



GLOBAL JOURNAL OF RESEARCHES IN ENGINEERING: F
ELECTRICAL AND ELECTRONICS ENGINEERING
Volume 23 Issue 4 Version 1.0 Year 2023
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4596 & Print ISSN: 0975-5861

Correlations between Meteorological Phenomena and Interruptions in the Supply of Electric Energy in a Power Supplier

By Marcos Antonio Lopes Freixo Filho & Ana Julia Righetto

Abstract- There are several technologies in the area of engineering, new and old, currently used with the objective of reducing the impacts of meteorological phenomena on electrical energy distribution networks. However, they still constitute one of the biggest challenges for companies around the world, the mitigation of the number of interruptions in the energy supply, in the face of the weather. This research work seeks something similar, which is to offer another analysis, from the point of view of data science, studying the correlation between the meteorological phenomena recorded by the Climatempo Monitoring and Alert System and the interruptions in the energy supply in two main cities of Maranhão State, with unknown causes, recorded by the control system used by the integrated operations center of the local power supplier, in an attempt to find some kind of fragility in the electrical system, not yet identified, that allows its improvement.

Palavras-chave: DEC, FEC, clima, transformadores de distribuição.

GJRE-F Classification: FOR Code: 0906



Strictly as per the compliance and regulations of:



© 2023. Marcos Antonio Lopes Freixo Filho & Ana Julia Righetto. This research/review article is distributed under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0). You must give appropriate credit to authors and reference this article if parts of the article are reproduced in any manner. Applicable licensing terms are at <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

Correlations between Meteorological Phenomena and Interruptions in the Supply of Electric Energy in a Power Supplier

Correlações entre Fenômenos Meteorológicos e Interrupções no Fornecimento de Energia Elétrica em uma Distribuidora

Marcos Antonio Lopes Freixo Filho ^α & Ana Julia Righetto ^ο

Resumo- Há diversas tecnologias no campo da engenharia, novas e antigas, utilizadas atualmente com o intuito de reduzir os impactos dos fenômenos meteorológicos nas redes de distribuição de energia elétrica. Porém, ainda assim constituem um dos maiores desafios para as companhias ao redor do mundo, a mitigação do número de interrupções no fornecimento de energia, frente às intemperes. Este trabalho de pesquisa busca algo semelhante, que é oferecer outra análise do ponto de vista da ciência de dados, estudando a correlação entre os fenômenos meteorológicos registrados pelo Sistema de Monitoramento e Alertas do Climatempo, e as interrupções no fornecimento de energia em duas principais cidades do Maranhão com causas desconhecidas, registradas pelo sistema de controle utilizado pelo centro de operações integrados da distribuidora de energia local, na tentativa de encontrar algum tipo de fragilidade no sistema elétrico, ainda não identificada, que permita o seu aprimoramento.

Palavras-chave: DEC, FEC, clima, transformadores de distribuição.

Abstract- There are several technologies in the area of engineering, new and old, currently used with the objective of reducing the impacts of meteorological phenomena on electrical energy distribution networks. However, they still constitute one of the biggest challenges for companies around the world, the mitigation of the number of interruptions in the energy supply, in the face of the weather. This research work seeks something similar, which is to offer another analysis, from the point of view of data science, studying the correlation between the meteorological phenomena recorded by the Climatempo Monitoring and Alert System and the interruptions in the energy supply in two main cities of Maranhão State, with unknown causes, recorded by the control system used by the integrated operations center of the local power supplier, in an attempt to find some kind of fragility in the electrical system, not yet identified, that allows its improvement.

Author α: Analista de Estratégia e Investimentos Sênior. Alameda A, 100 – Quitandinha; São Luís, Maranhão, Brasil.

e-mail: marcos.freixo@gmail.com

ORCID 0009-0002-4391-7583.

Author ο: Head in Statistics and Customer Experience, ALVAZ Agritech, Av. Ayrton Senna da Silva, Londrina- Paraná, Brasil.

ORCID is 0000-0003-4879-0710.

I. INTRODUÇÃO

O setor elétrico brasileiro, para que a energia possa chegar até os seus consumidores, passa por quatro etapas: a geração, a transmissão, a distribuição e a comercialização. A distribuição é responsável pelo varejo, vendendo a energia comprada em leilões e as recebendo pela transmissão. Este é um mercado regulado, com a participação de empresas públicas e privadas, de capital aberto, ou fechado. Segundo o Portal da Indústria (2021), existem no Brasil 131 concessionárias, permissionárias e cooperativas de distribuição de energia elétrica.

A Agência Nacional de Energia Elétrica [ANEEL], é o órgão regulador do mercado de energia elétrica. Dentre muitas de suas atribuições, cabe a ela regular a qualidade no fornecimento e expansão do sistema elétrico nacional. Para medir a qualidade, ela utiliza dois indicadores regulatórios que são metas estabelecidas à todas as distribuidoras do país.

O [DEC], sigla para Duração Equivalente de Interrupção por Unidade Consumidora, mede a duração equivalente de interrupção por unidade consumidora, ou o tempo que, em média, no período de observação, cada unidade consumidora ficou sem energia elétrica. Já o [FEC], sigla para Frequência Equivalente de Interrupção por Unidade Consumidora, se refere a frequência equivalente de interrupção por unidade consumidora, o número de interrupções ocorridas, em média, no período de observação. É de grande importância para as distribuidoras atender a estas regras regulatórias por apresentar um risco financeiro, receber ocasionalmente uma multa pela má gestão destes indicadores, como o caso de uma distribuidora de Minas Gerais, que recebeu um auto de infração em mais de doze milhões de reais, como informou o Canal Energia (2018).

Dentre os fatores que podem levar a uma interrupção no fornecimento de energia, temos o baixo investimento e manutenção do sistema elétrico, acidentes quaisquer com a rede de distribuição, ou mesmo na transmissão, o que normalmente é

expurgado dos resultados das distribuidoras de energia nos cálculos da agência reguladora, e os fenômenos climáticos, ou meteorológicos que por ventura derrubem o sistema, como descargas elétricas, ventanias, entre outros fatores.

Segundo Silva (2013), em nosso país, aproximadamente 70% dos desligamentos na transmissão e 40% na distribuição, são provocados por raios, sendo o número de transformadores queimados por descargas elétricas atmosféricas em torno de 40%. Tais números causam um impacto considerável no fornecimento de energia, o que pode ser constatado pela alta correlação entre a incidência de descargas elétricas e os índices de qualidade da maioria das empresas do setor elétrico.

No trabalho de Wronski et al (2003), após a utilização de modelos estatísticos, foram encontradas evidências de maior número de atendimentos e maior tempo médio de atendimento nas chamadas emergenciais de interrupção no fornecimento para a condições meteorológicas de vento e chuva, enquanto houve um menor número para a condição temperatura.

