

ANALISE DA ASSOCIAÇÃO DAS CONDIÇÕES CLIMÁTICAS E QUANTIDADE DE OCORRÊNCIAS DE INTERRUPÇÃO DE ENERGIA ELÉTRICA

Emanuel Rocha Santos
emanuel.rsantos@enel.com
IFF

Luiz Eduardo Granado Cardoso
luizeduardo_granado@hotmail.com
IFF

Henrique Rego Monteiro da Hora
henrique.dahora@iff.edu.br
IFF

Resumo: O trabalho teve como objetivo estudar a associação entre a quantidade de ordens emergenciais no sistema de distribuição de energia e as variáveis climáticas de temperatura, umidade, vento e radiação na região de Farol de São Tomé. O clima tem forte influência no número de ocorrências de interrupção de energia, afetando diretamente a confiabilidade do sistema de distribuição de energia elétrica. Eventos como rajada de vento forte ou variações de temperatura podem aumentar a quantidade de interrupções de fornecimento e gerar impactos negativos na qualidade da energia fornecida aos consumidores. Neste trabalho foi analisada a associação entre as variáveis independentes: velocidade do vento, rajada de vento, radiação, temperatura, umidade e a variável dependente: quantidade de ordens. A técnica aplicada para modelagem dos dados foi a Regressão Linear. Para a etapa de determinação da técnica de modelagem, foi utilizado o software CurveExpert, o qual calcula os parâmetros estatísticos da distribuição dos dados, bem como sugere técnica de modelagem baseado em indicadores estatísticos de confiança. Foi utilizado também o software R, onde foi aplicado para verificação da linearidade e normalidade do modelo, bem como cálculo dos parâmetros da regressão linear. Através de aquisição e tratamento dos dados, bem como aplicação da técnica de regressão linear foi possível determinar, com grau de confiança de 36%, as variáveis e coeficientes associados, os quais explicam o volume de ordens emergenciais em função das variáveis climáticas. É importante destacar que um modelo preciso é uma

ferramenta útil para as concessionárias, permitindo que elas alterem seus planos de contingência, em função das variações das condições climáticas, melhorando sua resposta às mudanças climáticas e, conseqüentemente, à confiabilidade.

Palavras Chave: Interrupção Energia - Regressão Linear - Condições Climáticas - Modelagem - Dados

1. INTRODUÇÃO

Zelar pelo fornecimento contínuo de energia, é a atribuição principal das concessionárias de energia elétrica, proporcionando ao consumidor, qualidade no fornecimento (ANEEL, 2018a). Entretanto o número de ocorrências de interrupções de energia é altamente dinâmico em um sistema de distribuição de redes áreas, o que traz grandes desafios para as prestadoras de serviços em função da imprevisibilidade de incidências, que são fortemente influenciadas pelas condições climáticas, tais como: calor, fortes chuvas e vento intenso. Os fatores climáticos geram problemas desde defeitos em conexões até distúrbios causados por vegetação sobre a rede (AMORIM, 2010; FANT et al., 2020)

O setor elétrico tem sua eficiência medida e auditada pela Agência Nacional de Energia Elétrica (ANEEL), que é responsável por regular desde a geração até as redes distribuição de média tensão (MT) e baixa tensão (BT) (CARDOSO, 1996). Estabelece normas e padrões de operação e controle, para tal, edita limites relativos à Qualidade da Energia Elétrica (QEE), e compeli as concessionárias de energia a enviar os dados de forma oficial para a agência. E estabelece ainda os limites máximos admitidos para os indicadores, relativos à continuidade, valor de tensão, tempo de normalização, frequência das de interrupções, e definindo penalização sob forma de multa, para indicadores coletivos, e compensações aos consumidores, para indicadores individuais (ANEEL, 2018b). Observando estes fatos podemos perceber que perdas econômicas ocasionadas pela interrupção no fornecimento são muito significativas geram custo de dois gêneros, custo financeiros e sociais. Os financeiros englobam a ausência de faturamento, pertinente a energia não vendida e o ressarcimento aos consumidores pelo período sem fornecimento e os sociais relacionados a satisfação dos clientes em relação a imagem da concessionária (ALMEIDA et al., 2010).

Fanucchi et al., (2016) faz uma análise da previsão da taxa de falhas sob condições climáticas diferentes utilizando o método regressão binomial negativa e demonstra sua eficiente em conjunto de dados com superdispersão e *outliers*, para verificar a recorrência de interrupções em um sistema de elétrico de distribuição e destacando aspectos climáticos, como velocidade da rajada de vento, número de descargas atmosféricas e precipitação de chuva, têm um papel importante na índice de falhas no fornecimento de elétrica, os dados climáticos estes utilizados tem similaridade com base deste trabalho.

Já o modelo proposto por Jaech et al. (2018) demonstra bons resultados a partir de correlação entre características climáticas e várias causas com base em registros históricos, ele prever o tempo de duração da interrupção através de um processamento de linguagem neural eficaz nas causas de interrupções com fatores relacionados diretamente a vento e a raios.

Não foram encontradas na literatura, trabalhos recentes abordando a influência condições climáticas associadas a número de ocorrências de interrupção de energia elétrica, utilizando como referência redes de distribuição com topologia aéreas convencionais aplicados a uma região litorânea no período compreendido 2019 a 2020.

Além de demonstrar a aplicabilidade do modelo, o objetivo deste trabalho é analisar, apurar e identificar o comportamento no número de ocorrências de interrupções de energia elétrica frente a variações climáticas utilizando regressão linear. Sua importância se dá no fato de aplicar a regressão linear para extrair potenciais informações da grande quantidade de dados armazenados por uma concessionária brasileira, do norte fluminense mais especificamente em sua região litorânea, com o intuito de explicitar características comuns a estas interrupções, permitindo, por parte da distribuidora de energia elétrica a criação de políticas de mitigação a estas interrupções, focadas em períodos que possuam as características identificadas.

