplitude de -100 a 100mV aumenta e no período de seca ocorre uma redução na quantidade de

descargas e mesmo não considerando as ocorrências de DP nessa classe de amplitude, o efeito da cheia é significativo.

Pode-se dizer que o estudo de descargas parciais através de Modelos Lineares Generalizados para dados de contagem se mostrou bastante viável para o propósito deste trabalho, embora devido à grande variabilidade dos dados alguns dos pressupostos do modelos avaliados foram prejudicados.

Além disso, é notório que os resultados obtidos neste trabalho são consonantes ao preconizado por Wang *et.al.*(2016), que em seus estudos comprovam cientificamente que o alto índice da umidade relativa do ar modifica sobremaneira a forma como as DPs se apresentam, promovendo uma diminuição da amplitude de inicialização da tensão das descargas, todavia, por mais que ocorram tardiamente neste ambiente em detrimento do seco, pela natureza eletronegativa das moléculas de águas oriundas do ambiente úmido aumentarem o “coeficiente efetivo de ligação”, reduzindo assim a eficiência de ionização na região, o que leva à redução de elétrons livres disponíveis causadores das descargas, e aumenta o campo elétrico ao redor da superfície isolante, aumentando também a condutividade desta superfície, fato que influencia diretamente na mobilidade das cargas depositadas, espalhando-as em área maior neste local, acarretando uma elevação do número de pulsos de descargas parciais.

Nas medições na cheias, período de janeiro a julho, onde se tem um maior volume de chuvas e conseqüentemente a umidade elevada, houve o aumento de 1264,1% em relação ao período de seca (julho a dezembro), das DPs em amplitudes compreendidas entre -100mV e 100mV, tidas como pequenas em relação ao universo de -800mV a 800mV, no acoplador esquerdo e 737,8% no sensor direito, no mesmo intervalo. Este fato confirma que o ciclo hidrológico exerce grande influência no comportamento das descargas parciais aferidas no hidrogerador de interesse, que pertence a Usina Hidrelétrica de Tucuruí, no estado do Pará.

Para dar seguimento a este estudo serão utilizados os modelos de contagem considerando a superdispersão de dados tais como o modelos de contagem Gama e o modelo semi-paramétrico Quasi-Poisson, bem como tentar caracterizar os tipos de padrão de defeitos das descargas estudadas.

REFERÊNCIAS

- ALVES, M. P. *Caracterização de padrões de descargas parciais em hidrogeradores utilizando técnicas de inteligência Computacional*. Dissertação do Programa de Pós-Graduação em Engenharia Elétrica. UFPA, 2015.
- ATLAS de energia elétrica do Brasil. 3. Ed. Brasília: ANEEL, 2008.
- BRASIL, F. S. *Experimental de técnicas de medição de descargas parciais nos enrolamentos estáticos de hidrogeradores*. Dissertação de Mestrado – Programa de Pós-Graduação em Engenharia Elétrica. UFPA. 2013.
- BRASIL, F. S. *Desenvolvimento de Sensores Eletromagnéticos para Medição de Descargas Parciais nos Enrolamentos Estáticos de Hidrogeradores*. Tese de Doutorado – Programa de Pós-Graduação em Engenharia Elétrica. UFPA. 2016.
- BUSSAB, O. W.; MORETTIN, A. P. *Estatística Básica*, 5ª edição, São Paulo: Ed. Saraiva, 2005.
- CAMPBELL, S. R et al. Characteristics of Partial Discharge Pulses from Operating Rotating Machines. Conference Record of the IEEE International Symposium on Electrical Insulation. Pittsburgh-USA. 1994.
- CONSEIL INTERNATIONAL DES GRANDS RÉSEAUX ÉLECTRIQUES (CIGRÉ). 392 *Survey of Hydrogenerator Failures* Working Group A1.10, 2009.
- FERREIRA, F. *Uma abordagem clássica e Bayesiana para o Modelo Tweedie*. Dissertação de Mestrado do Programa de Pós-Graduação em Estatística. UFSCar, 2005.
- FRUTH, B. A; GROSS D. W. Partial discharge signal generation transmission and acquisition. *IEE Proc. Sci. Meas. Technol.*, vol. 142, (1), 1995.
- GOMES, F. J. S. *Análise de descargas parciais por meio de modelos para dados de contagens*. Dissertação de mestrado do Programa de Pós-Graduação em Matemática e Estatística. UFPA, 2016.
- GROSS, D. W.; FRUTH, B. A. Characteristics of Phase Resolved Partial Discharge Pattern in Spherical Voids. Conference on *Electrical Insulation and Dielectric Phenomena*, Atlanta, GA, USA 1998.
- GUILLEN, D.; IDARRAGA-OSPINA, G.; MOMBELLO, E.; CABRAL, S. Partial discharges location in transformer winding using wavelets and Kullback–Leibler divergence. *Electric Power Systems Research*, Vol. 136, pp.398-405, 2016.