A distribuidora de energia do Maranhão, que durante anos permaneceu entre as três melhores empresas do país no quesito qualidade no fornecimento de energia, despencou em 2021 para os últimos lugares, de acordo com o último ranking divulgado pela [ANEEL] (2022). A empresa acredita que isto não se deve aos investimentos realizados na rede, que na verdade, ano após ano são cada vez maiores, mas se deve a um ano atípico de muitos registros de

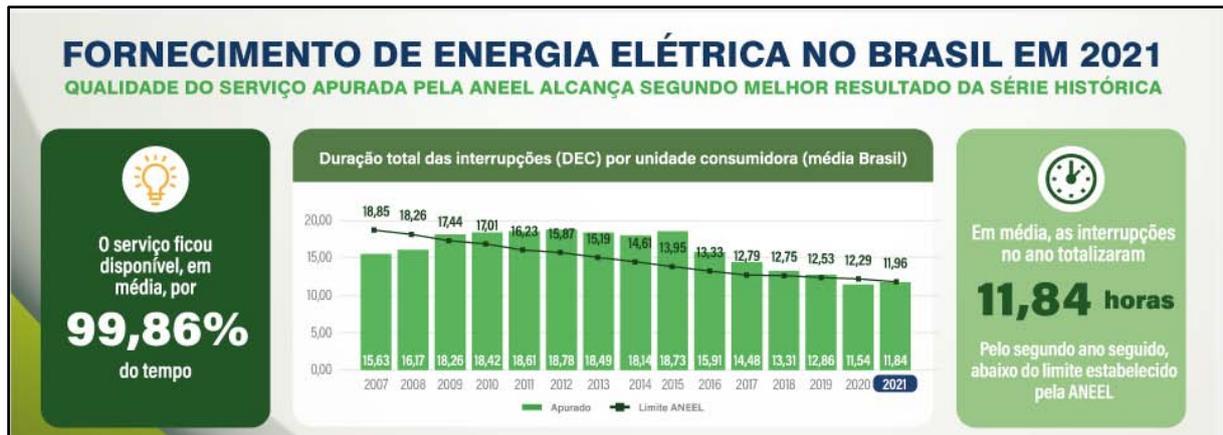
fortes precipitações, enchentes, ventanias, descargas elétricas atmosféricas, entre outros fenômenos climáticos.

Este trabalho tem como objetivo verificar e quantificar a correlação entre os fenômenos meteorológicos e a quantidade de ocorrências de interrupção no fornecimento de energia elétrica registradas pelo centro de operações integradas da distribuidora energética do Estado do Maranhão, para os municípios de São Luís e Imperatriz.

II. MATERIAL E MÉTODOS

O princípio que irá nortear este trabalho de pesquisa, a formar a base de dados final para estudos, ou seja, após a fase de “data wrangling”, é que esta base de dados reflita muito bem o fenômeno a ser estudado, que é a interrupção no fornecimento de energia, que seja capaz de explicar o preceito principal de que, apesar das dificuldades enfrentadas particularmente no país, em suma, o fornecimento de energia elétrica passa mais tempo em continuidade do fornecimento, do que interrompido, ou suspenso.

Para se ter uma dimensão sobre tal fato, o indicador do [DEC] em 2021, valor apurado na média para todo o Brasil, segundo o Governo Federal (2022), ficou em 11,84 horas. Em outras palavras, de um total de 8.760 horas do ano de 2021, em 99,86% do tempo a energia esteve em pleno fornecimento, sem interrupção, no Brasil inteiro, como pode ser constatado na Figura 1.



Fonte: Governo Federal (2022)

Figura 1: Recorte do Resultado do [DEC] em 2021

Ainda que o resultado para o Maranhão tenha ficado acima deste valor, registrando 29 horas, estando 99,67% do tempo em pleno fornecimento, conforme informado pelo distribuidora de energia em seu site de relacionamento com os investidores, em março de 2022, este número continua apontando para um fato supra relevante: a base de dados deve dar alicerce para realizar o estudo de um fenômeno raro, ou seja,

que além dos dados sobre as interrupções no fornecimento de energia, a base deve possuir cada hora em que o fenômeno não aconteceu, o período em que o fornecimento de energia não foi interrompido.

Seguindo esta premissa, este trabalho de pesquisa experimental, vem trabalhar a hipótese da existência de uma correlação entre os fenômenos meteorológicos, e a quantidade de vezes em que o

sistema elétrico teve sua distribuição interrompida, porém somente dos dados sinalizados como causa desconhecida, atendida por transformadores de distribuição, em dois municípios do Estado do Maranhão.

O centro de operações integradas[COI], setor responsável por manter o sistema de distribuição de energia elétrica em pleno atendimento para a população maranhense, possui em seus controles, uma base de registros de interrupções do fornecimento de energia, e seus equipamentos afetados, chamados de componentes, com informações que caracterizam em detalhe a sua tipificação, informações a princípio muito relevantes para esta pesquisa para verificarmos se há um padrão de componentes mais suscetíveis, ou não, às intempéries climáticas.

Com autorização da empresa, foram colhidos dados dos últimos doze meses, compreendidos entre agosto de 2021 e julho de 2022, somente dos equipamentos de transformadores de distribuição, já que toda interrupção de energia, que afete o bem-estar dos consumidores, acaba passando por este tipo de equipamento, segundo o especialista do próprio setor.

Outro fato importante é que, com a ajuda deste especialista, foram excluídas desta amostra de dados, todas as interrupções de energia derivadas da intervenção humana, como acidentes de trânsito, desligamentos programados, sobrecargas do equipamento causadas pelo consumo, não realização de poda de árvores abaixo da rede elétrica, para citar alguns, o que deixa a base de informações mais limpa para a realização deste estudo, já que todas as interrupções citadas têm a sua causa conhecida, e registrada no sistema.

Coletada a primeira parte dos dados, a segunda parte foi fornecida junto com a área conhecida como geoprocessamento, e lá foram coletadas as informações a respeito do clima, para as mesmas cidades. A companhia de energia possui um contrato de prestação de serviços com a empresa Climatempo, e através do Sistema de Monitoramento e Alertas do Climatempo, com a autorização da distribuidora, foram coletadas também informações do mesmo período da base coletada anteriormente, referentes a pluviometria, descargas elétricas atmosféricas, rajadas de vento, velocidade média do vento ao longo do dia, temperatura média, máxima e mínima.

As variáveis possuem precisões diferentes para relatar o período em que aconteceram, variando entre a precisão da hora, minuto e segundo em que ocorreram, como os dados de interrupções de energia, e os dados das descargas elétricas atmosféricas, e até informação que trata de resultado válido pelo dia inteiro, como velocidade do vento, e as informações sobre temperatura. Todas foram colocadas na mesma escala de tempo, demonstrando os fatos ocorridos a cada uma hora do dia, durante 365 dias. Assim, foram

criadas quatro variáveis para identificar o tempo: ano, mês, dia e hora.