2. METODOLOGIA

2.1 CLASSIFICAÇÃO DA PESQUISA

A pesquisa pode ser classificada como quantitativa e descritiva, pois tem o intuito de representar a influência das condições climáticas adversas na operação do sistema de redes de distribuição aérea a partir do número de registros de ocorrências de falhas no fornecimento. Para a análise dos dados foram utilizados os softwares *CurveExpert* e *R*, e com base na modelagem dos dados chegou-se à técnica de regressão linear para explicar a associação entre os dados climáticos e a quantidade de ordens.

2.2 PROCEDIMENTOS TÉCNICOS

Foi adotado no trabalho a metodologia baseada nas macroetapas deliberadas no trabalho de Cardoso (2019). Para alcançar o objetivo desse trabalho e empregado as seguintes etapas apresentadas na Figura 1.

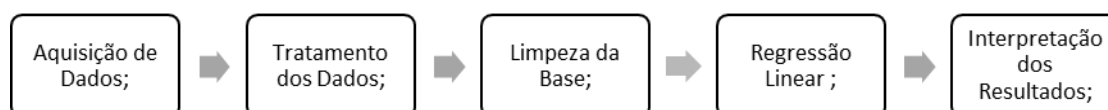


Figura 1: Fases da Pesquisa

Fonte: Cardoso (2019).

2.2 AQUISIÇÃO DE DADOS

Na primeira fase é realizada obtenção dos dados. Nesse trabalho os dados são provenientes de duas bases distintas: a base do sistema de atendimento ao consumidor (SAC) da distribuidora de energia e a base de dados do Instituto Nacional de Meteorologia (INMET). Na figura 2 podemos visualizar a localização da subestação da distribuidora e também coordenada da estação do INMET de Farol de São Tomé. Percebe-se que estas são próximas fisicamente.



Figura 2 : Localiza da Subestação e Estação Meteorológicos

Fonte: Elaboração Própria (2020).

Os dados das interrupções são compostos por 4.929 registros de ocorrências, extraídos da base da concessionária local energia, normatizados pelo Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional Módulo 8 (PRODIST 8), registra as informações necessárias relacionadas às interrupções do circuito, incluindo área de serviço, número de referência do circuito, causa da interrupção, duração, hora, data, número de clientes afetados, equipamentos disparados, ou do alimentador, de três subestações de uma concessionária brasileira do norte fluminense de 01 de Abril de 2019 a 31 de março de 2020. Toda essa malha do sistema e composta por 3 subestações, 8 alimentadores 250 quilômetros de linhas aéreas, fornecendo energia para 40.000 consumidores. A figura 3 apresenta o fluxo de limpeza dos dados retirados da base do SAC.

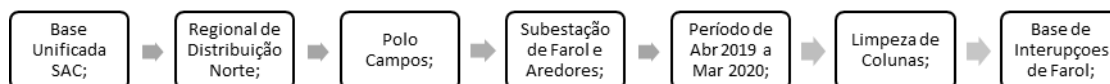


Figura 3 :Fluxo da Higienização dos Dados

Fonte: Elaboração Própria (2020).

Já os dados meteorológicos foram extraídos da base INMET, tomando como referência a estação A620 localizada na região de Farol de São Tomé, a qual disponibiliza os dados dos últimos 365 dias. A base é composta das seguintes variáveis: data, horário, ponto de orvalho, temperatura, umidade, Pressão, Velocidade do Vento, Radiação e Precipitação. Esses dados são fundamentais para análise da associação entre clima e interrupções não programadas. Como podemos observar figura 4 apresenta os dados disponibilizados pela estação A620 do INMET.

