



Universidade Estadual de Londrina  
Centro de Tecnologia e Urbanismo  
Departamento de Engenharia Elétrica  
Programa de Pós-Graduação em Engenharia Elétrica

Leonardo Yuji Ishizaki

Proposta de planejamento direcionado à melhoria dos indicadores de desempenho em sistemas elétricos de distribuição via ferramentas de regressão

Londrina,  
29 de março de 2022



Leonardo Yuji Ishizaki

Proposta de planejamento direcionado à melhoria dos indicadores de desempenho em sistemas elétricos de distribuição via ferramentas de regressão

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica na Universidade Estadual de Londrina como parte dos requisitos necessários para a obtenção do Título de Mestre em Engenharia Elétrica.

Área: Qualidade de energia

Orientadora: Dra. Maria Bernadete de Moraes França

Coorientadora: Dra. Juliani Chico Piai

Londrina,  
29 de março de 2022

Ishizaki, Leonardo Yuji

Proposta de planejamento direcionado à melhoria dos indicadores de desempenho em sistemas elétricos de distribuição via ferramentas de regressão. Londrina, 2022. 95 p.

Orientadora: Dra. Maria Bernadete de Moraes França

Coorientadora: Dra. Juliani Chico Piai

Dissertação (Mestrado) – Departamento de Engenharia Elétrica –  
Universidade Estadual de Londrina

1. Regressão linear múltipla, 2. Qualidade de energia, 3. Indicadores da distribuição

Leonardo Yuji Ishizaki

Proposta de planejamento direcionado à melhoria dos indicadores de desempenho em sistemas elétricos de distribuição via ferramentas de regressão

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Estadual de Londrina como parte dos requisitos necessários para a obtenção do título de Mestre em Engenharia Elétrica.

Área: Qualidade de energia

Banca Examinadora

---

Dra. Maria Bernadete de Moraes França  
Departamento de Engenharia Elétrica  
Universidade Estadual de Londrina  
Orientadora

---

Dr. Luis Alfonso Gallego Pareja  
Departamento de Engenharia Elétrica  
Universidade Estadual de Londrina  
Membro

---

Dr. Rodrigo Zempulski Fanucchi  
Departamento de Engenharia Elétrica  
Universidade Estadual de Londrina  
Membro externo

Londrina,  
29 de março de 2022



*Dedico este trabalho ao meu pai, Edson e minha mãe, Mery.*





# Agradecimentos

Primeiramente, agradeço meus pais por todo carinho e incentivo, que foram meus pilares no desenvolvimento deste trabalho. Agradeço a minha orientadora por todo apoio e pela oportunidade de realizar um estudo de mestrado. Agradeço também minha coorientadora por todas as reuniões e ensinamentos. Por fim, agradeço meus amigos, minha tia Paula e meu irmão Maurício que, mesmo distantes, me acompanharam nesta caminhada.



# Resumo

Nos dias atuais, de acordo com as normas e metodologias impostas nos Procedimentos de Distribuição (PRODIST), estabelecido pela Agência Nacional de Energia Elétrica (ANEEL), as empresas concessionárias do setor de distribuição estão sujeitas a sanções financeiras e até perdas de concessão, em casos de transgressão dos limites estabelecidos pela própria ANEEL. Vinculada à esse fato, as concessionárias de energia buscam constantemente por estratégias eficazes para garantir a qualidade de energia adequada, de forma a garantir o atendimento dentro dos limites estabelecidos, com custos abaixo do disponibilizado na tarifa e consumidores satisfeitos com os serviços prestados. Dado o cenário atual, em que determinadas empresas sofrem com altos valores de compensações financeiras, o trabalho em questão teve por objetivo a implementação de ferramentas de regressão linear múltipla (RLM) que auxiliem as concessionárias na identificação das causas que geram a insatisfação dos consumidores residenciais. A metodologia proposta consistiu na construção de modelos de RLM, baseado na estrutura de dados em painel, para avaliar o impacto, no Índice ANEEL de satisfação do consumidor (IASC), de diferentes indicadores da distribuição, além de outras variáveis do setor como consumo de energia elétrica, número de consumidores e o valor das compensações financeiras pagas de uma determinada concessionária. Os resultados com os modelos propostos confirmaram algumas discussões já apresentadas na literatura especializada, além de apresentar novas contribuições para a metodologia do cálculo do IASC.

**Palavras-chave:** Regressão linear múltipla, Qualidade de energia, Indicadores da distribuição



# Abstract

Nowadays, according to the rules and methodologies imposed in the Procedures for Distribution (PRODIST), established by the National Agency of Electric Energy (ANEEL), the concessionaires in the distribution sector are subject to financial sanctions and even loss of concession, in cases of transgression of the established limits. Linked to this fact, the energy concessionaires are constantly searching for effective strategies to ensure the quality of adequate energy, in order to guarantee service within the established limits, with costs below the available tariff and consumers satisfied with the services provided. Given the current scenario, in which certain companies suffer from high values of financial compensations, the objective of this study was to implement multiple linear regression (MLR) tools to assist utilities in identifying the causes of residential customer dissatisfaction. The proposed methodology consisted in the construction of RLM models, based on the panel data structure, to assess the impact, on the ANEEL Consumer Satisfaction Index (IASC), of differences current indicators of the distribution, in addition to other variables in the sector such as consumption of electricity, number of consumers and the amount of financial compensation paid for a particular dealership. The results with the proposed models confirmed some discussions have already been presented in the specialized literature, in addition to presenting new contributions to the IASC calculation methodology.