Um fato importante é que tal base neste momento contém apenas informações relativas às interrupções, e como estava o clima naquele exato momento, entretanto, como já mencionado, ainda que uma distribuidora no território nacional tenha os piores indicadores de continuidade no fornecimento de energia, mesmo assim o sistema elétrico desta distribuidora hipotética estaria em pleno funcionamento na esmagadora parte do tempo. Desta forma, foram reproduzidos os dados que representam o período em que a energia estava em plena distribuição, para cada tipo diferente de transformador de distribuição, trazido na amostra de interrupções, com o mesmo período de tempo da amostra, em dias e horas, e devidamente cruzada com a base climatológica, para se conhecer as exatas condições do tempo neste mesmo exato período. Para diferenciar as duas informações na base de dados, foi criada uma variável chamada de "interrupções", do tipo numérica, atribuindo a quantidade "1" para as observações relativas a uma interrupção no fornecimento de energia, e "0" para as observações que representam o período em que a energia estava em pleno funcionamento.

Cabe salientar que foram excluídas desta base de dados de observações em que a energia estava em plena distribuição, as observações que coincidem temporalmente com a base de interrupções, afinal, exatamente neste exato período, a energia fortuitamente estava suspensa em seu fornecimento. Também, no intuito de explorar uma outra possibilidade que pode ser relevante para a pesquisa, foi criada uma variável chamada "raio", do tipo binária, para indicar a presença de raios naquela exata hora relada atribuindo o valor 1, e 0 quando o fato não havia ocorrido, o que é diferente da variável "descargas_atm", que apresenta a quantidade de descargas elétricas atmosféricas dentro daquela mesma observação.

Finalmente com as duas bases de dados unidas (com e sem interrupções), encerrando a etapa de "data wrangling" da coleta de dados, para que se garanta a aleatoriedade da informação, quanto ao objeto desta pesquisa experimental, e permitindo analisar se este estudo se trata de um caso em que o melhor modelo estatístico deva levar em conta o estudo de um fenômeno raro, considerando o fato de que energia elétrica passa mais tempo em plena distribuição, do que suspensa.

Por fim, em razão do tamanho da base de dados para este estudo, se fossem considerados todos os municípios do Estado do Maranhão, ainda que utilizando dados relativos ao período de apenas um ano, foram escolhidos apenas dois dos principais municípios do Estado: São Luís, e Imperatriz. O estudo poderia se limitar a capital do Estado (São Luís), por seu tamanho e importância. Todavia, a cidade de

Imperatriz possui características completamente diferentes, no que se diz respeito ao clima, relevo, economia, ou outras não citadas, o que torna ainda mais interessante checar neste estudo, se há semelhanças nos resultados obtidos entre os dois municípios, o que poderia levantar outras hipóteses, para outros trabalhos futuros.

III. MODELO DE REGRESSÃO DE POISSON

A metodologia estatística que foi adotada neste estudo foi a aplicação do modelo de regressão de Poisson. Os modelos de regressão Poisson e o modelo de regressão binomial negativo fazem parte do que é conhecido por modelos de regressão para dados de contagem, que têm por objetivo analisar o comportamento, em função de variáveis preditoras, de

Expressão do modelo de Poisson:

$$\ln(\hat{Y}_i) = \ln(\lambda_{\text{poisson } i}) = \alpha + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \dots + \beta_k \cdot X_{ki} \quad \text{form. (1)}$$

Em que $\ln(\hat{Y}_i)$ ou $\ln(\lambda_{\text{poisson } i})$ é definido a partir de uma função de ligação canônica que é proposta com base na definição dos logaritmos das funções de verossimilhança oriunda da função densidade/probabilidade da distribuição Poisson, α representa os termos do intercepto, β_k são os coeficientes de cada variável explicativa e correspondem aos parâmetros a serem estimados, X_{ki} são as variáveis explicativas.

Fórmula da probabilidade da distribuição Poisson:

$$p(Y_i = m) = \frac{e^{-\lambda_i} \lambda_i^m}{m!} \quad \text{form. (2)}$$

Em que $p(Y_i = m)$ é a função da probabilidade de um evento, no modelo de regressão de Poisson, determinada por uma observação i ($i = 1, 2, \dots, n$), "n" é o tamanho da amostra, e possui a probabilidade de ocorrência de uma contagem "m", em uma

Expressão do modelo Binomial Negativo:

$$\ln(\hat{Y}_i) = \ln(\lambda_{\text{bneg } i}) = \alpha + \beta_1 \cdot X_{1i} + \beta_2 \cdot X_{2i} + \dots + \beta_k \cdot X_{ki} \quad \text{form. (3)}$$

Em que $\ln(\hat{Y}_i)$ ou $\ln(\lambda_{\text{bneg } i})$ é definido a partir de uma função de ligação canônica que é proposta com base na definição dos logaritmos das funções de verossimilhança oriunda da função densidade/probabilidade da distribuição Poisson Gama, ou Binomial Negativa, α representa os termos do intercepto, β_k são os coeficientes de cada variável explicativa e correspondem aos parâmetros a serem estimados, e X_{ki} são as variáveis explicativas.

Fórmula da probabilidade da distribuição Binomial Negativa:

$$p(Y_i = m) = \frac{\delta^\theta \cdot m^{\theta-1} \cdot e^{-m \cdot \delta}}{(\theta-1)!} \quad \text{form. (4)}$$

Em que $p(Y_i = m)$ é a função de probabilidade no modelo binomial negativo determinada por uma

uma determinada variável dependente que se apresenta na forma quantitativa, com valores discretos e não negativos, exatamente como o objeto de estudo deste trabalho, onde a princípio não é possível fracionar o evento "interrupção do sistema elétrico" (não considerase neste estudo os casos em que a energia esteve em oscilação quanto ao fornecimento), tendo sempre sua quantificação como números discretos e positivos.

Considerando o que afirma Fávero e Belfiore (2022, p.149), ao mencionar que no caso de ocorrência de fenômenos raros, com baixa probabilidade de sucesso ($p \rightarrow 0$), sob determinada exposição (unidade temporal, espacial, social, etc.), como no caso desta pesquisa experimental, em determinado intervalo de tempo, o modelo de distribuição Poisson se torna o prioritário a ser explorado.

determinada exposição (período, área, região, entre outros exemplos), em que "λ" é o número esperado de ocorrências ou a taxa média estimada de incidência do fenômeno em estudo para uma dada exposição

A título de comparação de melhor modelo, também foi aplicado neste estudo o modelo de regressão binomial negativo, também conhecido como Poisson Gama, outro modelo de contagem aqui citado. A diferença básica entre os dois modelos, é que o modelo binomial negativo é mais utilizado para os casos de superdispersão, ou inflação de zeros da amostra, onde a variância da variável dependente é maior que sua própria média. Segundo Fávero e Belfiore (2022, p.696), a superdispersão é comumente gerada pela presença de maior heterogeneidade nos dados entre observações da amostra.

observação i ($i = 1, 2, \dots, n$), "n" é o tamanho da amostra, onde θ é chamado de parâmetro de forma e deve ser maior que zero, e δ é chamado de parâmetro de taxa de decaimento, também maior que zero.

A hipótese nula para este estudo será a baixa ou nenhuma correlação entre os fenômenos meteorológicos e as interrupções no fornecimento de energia elétrica em uma distribuidora, abrindo a possibilidade de estudar outros fatores, como a fadiga dos equipamentos, vida útil da rede de distribuição, ou mais fatores, que possam ter preponderância maior na suspensão ocasional da distribuição energética de eletricidade.