Figura 4 : Base de Dados Estação A620

Consulta Dados da Estação Automática: Campos - São Tomé (RJ)																	Fechar	
Observação: Estes são dados brutos e sem consistência com o único objetivo de deixá-los disponíveis de forma imediata.																		
Data Inicial: 20/04/2020		Data Final: 20/04/2020		Nova Pesquisa			Download de Dados											
Data	Hora	Temperatura (°C)			Umidade (%)			Pto. Orvalho (°C)			Pressão (hPa)			Vento (m/s)			Radiação (kJ/m²)	Chuva (mm)
		Inst.	Máx.	Mín.	Inst.	Máx.	Mín.	Inst.	Máx.	Mín.	Inst.	Máx.	Mín.	Vel.	Dir. (°)	Raj.		
20/04/2020	00	23.9	23.9	23.8	97	100	97	23.4	23.9	23.3	1019.9	1019.9	1019.4	2.1	190	4.6	-3.53	0.0
20/04/2020	01	23.9	24.0	23.8	95	100	95	23.1	23.8	23.1	1019.5	1020.0	1019.5	2.5	196	4.5	-3.49	0.0
20/04/2020	02	23.7	23.9	23.7							1019.3	1019.5	1019.2	2.8	213	5.1	-3.53	0.0
20/04/2020	03	22.8	23.9	22.8							1019.2	1019.4	1019.2	2.6	144	5.3	-3.18	0.0
20/04/2020	04	23.0	23.0	22.7							1018.8	1019.2	1018.7	2.0	156	6.2	-3.41	0.0
20/04/2020	05	23.2	23.2	23.0							1018.0	1018.8	1018.0	1.2	159	4.3	-3.00	0.0
20/04/2020	06	23.2	23.3	23.1							1017.7	1018.1	1017.7	1.5	221	3.7	-3.17	0.0
20/04/2020	07	21.4	23.2	21.4							1017.6	1017.9	1017.6	1.8	289	4.5	-2.12	0.0
20/04/2020	08	20.8	21.4	20.6							1017.6	1017.8	1017.4	2.2	261	4.3	-3.48	0.0
20/04/2020	09	20.9	21.2	20.8							1018.0	1018.1	1017.6	3.2	260	6.1	-2.16	0.0
20/04/2020	10	21.6	21.6	20.9							1018.6	1018.8	1018.0	2.5	256	5.8	115.1	0.0
20/04/2020	11	22.4	22.4	21.4							1019.1	1019.1	1018.6	4.7	252	8.2	744.9	0.0
20/04/2020	12	22.8	23.4	22.3							1019.6	1019.8	1019.1	3.6	254	8.2	1457.	0.0
20/04/2020	13	24.6	24.8	22.8							1019.7	1019.8	1019.6	3.3	238	6.9	1553.	0.0
20/04/2020	14	24.6	24.9	24.0							1019.0	1019.7	1019.0	4.0	227	7.3	1628.	0.0
20/04/2020	15	25.2	25.4	24.3	100	100	100	25.2	25.2	24.3	1018.0	1019.0	1018.0	4.2	226	7.8	2239.	0.0
20/04/2020	16	24.7	25.3	24.5	97	100	95	24.1	25.1	23.7	1017.4	1018.0	1017.4	4.3	227	8.0	1856.	0.0
20/04/2020	17	25.4	25.4	24.5	87	97	87	23.2	24.4	23.0	1016.9	1017.4	1016.9	4.1	225	7.8	2177.	0.0
20/04/2020	18	25.0	25.6	24.8	86	88	84	22.4	23.1	22.0	1016.6	1016.9	1016.6	4.3	222	7.8	2173.	0.0
20/04/2020	19	24.5	25.2	24.4	92	93	85	23.2	23.2	22.1	1016.6	1016.7	1016.6	4.4	213	7.6	1392.	0.0
20/04/2020	20	24.1	24.7	24.0							1017.0	1017.1	1016.6	3.9	210	7.6	548.3	0.0
20/04/2020	21	23.7	24.1	23.7							1016.9	1017.1	1016.9	3.0	208	7.3	28.50	0.0
20/04/2020	22	23.6	23.7	23.6							1017.3	1017.3	1016.9	3.4	206	6.8	-3.54	0.0
20/04/2020	23	23.8	23.8	23.6							1018.0	1018.0	1017.3	3.9	212	6.9	-3.54	0.0

Estação: Campos - São Tomé-A620
 Código OMM: 86890
 Registro: 23 UTC
 Temp. Max.: 21.6 °C
 Temp. Min.: 21.3 °C
 Umidade: ///%
 Pressão: 1009.6 hPa
 Precipitação: 0.0 mm
 Vento Dir: 12 °
 Vento Vel: 2.9 m/s



11014
000011
000101

Dados Gráfico

Aberta em: 13/06/2008
 Latitude: -22.041647°
 Longitude: -41.051871°
 Altitude: 7 metros

Fonte: INMET (2020).

Para o estudo foi extraído 11 atributos das bases dados, 6 da estação A620 INMET e 5 do SAC da concessionária de energia. Da base da distribuidora foi retirado a data, número da ordem, causa, subestação e alimentador. Entretanto na base do INMET foi extraído a data, temperatura instantânea, umidade instantânea, velocidade do vento, rajada de vento e radiação. Após obter os dados das duas bases, foi necessário uni-las em uma base única, e para isso, o atributo a data foi utilizado como referência. As etapas posteriores serem realizadas nessa nova base.

2.4 TRATAMENTO DE DADOS

Nessa fase foi realizado, tratamentos na base dados este tem o motivo facilitar a obtenção de modelos durante a realização da regressão linear, os tratamentos realizados foram:

- Agrupamento do número de registro das ordens de acordo com nomenclatura alimentador e da subestação;
- Agrupamento das ocorrências em cada subestação por hora;

- Agrupamento das ocorrências em cada subestação por data;
- Agrupamento das ocorrências em cada subestação por semana;
- Cálculo da média da temperatura instantânea, da umidade instantânea, da velocidade do vento, da rajada de vento, radiação todos foram calculados através 24 referencias hora disponibilizadas pela base INMET.

2.5 LIMPEZA DOS DADOS

Nesta terceira fase e realizado pré-processamento dos dados que inclui a higienização destes, pois alguns registros estavam incompletos e outras colunas possuem valores não estruturados ou ainda as anomalias onde o cliente teve uma interrupção de energia por um defeito interno na instalação não sendo atribuído a responsabilidade a concessionária. A limpeza realizada na base foi:

- Exclusão dos registros de registros anulados;
- Exclusão dos registros com dados faltante;
- Exclusão dos registros que são anomalias;
- Exclusão 62 colunas entre estes pontos elétricos, viatura, afetação e etc.;
- Exclusão dos registros de ponto de orvalho e pressão.

2.4 REGRESSÃO LINEAR

Nesta quarta fase iremos comentar a regressão modelo de regressão linear , está e uma metodologia estatística que permite estabelecer uma relação linear entre duas variáveis, usualmente representas por X e Y , ambas de natureza quantitativa, onde Y representa a variável dependente ou variável resposta, cujo valor se pretende explicar e X representa a variável independente ou variável explicativa, cujo valor é conhecido. Segue o exemplo de equação regressão linear (PETERNELLI, 2018; TANNI et al., 2020).

Neste trabalho será analisada a correlação entre as variáveis independentes: velocidade do vento, rajada de vento, precipitação, temperatura, umidade e a variável na influência na variável dependente quantidade de ordens.