**Keywords:** Multiple linear regression, Power quality, Distribution indicators



# Lista de abreviaturas e siglas

<b>ABRADEE</b>	Associação Brasileira de Distribuidoras de Energia Elétrica
<b>ABNT</b>	Associação Brasileira de Normas Técnicas
<b>ANEEL</b>	Agência Nacional de Energia Elétrica
<b>BLUE</b>	<i>Best linear unbiased estimators</i>
<b>BT</b>	Baixa tensão
<b>CRC</b>	Conta de Resultados a Compensar
<b>DNAEE</b>	Departamento Nacional de Águas e Energia Elétrica
<b>ELETOBRAS</b>	Centrais Elétricas Brasileiras S. A.
<b>IEEE</b>	<i>Institute of Electrical and Electronics Engineers</i>
<b>FRP</b>	Função de regressão populacional
<b>MAE</b>	Mercado Atacadista de Energia
<b>MT</b>	Média tensão
<b>MQG</b>	Mínimos quadrados generalizados
<b>MQO</b>	Mínimos quadrados ordinários
<b>ODS</b>	Objetivos de Desenvolvimento Sustentável
<b>PIB</b>	Produto Interno Bruto
<b>PRODIST</b>	Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional
<b>RESEB</b>	Reestruturação do Setor Elétrico Brasileiro
<b>RLM</b>	Regressão linear múltipla
<b>SEB</b>	Setor elétrico brasileiro





# List of Notations

- $\hat{x}$  : Valor estimado da variável  $x$ ;
- $\bar{x}$  : Valor médio da variável  $x$ ;
- $E(x|y)$  : Esperança da variável  $x$ , condicionada à variável  $y$ ;
- $\Sigma$  : Somatório;
- $\ln(x)$  : Logaritmo neperiano de  $x$ ;
- $Corr(x, y)$  : Correlação entre as variáveis  $x$  e  $y$ ;



# List of Symbols

- ACSI* American Customer Satisfaction Index;
- AENS* Average Energy not Supplied;
- ASAI* Average Service Availability Index;
- CAIDI* Customer Average Interruption Duration Index;
- CAIFI* Customer Average Interruption Frequency Index;
- DEC* Duração Equivalente de Interrupção por Unidade Consumidora;
- DER* Duração Equivalente de Reclamação;
- DIC* Duração de Interrupção Individual por Unidade Consumidora ou por Ponto de Conexão;
- DICRI* Duração de Interrupção Individual ocorrida em Dia Crítico por Unidade Consumidora ou por Ponto de Conexão;
- DMIC* Duração Máxima de Interrupção Contínua por Unidade Consumidora ou por Ponto de Conexão;
- DRC* Duração Relativa da Transgressão para Tensão Crítica;
- DRCE* Duração Relativa da Transgressão para Tensão Crítica Equivalente;
- DRP* Duração Relativa da Transgressão para Tensão Precária;
- DRPCE* Duração Relativa da Transgressão para Tensão Precária e Crítica Equivalente;
- DRPE* Duração Relativa da Transgressão para Tensão Precária Equivalente;
- ECSI* European Customer Satisfaction Index;
- ENS* Energy not Supplied;
- FIC* Frequência de Interrupção Individual por Unidade Consumidora ou por Ponto de Conexão;
- FEC* Frequência Equivalente de Interrupção por Unidade Consumidora;
- FER* Frequência Equivalente de Reclamação a cada mil Unidades Consumidoras;
- $H_1$  Hipótese Alternativa;
- $H_0$  Hipótese Nula;
- IASC* Índice ANEEL de Satisfação do Consumidor;
- $N_L$  Número total de unidades consumidoras avaliadas em uma determinada concessionária;
- $R^2$  Coeficiente de determinação;