Ainda, foram feitas análises de correspondências múltiplas, que se trata de uma técnica multivariada que possibilita a investigação de

associação com mais de duas variáveis categóricas, conforme Fávero e Belfiore (2022, p.463), no intuito de perceber a relação de interdependência entre as variáveis utilizadas neste estudo, trabalhando somente com as observações onde ocorreram as interrupções do fornecimento de energia elétrica, nas duas amostras os dados.

As técnicas de análise de correspondência são métodos de representação de linhas e colunas de tabelas cruzadas de dados, como coordenadas em um gráfico, chamado de mapa perceptual, a partir do qual se podem interpretar as similaridades e diferenças de comportamento entre variáveis, e entre categorias.

Essa técnica tem como principal objetivo avaliar a significância dessas similaridades, reduzindo suas dimensões, de modo a determinar coordenadas das categorias com base na distribuição dos dados em tabelas cruzadas para, a partir dessas coordenadas, construir mapas perceptuais, que nada mais são que diagramas de dispersão que representam as categorias das variáveis na forma de pontos em relação a eixos de coordenadas ortogonais, fazendo delas, na realidade, mapas de categorias.

O método para realizar a análise consiste em transformar as variáveis qualitativas em variáveis binárias, obtendo daí uma matriz binária. Com base nesta matriz binária, é obtido a inércia principal total na [ACM]. Supondo que a matriz binária seja semelhante a uma tabela de contingência da análise de correspondência, é possível obter a inércia principal parcial das dimensões, seus autovalores, autovetores, e posterior coordenadas desta matriz.

Outro método alternativo é a combinação em uma única matriz, as tabelas de contingências com os cruzamentos de todos os pares de variáveis. Essa matriz resultante, quadrada e simétrica, é conhecida por matriz de Burt, como explica Baltar (2005).

Só participaram da [ACM] as variáveis que apresentaram associação estaticamente significativa, com pelo menos uma variável contida na análise, por meio do teste do X^2 . Caso alguma delas não apresentasse associação com outras, estas ficaram excluídas da análise de correspondência. Ao elaborar o mapa perceptual, são representadas no mapa as

coordenadas das dimensões que apresentam a inércia principal parcial maior do que a média da inércia principal total por dimensão.

A aplicação dos modelos estatísticos se deu através da linguagem R, utilizando o software RStudio versão 1.4.1106. A base final foi exportada para o formato de planilha, e tratada em banco de dados SQL Server.

IV. RESULTADOS E DISCUSSÃO

a) São Luís

Foram carregados os dados da amostra de São Luís, coletada na distribuidora de energia no R com um total de 272.151 observações e 12 variáveis. Os dados podem ser vistos através da função "glimpse()", conforme apêndice.

Para reforçar essa possibilidade conceitual, de que o objeto de estudo deste trabalho de pesquisa, se trata de um fenômeno raro, foi realizado o diagnóstico preliminar para observação de eventual igualdade entre a média e a variância da variável dependente INTERRUPCAO, cujos resultados foram respectivamente 0,0180268 e 0,0177019, valores muito próximos, o que indica fortemente que de fato, o modelo estatístico mais indicado para explicar a contagem de interrupções de energia elétrica da amostra de dados de São Luís, seja um modelo de regressão Poisson.

Utilizando a função "glm()", fez-se a estimação deste modelo, colocando a variável INTERRUPCAO como a variável dependente, e as outras onze variáveis, já citadas, como variáveis explicativas.

O primeiro resultado apurado com base no modelo de regressão de Poisson, confirma a hipótese desta pesquisa experimental, de que existe de fato uma correlação entre os fenômenos meteorológicos e as vezes em que o fornecimento de energia elétrica foi interrompido aos consumidores de São Luís, pois pelo menos um beta de uma variável com informações da meteorologia no momento em que acontece a interrupção, teve o seu p-valor menor que 0.05, ou seja, podemos admitir o modelo estatísticos perfeitamente correlacionado, para fins preditivos, como vê-se na Tabela 1, através da função "summary()".

Tabela 1: Summary do Modelo Poisson da base Coletada de São Luís

Variáveis	Estimativa	Erro padrão	Valor de Z	Pr(> z)
(Intercepto)	0,7762	0,6353	1,2220	0,2218
FASESABC	-0,3258	0,4756	-0,6850	0,4934
FASESAC	-2,0067	0,5547	-3,6170	0,0003
FASESB	-4,9644	0,5544	-8,9540	<2e-16
KVA10	-5,8039	0,4141	-14,0140	<2e-16
KVA1000	-5,3779	0,4167	-12,9070	<2e-16
KVA112,5	0,0339	0,0417	0,8140	0,4158

KVA15	-3,0510	0,1446	-21,0960	<2e-16
KVA150	-0,4834	0,0484	-9,9800	<2e-16
KVA1500	-7,1696	1,0024	-7,1520	0,0000
KVA1750	-6,0710	0,5833	-10,4080	<2e-16
KVA2000	-6,4765	0,7120	-9,0960	<2e-16
KVA225	-2,9794	0,1396	-21,3490	<2e-16
KVA25	-7,1520	1,0015	-7,1410	0,0000
KVA250	-7,1690	1,0024	-7,1520	0,0000
KVA2600	-7,1696	1,0024	-7,1520	0,0000
KVA30	-2,9908	0,1431	-20,9010	<2e-16
KVA300	-3,9313	0,2119	-18,5510	<2e-16
KVA3000	-7,1696	1,0024	-7,1520	0,0000
KVA350	-6,4765	0,7120	-9,0960	<2e-16
KVA3500	-7,1690	1,0024	-7,1520	0,0000
KVA45	-1,2695	0,0760	-16,7130	<2e-16
KVA5	-5,9503	0,4556	-13,0590	<2e-16
KVA500	-4,6847	0,3004	-15,5930	<2e-16
KVA612,5	-6,4765	0,7120	-9,0960	<2e-16
KVA725	-6,4765	0,7120	-9,0960	<2e-16
KVA75	-0,1176	0,0463	-2,5430	0,0110
MEDIA_CLIENTES	-0,0006	0,0004	-1,3650	0,1722
PLUVIOMETRIA	0,0154	0,0025	6,0680	0,0000
TEMP_MIN	-0,0720	0,0149	-4,8420	0,0000
TEMP_MAX	-0,0200	0,0132	-1,5130	0,1303
TEMP_MEDIA	NA	NA	NA	NA
DESCARGAS_ATM	0,0015	0,0007	2,0870	0,0369
RAIO	0,3236	0,0633	5,1090	0,0000
RAJADA	0,0042	0,0010	4,0830	0,0000
VENTOS_KMH	0,0021	0,0015	1,4400	0,1499