Para as etapas de modelagem e cálculo da regressão linear, foi utilizado o *software CurveExpert*, no qual e focado na análise de regressão não-linear e ajuste de curva e suavização de dados e o R. e uma linguagem de programação plataforma aberta e consolidado no mercado para utilização de desenvolvimento e análises estatísticas.

3. RESULTADOS E DISCUSSÃO

Para definição da técnica de modelagem, base de dados e modelo, foram realizadas quatro etapas de modelagem e ajustes na base de dados. A seguir serão detalhadas as etapas.

Na primeira etapa base de dados foi consolidada com a variável a ser explicada: quantidade de ordens e variáveis independentes: Temperatura Instantânea (°C), Umidade Instantânea (%), Velocidade do vento (m/s), rajada de vento (m/s) e radiação (KJ/m²). Os dados foram agrupados por hora, dia, mês e ano, conforme tabela da figura 5. Abaixo o detalhamento das unidades de medida de cada variável.

Graphs and Data						
Data	Notes	Data Plot	Top Results	+		
	temp_inst	umid_inst	vento_vel	vento_rajada	radiacao	Quantidade de Ordens
1	24.200000	83.000000	0.500000	1.500000	-3.540000	1.000000
2	21.500000	59.000000	1.200000	4.500000	-3.530000	3.000000
3	22.300000	79.000000	5.200000	10.500000	-3.540000	1.000000
4	25.200000	86.000000	4.500000	10.200000	1027.000000	1.000000
5	25.400000	77.000000	1.700000	3.400000	916.800000	1.000000
6	23.600000	65.000000	7.500000	12.600000	1231.000000	0.000000
7	20.100000	72.000000	5.400000	11.600000	841.900000	4.000000
8	23.900000	70.000000	5.700000	9.700000	1598.000000	4.000000
9	20.900000	79.000000	3.700000	11.500000	10.900000	2.000000
10	24.200000	85.000000	9.000000	15.400000	-3.150000	1.000000
11	21.800000	86.000000	1.100000	2.600000	-3.540000	2.000000
12	24.400000	87.000000	4.100000	10.400000	-1.860000	0.000000
13	24.100000	85.000000	7.200000	13.300000	-3.540000	2.000000
14	21.900000	89.000000	2.800000	5.000000	-2.050000	3.000000

Figura 5: Modelagem dos dados agrupados por hora

Fonte: Elaboração Própria (2020).

Os dados foram aplicados no *CurveExpert* para buscar a melhor técnica de modelagem. Com o agrupamento das variáveis até o nível de detalhe de hora, obteve-se o resultado da figura 6.

Results			
Name	Kind	Score	R ²
Linear	Regressi...	453	0.0405...
Savitzky-Golay Smoothing [window=823, d...	Smoother	0.0	0.0000...
Lowess Smoothing	Smoother	0.0	0.0000...

Figura 6: Resultado da primeira etapa de modelagem

Fonte: Elaboração Própria (2020).

A melhor técnica apontada foi a regressão linear, porem com um baixo grau de confiança, que foi medido através do R² de 4,05%. Diante disso foi realizado ajuste na base de dados e realizada a segunda etapa.

Já na segunda etapa na segunda etapa, os dados foram agrupados por dia e não mais por hora como na etapa anterior. As variáveis foram mantidas, porem no agrupamento houve a transformação para média dos valores por dia, grau hierárquico do agrupamento. Na figura 7, é apresentada a nova base de dados após a transformação dos dados.

Graphs and Data						
Data	Notes	Data Plot	Top Results	+		
	Média de temp_inst	Média de umid_inst	Média de vento_vel	Média de vento_rajada	Média de radiacao	Quantidade de Ordens
1	24.770000	83.950000	1.750000	3.930000	606.130000	15.000000
2	23.080000	58.210000	2.490000	4.870000	1385.110000	15.000000
3	23.550000	76.380000	4.110000	7.880000	958.670000	7.000000
4	26.030000	83.710000	4.310000	8.340000	840.640000	11.000000
5	25.530000	73.330000	3.040000	6.150000	1059.190000	9.000000
6	23.280000	69.920000	6.450000	11.900000	946.050000	16.000000
7	20.490000	67.080000	6.110000	12.420000	802.730000	27.000000
8	21.210000	82.960000	3.230000	6.450000	734.610000	13.000000
9	21.540000	85.420000	4.200000	8.920000	283.740000	10.000000
10	24.060000	84.330000	5.400000	10.930000	277.350000	26.000000
11	23.680000	81.420000	2.310000	5.100000	546.650000	8.000000
12	25.690000	80.920000	4.110000	7.960000	546.180000	8.000000
13	23.300000	81.420000	3.370000	7.310000	668.360000	14.000000
14	24.110000	68.710000	6.190000	11.390000	1487.360000	13.000000

Figura 7: Modelagem dos dados agrupados por dia

Fonte: Elaboração Própria (2020).

Aplicando a função *CurveFinder* do programa *CurveExpert* foi apontado como melhor técnica a Regressão Linear com um grau de confiança ainda baixo, mostrado através do R² cujo valor foi de 16,6%. A Figura 8 apresenta o resultado da aplicação *CurveFinder*.

Figura 8: Resultado da modelagem na segunda etapa

Fonte: Elaboração Própria (2020).

Name	Kind	Score	R^2
Linear	Regression	453	0.166105
Savitzky-Golay Smoothing [window=37, degree=4]	Smoother	0.0	0.000000
Lowess Smoothing	Smoother	0.0	0.000000

Com o intuito de melhorar o grau de confiança, foi realizada uma nova transformação na base de dados, nesta terceira etapa os dados foram agrupados por semana, onde o número de ordens foi somado e as demais variáveis foram transformadas para média. Na tabela da figura 9 é apresentado o agrupamento por semana realizado.