- GURURAJ, B. e SATISH. L. Use of hidden Markov models for partial discharge pattern classification. *IEEE Transactions on Electrical Insulation*, vol.28 (2), pp. 172-182, 1993.
- IEC 60270. *High-voltage test techniques – Partial discharge measurements*. 2000.
- JORGENSEN, B.; SOUZA, M. C. P. Fitting tweedies.s compound poisson. *Model to insurance claims data*. Informes de Matemática, IMPA, Série B-072/92, 1992.
- MA, H; CHAN, J. C.; SAHA, T. K; EKANAYAKE C. Pattern Recognition Techniques and Their Applications for Automatic Classification of Artificial Partial Discharge Sources. *IEEE Transactions on Dielectrics and Electrical Insulation* vol. 20 (2), pp. 468-478, 2013.
- MAS'UD, A. A. STEWART, B.G; McMEEKIN, S.G. Application of an ensemble neural network for classifying partial discharge patterns. *Electric Power Systems Research*, vol.110, pp.154-162, 2014.
- OPERADOR NACIONAL DO SISTEMA ELÉTRICO. Disponível em: <<http://www.ons.org.br/pt/paginas/energia-agora/carga-e-geração>> Acesso em 24 de novembro de 2017.
- RAMOS, T.S.B.N; PROTÁZIO, J.M.B, MACIEL, M. C.; OLIVEIRA, T. F. ; SILVEIRA, A.M.S. *Séries Temporais Fuzzy aplicada na predição da vazão afluente em reservatório de hidrelétricas: um estudo de caso*. In: XV EMR-Escola de Modelos de Regressão, 2017, Goiânia. Anais da XV EMR-Escola de Modelos de Regressão, 2017.
- RAYMOND W. J. K, ILLIAS H. A., BAKAR A.H.A., MOKHLIS H. Partial discharge classifications: Review of recent progress. *Measurement* vol.68, pp.164–181,2015.
- SATHIYASEKAR K.; THYAGARAJAH, K. ; KRISHNAN, A. Neuro fuzzy based predict the insulation quality of high voltage rotating machine. *Expert Systems with Applications* vol.38 (1), pp. 1066–1072, 2011.
- SU.S.; CHIA C.C.; CHEN C.Y.; Chen J.F. Classification of partial discharge events in GILBS using probabilistic neural networks and the fuzzy c-means clustering approach. *Electrical Power and Energy Systems* , vol.61, pp.173–179,2014.
- SUWARNO; MIZUTANI T. Diagnosis of Insulation Conditions: Interpretation of Partial Discharges from ϕ -q-n pattern, Pulse-Sequence and Pulse Waveform. *International Conference on Condition Monitoring and Diagnosis*. China, 2008.
- The R foundation for statistical computing; *R Program. A Free Software for Statistical Computing*; Vienna, Austria, 2014.

WANG X.; TAYLOR N., EDIN, H. Effect of humidity on partial discharge in a metal-dielectric air gap on machine insulation at trapezoidal testing voltages. *Journal of Electrostatics*, vol.83, p.88-96, 2016.

YAWOOTTI, A.; INTRA, P.; TIPPAYAWONG, N.; RATTANADECHO, P.. An experimental study of relative humidity and air flow effects on positive and negative corona discharges in a corona-needle charger. *Journal of Electrostatics*, vol. 77, p.116–122, 2015.