Fonte: Resultados originais da pesquisa

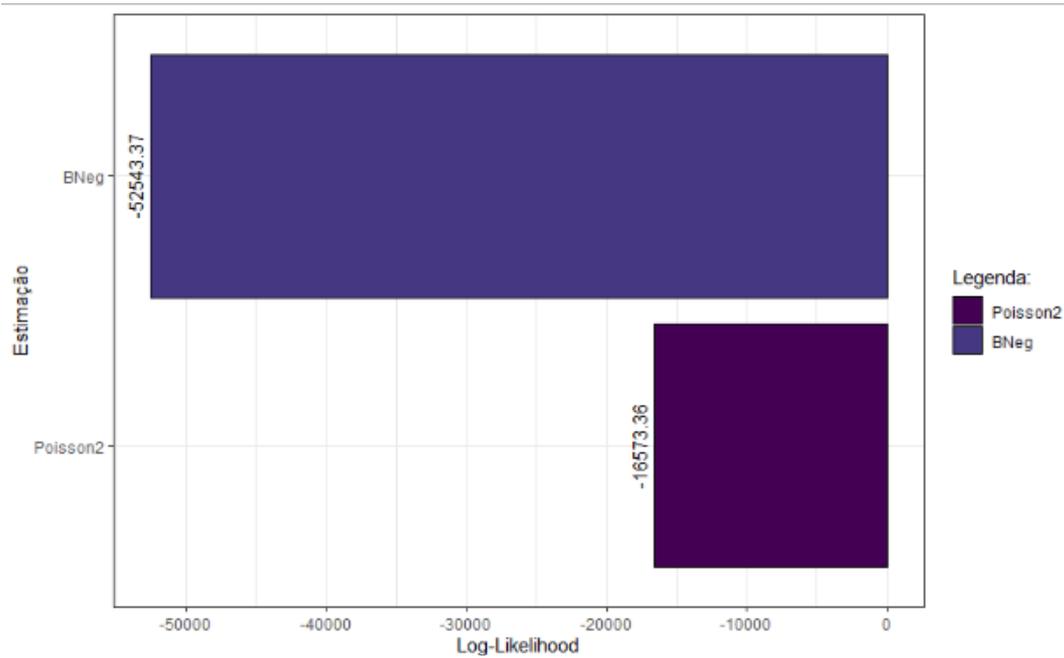
No intuito de evitar efeitos prováveis de multicolinearidade entre as variáveis meteorológicas que naturalmente tendem a ter resultados próximos, foi feito uma nova estimação do modelo, onde foram consideradas apenas as variáveis que foram estatisticamente significantes para o objeto deste estudo. Desta forma, foram retiradas as variáveis TEMP_MAX, TEMP_MEDIA, DESCARGAS_ATM, VENTOS_KMH, e atribuindo o nome de modelo_poisson2.

Para fins de comparação, e se aproximar da certeza de que o modelo estatístico não se trata de uma distribuição Binomial Negativa, corroborando com o diagnóstico preliminar que compara a média e a variância da base coletada de dados de São Luís, fez-se necessário carregar o modelo binomial negativo desta mesma base, utilizando a função "glm.nb()".

Segundo Bussab e Morettin (2005), o princípio da verossimilhança explica que se deve escolher aquele

valor do parâmetro desconhecido que maximiza a probabilidade de obter a amostra particular observada, ou seja, o valor que torna aquela amostra a "mais provável". Esse princípio foi enunciado por Ronald Aylmer Fisher pela primeira vez em 1912 e, em 1922, deu-lhe forma mais completa, introduzindo a expressão Log-Likelihood(verossimilhança).

Em outras palavras, o método de máxima verossimilhança consiste em estimar o parâmetro λ por uma amostra aleatória X_1, \dots, X_n independente e igualmente distribuída de X . Na Figura 2 é feita a comparação entre os dois modelos estatísticos, comparando o Log-Likelihood de ambos os modelos estatísticos aplicados na amostra de São Luís.



Fonte: Resultados originais da pesquisa

Figura 2: Comparação de Log Lik Entre os Modelos Poisson (2) e Binomial Negativo (Poisson Gama)

De acordo com Fávero e Belfiore (2022), quanto menor for o Log-Likelihood na comparação entre modelos estatísticos, melhor é o ajuste do modelo à distribuição da amostra. Assim, para este estudo, será considerado o modelo Poisson (modelo_poisson2) para explicar a contagem dos dados de interrupção de fornecimento de energia elétrica em São Luís do Maranhão.

Apenas para exemplificar a importância desta pesquisa experimental, tomando o modelo estatístico fundamentadamente adotado, através da função "predict", é possível estimar em aproximadamente 36,6%, a probabilidade de ocorrer uma interrupção do fornecimento de energia elétrica de um transformador de distribuição com três fases conectadas (ABC), de carga de 112,5Kva, sob uma chuva de 15 milímetros, com uma temperatura mínima de 22°C, na presença de raios, rajadas de vento de 30Km/h, conforme pode ser visto no apêndice deste trabalho.

Ainda, cabe neste trabalho explorar a relação que cada uma destas variáveis que explicam as interrupções no fornecimento de energia, na amostra utilizada no modelo de Poisson, se relaciona-se entre si. Para isto, foi gerada um segundo "dataframe", a partir da base denominada de final.

Em um novo "wrangling" dos dados, para que se possa realizar uma análise de correspondência, foram criadas novas variáveis para minimizar a quantidade de rótulos a serem exibidos num mapa perceptual que será feito com a análise de correspondência múltipla [ACM].

A variável PLUVIOMETRIA, foi substituída pela variável PLUV_TIPO, agora como uma variável

categórica, contendo uma classificação da chuva baseada na sua intensidade medida na quantidade de milímetros registrados naquela hora. A classificação da intensidade da chuva obedeceu aos critérios adotados pela Organização Meteorológica Mundial – OMM (2018), e para nossa amostra, se dividiu entre fraca, moderada e forte.

A variável RAJADA foi substituída pela variável RAJADA_TIPO, e suas categorias foram classificadas segundo a escala Beaufort, também disciplinada pela OMM, escala baseada na velocidade do vento registrada conforme Tabela 2. A variável VENTO foi excluída por possíveis relações de multicolinearidade na amostra com outras variáveis, como já mostrados neste trabalho.

A variável RAI0 foi mantida, mas para fins de exibição dos rótulos de dados no mapa perceptual da [ACM], os valores "0" e "1", foram substituídas respectivamente pelas categorias "RAIO_NAO" e "RAIO_SIM". A variável DESCARGAS_ATM foi excluída, sem substituta, por não ter sido encontrado quaisquer classificações advindas de instituições oficiais ou norma científica, nas pesquisas deste trabalho, que definissem a intensidade de descargas elétricas atmosféricas.