DIA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
SEMANA	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	C	C	C	C	C	C	C	C	D	D	D	D	D	D	D	D

Figura 9: Agrupamento dos dias em semanas

Fonte: Elaboração Própria (2020).

Na terceira etapa, com o agrupamento dos dados por semana, foi possível analisar graficamente a relação entre as variáveis climáticas e a quantidade de ordens. Na figura 10 é

destacada a relação entre a quantidade de ordens e rajada de vento. Percebe-se que com o aumento da rajada de vento ocorre também o aumento no volume de ordens emergenciais.

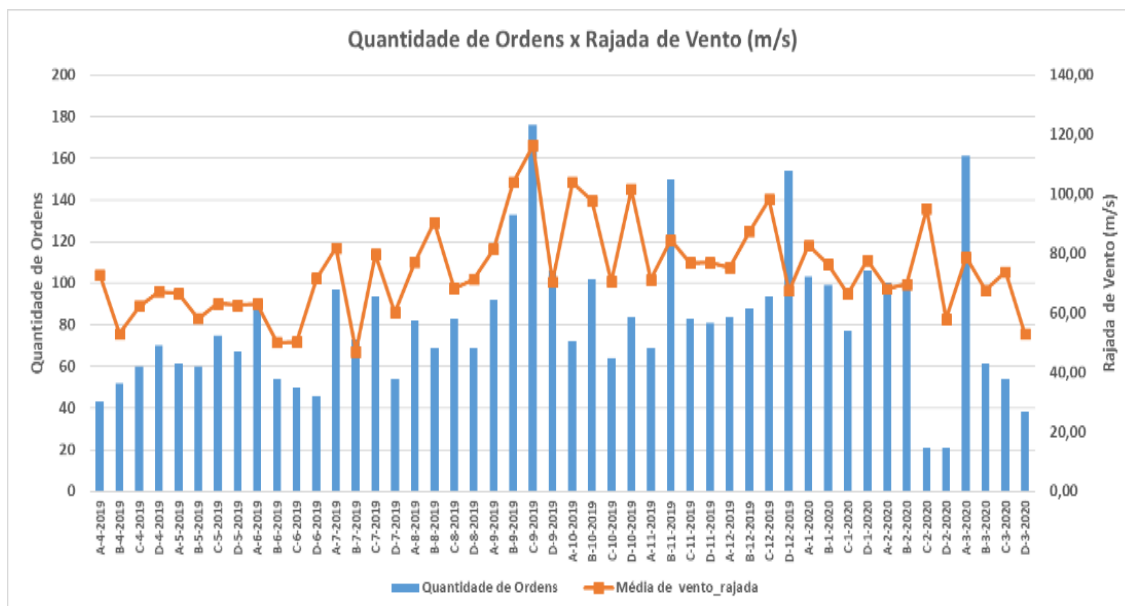


Figura 10: Quantidade de ordens versus rajada de vento

Fonte: Elaboração Própria (2020).

Os dados transformados foram inseridos novamente o *CurveExpert* para nova modelagem. A figura 11 de dados inserida no *Software*.

Graphs and Data						
Data	Notes	Data Plot	Top Results	+		
	Média de temp_inst	Média de umid_inst	Média de vento_vel	Média de vento_rajada	Média de radiacao	Quantidade de Ordens
1	24.590000	82.860000	3.330000	6.640000	907.860000	77.000000
2	23.130000	72.110000	3.980000	7.690000	976.200000	81.000000
3	22.690000	75.920000	3.980000	8.150000	762.180000	92.000000
4	26.070000	73.640000	3.340000	6.650000	857.300000	61.000000
5	23.490000	77.330000	5.540000	10.380000	895.980000	133.000000
6	20.900000	72.420000	3.640000	7.970000	793.700000	94.000000
7	20.900000	73.260000	2.270000	4.700000	749.190000	73.000000
8	24.460000	79.070000	4.220000	8.450000	746.190000	150.000000
9	23.750000	77.300000	2.480000	5.000000	648.860000	54.000000
10	25.620000	78.210000	4.600000	8.740000	885.920000	88.000000
11	24.430000	76.390000	4.960000	9.780000	939.850000	102.000000
12	21.410000	73.490000	3.430000	7.040000	704.090000	106.000000
13	23.990000	77.680000	3.800000	7.530000	912.290000	84.000000
14	25.870000	76.460000	3.410000	6.740000	1178.450000	154.000000

Figura 11: Modelagem de dados agrupados por semana

Fonte: Elaboração Própria (2020).

Aplicando a função *CurveFinder* do programa *CurveExpert* foi apontado como melhor técnica a regressão linear com grau de confiança de 36,1%, mostrado através do R^2 . A figura 12 apresenta o resultado do *CurveFinder*.

Name	Kind	Score	R ²
Linear	Regression	475	0.361051
Lowess Smoothing	Smoother	0.0	0.000000

Figura 12: Resultado da modelagem na terceira etapa

Fonte: Elaboração Própria (2020).

Analisando as possibilidades de transformação e também o grau de confiança apresentado na última modelagem, decidiu-se utilizar o modelo de regressão linear com o grau de confiança R² de 36%. Tal decisão foi considerada também do ponto de vista de aplicação, uma vez que atualmente a concessionária local não utiliza nenhum modelo para associação entre dados climáticos e quantidade de ordens. Analisando os gráficos e relação entre variáveis, percebeu-se que a velocidade e rajada do vento são dependentes entre si, de modo que selecionaremos somente uma para compor o modelo. Verificou-se, através da modelagem no *CurveExpert*, que a variável rajada de vento mante o melhor nível de confiança.

Entretanto na quarta etapa foi excluída a variável velocidade do vento e novamente foi feita a modelagem no *CurveExpert*. Na figura 13 as novas variáveis do modelo.

	Média de temp_inst	Média de umid_inst	Média de vento_rajada	Média de radiacao	Quantidade de Ordens
1	26.810000	73.400000	7.260000	986.920000	43.000000
2	25.730000	76.270000	5.300000	726.740000	52.000000
3	26.300000	75.940000	6.230000	814.530000	60.000000
4	26.510000	72.930000	6.720000	1017.530000	70.000000
5	26.070000	73.640000	6.650000	857.300000	61.000000
6	24.730000	74.750000	5.810000	574.860000	60.000000
7	23.990000	80.340000	6.320000	563.180000	75.000000
8	24.640000	76.350000	6.260000	684.400000	67.000000
9	22.520000	75.960000	6.300000	565.830000	92.000000
10	23.750000	77.300000	5.000000	648.860000	54.000000
11	23.390000	71.680000	5.030000	601.670000	50.000000
12	24.330000	76.260000	7.160000	663.790000	46.000000
13	22.500000	72.680000	8.190000	582.790000	97.000000
14	20.900000	73.260000	4.700000	749.190000	73.000000

Figura 13: Modelagem dos dados sem a variável velocidade do vento

Fonte: Elaboração Própria (2020).

Aplicando a função *CurveFinder* foi apontado como melhor técnica a regressão linear com grau de confiança de 36,02%, mostrado através do R². A imagem abaixo apresenta o resultado do *CurveFinder*.

Name	Kind	Score	R ²
Linear	Regression	476	0.360206
Lowess Smoothing	Smoother	0.0	0.000000

Figura 14: Resultado do modelo sem a variável velocidade do vento

Fonte: Elaboração Própria (2020).

Na figura 15 apresenta-se o resultado da modelagem com as variáveis e seus respectivos coeficientes.

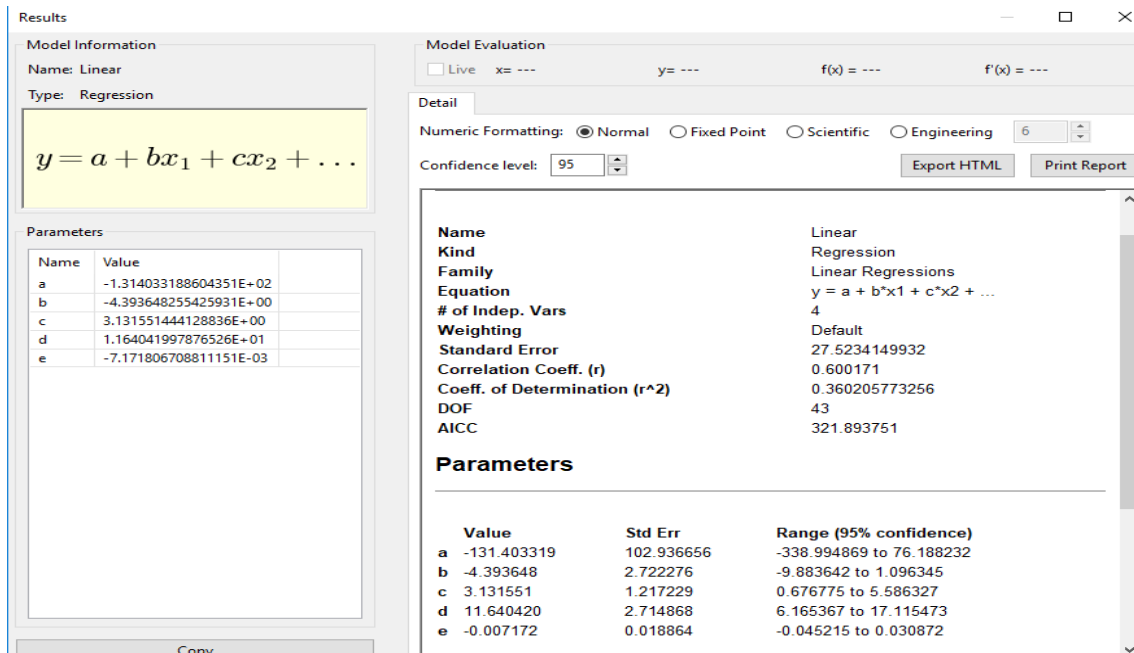


Figura 15: Resultado da modelagem

Fonte: Elaboração Própria (2020).

3.1 APLICAÇÃO DA TÉCNICA DE REGRESSÃO LINEAR

Os dados devem ser quantitativos para que a regressão linear tenha performance, caso contrário, para dados qualitativos, outros métodos devem ser utilizados (PETERNELLI, 2018). Tanto as variáveis independentes quanto as dependentes, alvo deste trabalho, são quantitativas. Como exemplo, no gráfico da figura abaixo encontram-se as variáveis de temperatura instantânea e quantidade de ordens.