Tabela 2: Escala Beaufort, Limitada Até A Maior Velocidade Registrada Nas Amostras

Grau	Designação	m/s	km/h	Efeitos em terra
0	Calmo	<0,3	<1	Fumaça sobe na vertical
1	Aragem	0,3 a 1,5	1 a 5	Fumaça indica direção do vento
2	Brisa leve	1,6 a 3,3	6 a 11	As folhas das árvores movem; os moinhos começam a trabalhar
3	Brisa fraca	3,4 a 5,4	12 a 19	As folhas agitam-se e as bandeiras desfraldam ao vento
4	Brisa moderada	5,5 a 7,9	20 a 28	Poeira e pequenos papéis levantados; movem-se os galhos das árvores
5	Brisa forte	8 a 10,7	29 a 38	Movimentação de grandes galhos e árvores pequenas
6	Vento fresco	10,8 a 13,8	39 a 49	Movem-se os ramos das árvores; dificuldade em manter um guarda-chuva aberto; assobio em fios de postes
7	Vento forte	13,9 a 17,1	50 a 61	Movem-se as árvores grandes; dificuldade em andar contra o vento
8	Ventania	17,2 a 20,7	62 a 74	Quebram-se galhos de árvores; dificuldade em andar contra o vento; barcos permanecem nos portos
9	Ventania forte	20,8 a 24,4	75 a 88	Danos em árvores e pequenas construções; impossível andar contra o vento
10	Tempestade	24,5 a 28,4	89 a 102	Árvores arrancadas; danos estruturais em construções
11	Tempestade violenta	28,5 a 32,6	103 a 117	Estragos generalizados em construções

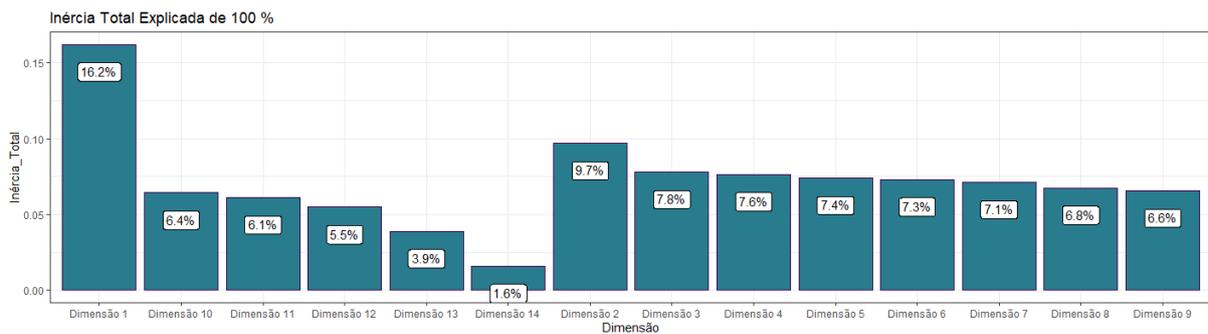
Fonte: Resultados originais da pesquisa

De forma semelhante, as três variáveis de temperatura foram substituídas por três novas variáveis categóricas, TEMP_MIN_FAIXA, TEMP_MAX_FAIXA e TEMP_MED_FAIXA, onde de forma arbitrária, por também não ter encontrado literatura semelhante às outras variáveis anteriormente citadas, que agrupam as temperaturas por faixa.

Por último, para fins demonstrativos, destinados a orientar estudos posteriores a este

trabalho, foram selecionados apenas três transformadores de distribuição na amostra, em razão de serem os mais utilizados pela Companhia, da igual forma, a fase escolhida a foi a “ABC”.

Através do comando “mca()”, obtemos a análise de correspondência múltipla da amostra. Foram obtidas 15 dimensões, e a inércia total explicada se deu conforme exposto no Figura 3.



Fonte: Resultados originais da pesquisa

Figura 3: Inércia Total Explicada da [ACM] da Amostra

Por fim, para fins de exemplo de estudos que possam ser explorados pelas áreas de engenharia, com outros equipamentos, em outras regiões do Estado, ou mesmo em subestações, linhas de transmissão, tem-se o mapa perceptual da Figura 4.

Tabela 3: Summary do Modelo Poisson da base Coletada de Imperatriz

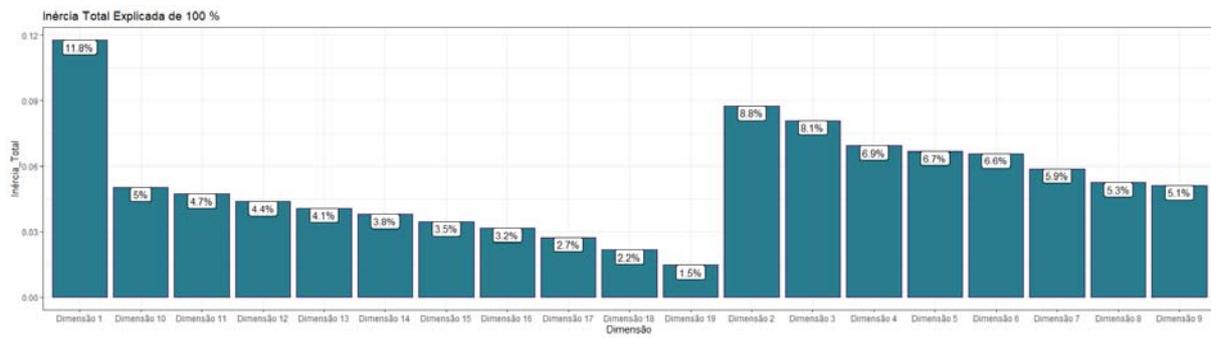
Variáveis	Estimativa	Erro padrão	Valor de Z	Pr(> z)
(Intercepto)	-6,6868	0,7542	-8,8660	<2e-16
FASESAB	0,6384	0,4976	1,2830	0,1996
FASESABC	1,1842	0,3659	3,2370	0,0012
FASESAC	0,2637	0,4207	0,6270	0,5308
FASESB	-1,0221	0,4310	-2,3710	0,0177
FASESBC	-0,0565	0,6097	-0,0930	0,9261
FADESC	-0,3513	0,5187	-0,6770	0,4982
KVA10	-3,2960	0,3605	-9,1420	<2e-16
KVA112,5	1,1633	0,1070	10,8690	<2e-16
KVA15	-1,9363	0,2319	-8,3490	<2e-16
KVA150	1,3345	0,1307	10,2060	<2e-16
KVA1500	-4,8743	1,0207	-4,7760	0,0000
KVA2000	-4,8743	1,0207	-4,7760	0,0000
KVA225	-2,3017	0,3194	-7,2070	0,0000
KVA30	-2,2306	0,3216	-6,9370	0,0000
KVA300	-3,4830	0,5302	-6,5690	0,0000
KVA3500	-4,8708	1,0162	-4,7930	0,0000
KVA45	-0,2392	0,1742	-1,3730	0,1698
KVA5	-2,6783	0,2943	-9,1000	<2e-16
KVA500	-3,2648	0,4937	-6,6130	0,0000
KVA600	-4,8743	1,0207	-4,7760	0,0000
KVA75	0,9328	0,1065	8,7580	<2e-16
MEDIA_CLIENTES	-0,0002	0,0014	-0,1340	0,8931
PLUVIOMETRIA	0,0083	0,0020	4,2340	0,0000
TEMP_MIN	0,0273	0,0222	1,2250	0,2204
TEMP_MAX	0,0311	0,0203	1,5300	0,1261
TEMP_MEDIA	-0,0253	0,0387	-0,6530	0,5137
DESCARGAS_ATM	-0,0010	0,0015	-0,6320	0,5275
RAIO	0,0320	0,1190	0,2690	0,7878
RAJADA	0,0163	0,0029	5,5700	0,0000
VENTOS_KMH	-0,0085	0,0044	-1,9390	0,0525

Fonte: Resultados originais da pesquisa

Também, para fins de comparação com o estudo anterior, o mesmo exemplo de “predict” feito para a mostra de São Luís, foi feito para a amostra de Imperatriz, com as mesmas variáveis, e o resultado foi bem diferente entre as duas cidades. Enquanto a probabilidade de uma interrupção no fornecimento de energia elétrica na primeira amostra foi de 36,6%, nas mesmas condições do tempo, e mesmas características de transformadores, o resultado foi uma probabilidade de apenas 5,16% do sistema elétrico ser interrompido pelos mesmos fatores (sete vezes menor).