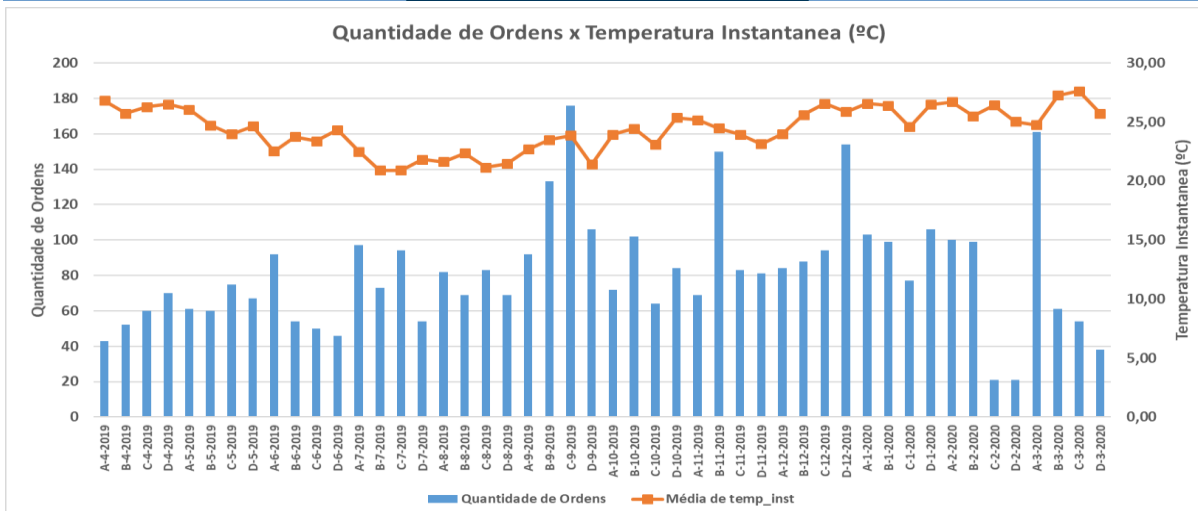


Figura 16: Quantidade de Ordens versus Temperatura

Fonte: Elaboração Própria (2020).

O Diagrama de Dispersão é simplesmente uma coleção de pontos num plano cujas duas coordenadas Cartesianas são os valores de cada membro do par de dados. A dispersão dos dados deve apresentar uma trajetória linear. Caso apresente uma outra trajetória, deve ser adequada a modelagem para o tipo de dispersão (exponencial por exemplo). Na figura 17 é apresentada a dispersão dos dados do modelo onde é verificada tendência linear de crescimento, respeitando assim uma condição para aplicação da regressão.

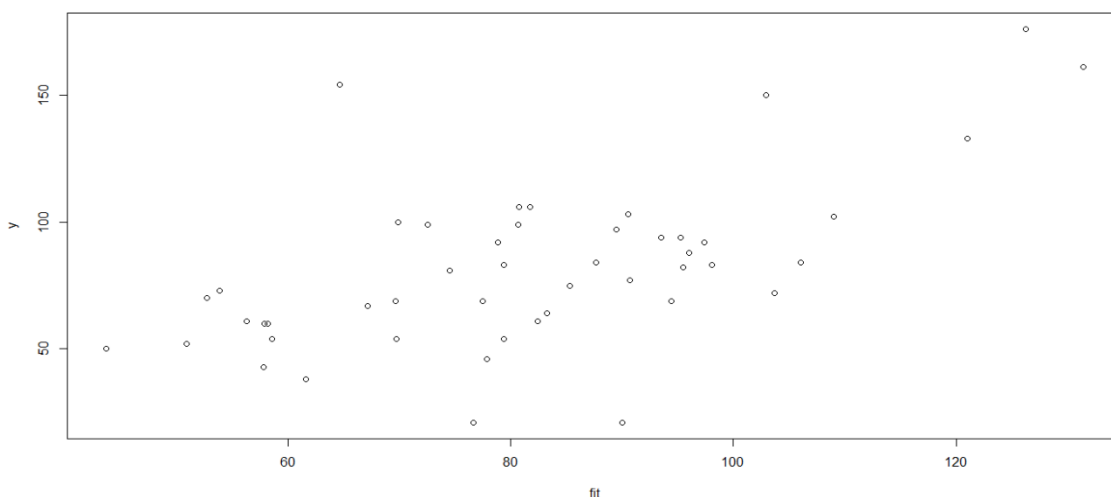


Figura 17: Gráfico de dispersão do Modelo versus dados reais

Fonte: Elaboração Própria (2020).

A distribuição dos Erros é a diferença entre a projeção e o valor real do ponto. Na aplicação da regressão deve-se observar se a distribuição dos erros segue um padrão normal. Conforme figura 18, é observada trajetória normal na distribuição.

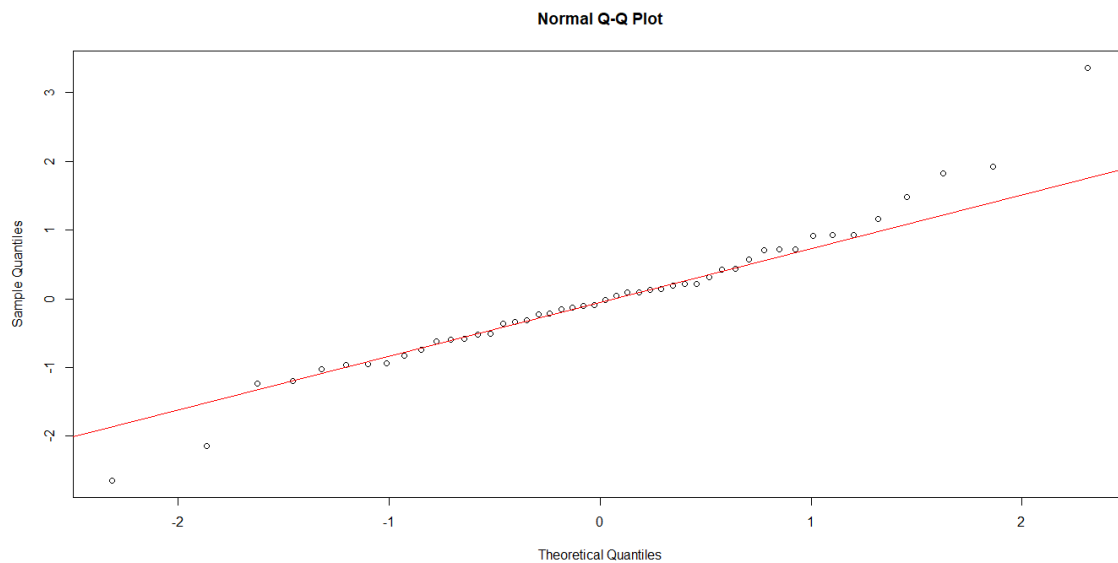


Figura 18: Gráfico de Normalidade

Fonte: Elaboração Própria (2020).