Ao realizar a análise de correspondências múltiplas, mais diferenças são observáveis entre os dois resultados. Em razão das disparidades climáticas

entre as duas cidades, mais categorias foram adicionadas para os cálculos, gerando 19 dimensões para a análise, três dimensões a mais que na amostra de São Luís, conforme o Figura 5, com a inércia total explicada da análise de correspondência múltipla.



Fonte: Resultados originais da pesquisa

Figura 5: Inércia Total Explicada da [ACM] da Amostra (Imperatriz) -

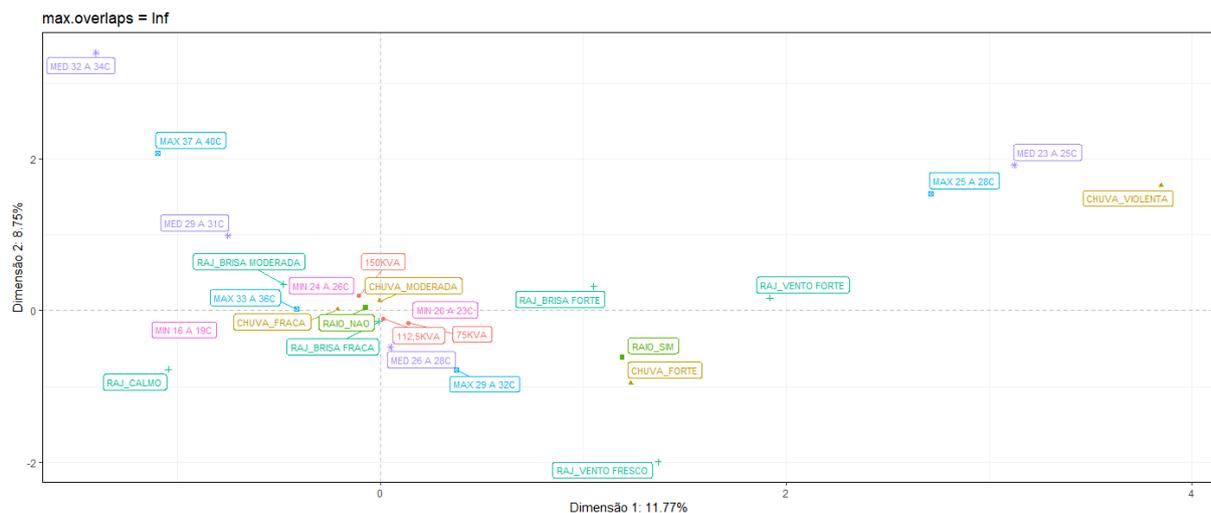
Assim como é minuciosamente explicado pelo site Weathers Park (2022), São Luís e Imperatriz possuem grandes diferenças de temperatura ao longo do ano, tendo Imperatriz como cidade com temperaturas mínimas registradas menores que a capital do Estado, e as temperaturas máximas registradas maiores que São Luís. Ainda de acordo com Weathers Park, há diferenças importantes quanto a velocidade dos ventos. São Luís rotineiramente registra maiores velocidades dos ventos do que na cidade de Imperatriz como descrito no referido portal.

De acordo com Lima e Ferreira (2013), separadas por aproximadamente 600km, uma da outra, as duas cidades de fato possuem climas bastante divergentes. São Luís, que fica ao norte do Estado, é uma cidade litorânea, mais sujeita a ventos fortes, e temperaturas mais equilibradas ao longo do ano, do que a cidade de Imperatriz, ao sul, com um tipo de clima misto, ora seco como o cerrado e ao sertão, ora chuvoso como o clima amazônico, por fazer fronteira entre estes dois biomas, além de ventos mais fracos, e temperaturas mais extremas do que na capital.

Ao analisar o mapa perceptual da amostra de Imperatriz na Figura 6, é interessante perceber que a chuva violenta (com mais de 50mm/h) tem pouquíssima relação com as interrupções no fornecimento de energia de Imperatriz. Isto pode mostrar o quanto o sistema elétrico na região foi bem preparado pela distribuidora para suportar as fortes intempéries do período chuvoso.

Já a chuva moderada (entre 2,5 a 10 mm/h) se mostra muito mais relevante para a interrupção do fornecimento de energia do que em São Luís. A rajada de brisa forte, que em São Luís se mostra importante, em Imperatriz é a rajada de brisa fraca (ventos repentinos de 12 a 19km/h) que se mostra mais preponderante.

Tal mapa pode auxiliar as áreas de engenharias a encontrar soluções que mitiguem a quantidade de suspensões do fornecimento de energia elétrica, segundo as condições apresentadas.



Fonte: Resultados originais da pesquisa

Figura 6: Mapa Perceptual da [ACM] da Amostra Para Exemplo de Estudo (Imperatriz)

V. CONCLUSÃO

É possível afirmar, através deste trabalho de pesquisa realizado com dados reais de uma distribuidora de energia elétrica do Estado do Maranhão, com duas de suas principais cidades, que há de fato, com um nível de significância de 5%, fortes correlações associadas aos fenômenos meteorológicos e as interrupções no fornecimento de energia, com causas desconhecidas, para ambas as amostras, com ressalvas que indicam que o estudo precisa ser de fato específico por região, e que, para fins preditivos, o modelo de distribuição de Poisson se encaixa perfeitamente.

Ainda, o estudo possibilita às distribuidoras, que possam trabalhar de modo preventivo, identificando as fragilidades do sistema elétrico quanto ao tema, e explorar ações que possam reduzir a frequência com que acontecem as interrupções, analisando não só transformadores de distribuição, mas equipamento por equipamento, ou até mesmo estruturas mais robustas, como subestações, através modelo de Poisson, e de mapas perceptuais de análise de correspondência múltipla, de modo a melhor o indicador do [FEC].

Por fim, também é possível criar um sistema de alertas que envie para o Centro de Operações Integradas, que tipo de equipamento, onde, e a que probabilidade, este possa ter sua distribuição de energia interrompida, de acordo com a previsão do tempo, combinado ao Sistema de Monitoramento de Alertas do ClimaTempo, com previsões em tempo real, podendo antecipadamente ao fenômeno, deslocar equipes de plantão a lugares estratégicos, para que estas possam tomar ações que reestabeçam o fornecimento de energia com mais celeridade. Desta forma, melhorando também o indicador do [DEC].

Por fim, este estudo sugere que outras variáveis sejam acrescentadas para um estudo semelhante, como a vida útil da rede, transformador de distribuição, ou o que queira ser estudado. Também seria interessante acrescentar o fabricante, e lote adquirido, para identificar se uma linha de materiais comprados pela distribuidora não possui mais fragilidade à fenômenos meteorológicos.

AGRADECIMENTO

A Deus por ter me dado toda a força deste mundo para chegar até aqui, com tantas pedras pelo caminho, incertezas que quase me levaram a desistir, dificuldades financeiras, mas principalmente a difícil arte de revezar o tempo com a família, amigos e trabalho, tarefa das mais complicadas que já passei nos últimos tempos. Só eu e Ele sabíamos o que estávamos construindo. Agradecer em especial aos amigos Evandro Lima Meireles, da Gerência Corporativa de Geoprocessamento, e Eliezer Silva de

Oliveira, da Gerência Corporativa de Operações, que com seus esforços na fase de mineração de dados, contribuíram enormemente com este trabalho, além também das pessoas que de alguma forma compreenderam este projeto de vida, e minimamente colaboraram para a conclusão deste trabalho.

REFERENCES RÉFÉRENCES REFERENCIAS

1. BALTAR, V. T. Análise fatorial múltipla para tabelas de contingência. 2005. Tese de Doutorado. Universidade de São Paulo.
2. BUSSAB, W.O.; MORETTIN, P.A. Estatística Básica. 5. ed. São Paulo: Saraiva, 2005.
3. CanalEnergia. Cemig é multada por indicadores de qualidade. Disponível em: <<https://www.canalenergia.com.br/noticias/53066819/cemig-e-multada-por-indicadores-de-qualidade>> Acesso em 09 de abr. de 2022.
4. FÁVERO, L.P.; BELFIORE, P. Manual de Análise de Dados. Rio de Janeiro: LTC, 2022.
5. Grupo Equatorial Energia [GEE]. Central de Resultados. Disponível em: <<https://ri.equatorialenergia.com.br/pt-br/divulgacao-e-resultados/central-de-resultados/>>. Acesso em 03 de ago. de 2022.
6. Imprensa Nacional [IM]. [ANEEL] divulga desempenho e ranking das distribuidoras sobre fornecimento de energia em 2021. Disponível em: <<https://www.gov.br/aneel/pt-br/assuntos/noticias/2022/aneel-divulga-desempenho-e-ranking-das-distribuidoras-sobre-fornecimento-de-energia-em-2021>> Acesso em 11 de abr. de 2022.
7. LIMA, R.M.B.F, FERREIRA, A.J.A. Estudos de Geografia do Maranhão. Maranhão. EDUFMA, 2013.
8. SILVA, J.C. A relação entre raios e chuvas nos arredores de Manaus: análise de casos extremos. Dissertação (Mestrado) – INPA/UEA, Manaus, AM. 2013.
9. WEATHERS PARK. Clima e condições meteorológicas médias em Imperatriz no ano todo. Disponível em: <<https://pt.weatherspark.com/y/30244/Clima-caracter%C3%ADstico-em-Imperatriz-Brasil-durante-o-ano>> Acesso em 15 de out. 2022.
10. WEATHERS PARK. Clima e condições meteorológicas médias em São Luís no ano todo. Disponível em: <<https://pt.weatherspark.com/y/30549/Clima-caracter%C3%ADstico-em-S%C3%A3o-Lu%C3%ADs-Brasil-durante-o-ano>> Acesso em 15 de out. 2022.
11. World Meteorological Organization [WMO]. Guide to Instruments and Methods of Observation: Volume I – Measurement of Meteorological Variables. Genebra, Suíça: Organização Meteorológica Mundial, no. 8, 7 ed. 2018. 618p. Disponível em: <https://library.wmo.int/doc_num.php?explnum_id=10616> Acesso em 04 de out. de 2022.

12. WRONSCKI, V.R. et al. METODOLOGIAS PARA IDENTIFICAR ASSOCIAÇÃO ENTRE PADRÕES CLIMÁTICOS E QUALIDADE DE FORNECIMENTO DE ENERGIA ELÉTRICA. In: Seminário Brasileiro sobre Qualidade da Energia Elétrica. Aracaju, SE. 2003.

Apêndice: Operações utilizadas no R Studio para a realização deste estudo

Base de informações do município de São Luís

```
> glimpse(base_slz) #visualização das observações e das especificações
Rows: 272,151
Columns: 14
$ INTERRUPTAO <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ~
$ FASES <chr> "ABC", "ABC", "ABC", "ABC", "ABC", "ABC", "ABC", "ABC", "AB~
$ KVA <chr> "0", "112,5", "112,5", "0", "0", "112,5", "112,5", "75", "1~
$ MEDIA_CLIENTES <dbl> 193, 68, 90, 151, 149, 131, 156, 207, 116, 288, 230, 108, 3~
$ PLUVIOMETRIA <dbl> 0.234375, 0.234375, 0.234375, 0.234375, 0.234375, 0.234375, ~
$ TEMP_MIN <dbl> 24, 24, 24, 24, 24, 24, 24, 24, 24, 24, 24, 24, 24, 24, ~
$ TEMP_MAX <dbl> 32, 32, 32, 32, 32, 32, 32, 32, 32, 32, 32, 32, 32, 32, ~
$ TEMP_MEDIA <dbl> 28.0, 28.0, 28.0, 28.0, 28.0, 28.0, 28.0, 28.0, 28.0, 28.0, ~
$ DESCARGAS_ATM <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
$ RAI0 <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
$ RAJADA <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
$ VENTOS_KMH <dbl> 18.72, 18.72, 18.72, 18.72, 18.72, 18.72, 18.72, 18.72, 18.~
```

Estimação do modelo Poisson para São Luís

```
#####
# ESTIMAÇÃO DO MODELO POISSON #
#####
#Estimação do modelo
modelo_poisson <- glm(formula = INTERRUPTAO ~ FASES + KVA
+ MEDIA_CLIENTES + PLUVIOMETRIA + TEMP_MIN
+ TEMP_MAX + TEMP_MEDIA + DESCARGAS_ATM
+ RAI0 + RAJADA + VENTOS_KMH
,
data = base_slz,
family = "poisson")
```

Probabilidade de interrupção de energia em condições hipotéticas em São Luís

```
> #Modelo Poisson:
> predict(object = modelo_poisson2,
+         newdata = data.frame(INTERRUPTAO = 1,
+                               FASES = "ABC",
+                               KVA = "112,5",
+                               PLUVIOMETRIA = 15,
+                               TEMP_MIN = 22,
+                               RAI0 = 1,
+                               RAJADA = 30),
+         type = "response")
1
0.3657294
```