O Resultados da Regressão: Na Figura 19, o resultado da modelagem no *Software R*.

```
Call:
lm(formula = quantidade_de_ordens ~ Media_de_temp_inst + Media_de_umid_inst +
  Media_de_vento_rajada + Media_de_radiacao, data = dados)

Residuals:
    Min       1Q   Median       3Q      Max
-69.142 -14.168  -1.364  12.616  89.624

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.314e+02  1.029e+02  -1.277   0.2086
Media_de_temp_inst  -4.394e+00  2.722e+00  -1.614   0.1139
Media_de_umid_inst   3.132e+00  1.217e+00   2.573   0.0136 *
Media_de_vento_rajada  1.164e+01  2.715e+00   4.288   0.0001 ***
Media_de_radiacao   -7.172e-03  1.886e-02  -0.380   0.7057
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 27.52 on 43 degrees of freedom
Multiple R-squared:  0.3602,    Adjusted R-squared:  0.3007
F-statistic: 6.052 on 4 and 43 DF,  p-value: 0.000591
```

Figura 19: Resultado da Regressão

Fonte: Elaboração Própria (2020).

Com base no resultado foi possível encontrar os coeficientes de cada variável e chegar a modelo que explica a variável Quantidade de Ordens em função das variáveis: Temperatura Instantânea, Umidade Instantânea, Rajada de Vento e Radiação. O Resultado da modelagem encontra-se da equação abaixo.

$$Ordens_{qtd} = -131,4 - 4,39 \times Temp_{inst} + 3,13 \times Umid_{inst} + 11,6 \times Rajada - 0,007 \times Radiação$$

4. CONSIDERAÇÕES FINAIS

O trabalho teve como objetivo estudar a associação entre a quantidade de ordens emergenciais no sistema de distribuição de energia e as variáveis climáticas de temperatura, umidade, vento e radiação na região de Farol de São Tomé. Esse objetivo foi atingido através de aquisição e modelagem dos dados da empresa de distribuição de energia local e também da estação meteorológica do INMET, bem como aplicação da técnica de regressão linear. Com a aplicação da técnica foi possível determinar, com grau de confiança de 36%, as variáveis e coeficientes associados, os quais explicam o volume de ordens emergenciais. Como possibilidade de trabalhos futuros, podem ser utilizadas outras técnicas de modelagem de dados, tais como: Regressão Logística e Regressão Binomial Negativa. Além disso, também podem ser utilizadas outras técnicas para tratamento dos dados, como exclusão dos pontos fora da curva (*Outliers*).

É importante destacar que um modelo preciso é uma ferramenta útil para as concessionárias, permitindo que elas alterem seus planos de contingência, em função das variações das condições climáticas, melhorando sua resposta às mudanças climáticas e, conseqüentemente, à confiabilidade.

5. REFERÊNCIAS

- ALMEIDA, R. S. et al. Qualidade do serviço, satisfação do consumidor e desempenho financeiro das empresas distribuidoras de energia elétrica brasileiras. **Anais do Congresso Brasileiro de Custos - ABC**, 2010.
- AMORIM, M. L. F. DE. **Otimização de atendimentos de emergência em redes de distribuição de energia elétrica**. Niterói: Universidade Federal Fluminense, 2010.
- ANEEL. **PRODIST – Módulo 8 – Qualidade da Energia Elétrica - Agência Nacional de Energia Elétrica**, 1 jan. 2018a. Disponível em: <<https://www.aneel.gov.br/modulo-8>>. Acesso em: 10 abr. 2020
- ANEEL. **PRODIST - Módulo 6 – Informações Requeridas e Obrigações - ANEEL**, 1 jan. 2018b. Disponível em: <<https://www.aneel.gov.br/modulo-6>>. Acesso em: 23 maio. 2020
- CARDOSO, F. H. LEI Nº 9.427. Institui a Agência Nacional de Energia Elétrica - ANEEL, disciplina o regime das concessões e atribuições. . 26 dez. 1996.
- CARDOSO, L. E. G. Análise das causas de atendimentos emergenciais em redes de distribuição de uma companhia de energia elétrica do norte e noroeste fluminense. p. 94, 2019.
- FANT, C. et al. Climate change impacts and costs to U.S. electricity transmission and distribution infrastructure. **Energy (Oxford, England)**, v. 195, 15 mar. 2020.
- FANUCCHI, R. Z. et al. **Failure rate prediction under adverse weather conditions in an electric Distribution System using Negative Binomial Regression**. 2016 17th International

Conference on Harmonics and Quality of Power (ICHQP). **Anais...** In: 2016 17TH INTERNATIONAL CONFERENCE ON HARMONICS AND QUALITY OF POWER (ICHQP). out. 2016

INMET. **INMET - Instituto Nacional de Meteorologia**. Disponível em:

<<http://www.inmet.gov.br/portal/>>. Acesso em: 20 abr. 2020.

JAECH, A. et al. Real-Time Prediction of the Duration of Distribution System Outages. **arXiv:1804.01189** [cs, math, stat], 29 jul. 2018.

PETERNELLI, L. A. **Regressão linear e correlação**, 2018. . Acesso em: 26 abr. 2020

TANNI, S. E. et al. Correlação vs. regressão em estudos de associação. **Jornal Brasileiro de E.**, v. 46, n. 1, 2020.