- SAIDI* *System Average Interruption Duration Index;*
- SAIFI* *System Average Interruption Frequency Index;*
- SCSI* *Swedish Customer Satisfaction Index;*
- SQE* Soma Quadrática da Regressão;
- SQR* Soma Quadrática dos Resíduos;
- SQT* Soma Quadrática Total;
- TMA* Tempo Médio de Atendimento às ocorrências emergenciais;
- TMD* Tempo Médio de Deslocamento às ocorrências emergenciais;
- TME* Tempo Médio de Execução às ocorrências emergenciais;
- TMP* Tempo Médio de Preparação às ocorrências emergenciais;
- COMP<sub>DIC</sub>* Valor das compensações financeiras correspondente a violação dos limites do indicador DIC;
- COMP<sub>DICRI</sub>* Valor das compensações financeiras correspondente a violação dos limites do indicador DICRI;
- COMP<sub>DMIC</sub>* Valor das compensações financeiras correspondente a violação dos limites do indicador DMIC;
- COMP<sub>FIC</sub>* Valor das compensações financeiras correspondente a violação dos limites do indicador FIC;
- COMP<sub>S</sub>* Valor das compensações financeiras correspondente a violação dos limites dos indicadores voltados à qualidade do serviço;
- COMP<sub>SC</sub>* Valor das compensações financeiras correspondente a violação dos limites dos indicadores voltados à qualidade do serviço ocorridas em dia crítico;
- COMP<sub>P</sub>* Valor das compensações financeiras correspondente a violação dos limites dos indicadores voltados à qualidade do produto;
- VIF* *Variance Inflation Factor;*

# Lista de ilustrações

Figura 1.1 – Desverticalização das atividades do setor elétrico. . . . .	28
Figura 2.1 – Estratificações das interrupções de longa duração. . . . .	41
Figura 2.2 – Faixas de tensão em relação à de tensão de referência $T_R$ . . . . .	43
Figura 2.3 – Modelo para o cálculo do IASC. . . . .	48
Figura 5.1 – Gráfico de probabilidade normal para o termo de erro composto referente ao modelo da equação 5.1. . . . .	72
Figura 5.2 – Histograma do termo de erro composto referente ao modelo da equação 5.1. . . . .	73
Figura 5.3 – IASC x DEC e IASC x FEC para o banco de dados utilizado. . . . .	80



# Lista de tabelas

Tabela 1.1 – Compensações financeiras pagas no período de 2010 até 2019 (Dados disponibilizados pela ANEEL). . . . .	30
Tabela 2.1 – Lista dos fatores geradores. . . . .	40
Tabela 2.2 – Resumo dos indicadores da distribuição aplicados no modelo. . . . .	46
Tabela 5.1 – Parâmetros estimados e testes de significância para o modelo 4.16. . . .	70
Tabela 5.2 – Parâmetros estimados e testes de significância para o modelo 4.16 sem DRPCE. . . . .	71
Tabela 5.3 – Resumo dos testes aplicados no modelo representado pela equação 5.1. . . .	71
Tabela 5.4 – Comparação erros padrões e erros padrões robustos para o modelo da equação 5.1. . . . .	72
Tabela 5.5 – Matriz de coeficientes de correlação para o modelo da equação 4.16. . . .	73
Tabela 5.6 – Resumo dos testes aplicados no modelo representado pela equação 4.17. . . .	75
Tabela 5.7 – Matriz de coeficientes de correlação para o modelo da equação 4.17. . . .	75
Tabela 5.8 – Parâmetros estimados e testes de significância para o modelo 4.17. . . .	75
Tabela 5.9 – Resumo dos testes aplicados no modelo representado pela equação 4.18. . . .	77
Tabela 5.10–Matriz de coeficientes de correlação para o modelo da equação 4.18. . . .	78
Tabela 5.11–Parâmetros estimados e testes de significância para o modelo 4.18. . . .	78
Tabela 5.12–Resumo dos testes aplicados no modelo representado pela equação 4.19. . . .	79
Tabela 5.13–Matriz de coeficientes de correlação para o modelo da equação 4.19. . . .	79
Tabela 5.14–Parâmetros estimados e testes de significância para o modelo 4.19. . . .	80





# Sumário

1	INTRODUÇÃO . . . . .	27
1.1	Histórico do cenário regulatório e reestruturação do setor elétrico brasileiro . . . . .	27
1.2	Objetivos do trabalho . . . . .	30
1.3	Organização do trabalho . . . . .	31
2	QUALIDADE DA ENERGIA ELÉTRICA . . . . .	33
2.1	Indicadores da distribuição . . . . .	35
2.1.1	Confiabilidade no fornecimento de energia elétrica . . . . .	37
2.1.2	Conformidade do nível de tensão . . . . .	43
2.1.3	Qualidade do tratamento de reclamações . . . . .	44
2.1.4	Resumo dos indicadores da distribuição . . . . .	45
2.2	Compensações financeiras pagas aos consumidores . . . . .	46
2.3	Índice ANEEL de satisfação do consumidor . . . . .	47
3	TEORIA ECONOMETRICA . . . . .	51
3.1	Estrutura de dados . . . . .	51
3.2	Análise de regressão linear múltipla . . . . .	52
3.3	Hipóteses clássicas do modelo de Regressão Linear Múltipla . . . . .	54
3.4	Modelos de Regressão Linear Múltipla com dados em painel . . . . .	56
4	METODOLOGIA . . . . .	61
4.1	Significância dos parâmetros de regressão e do modelo . . . . .	61
4.2	Critérios para seleção de modelos . . . . .	64
4.3	Modelos de Regressão Linear Múltipla propostos . . . . .	65
5	RESULTADOS E DISCUSSÕES . . . . .	69
5.1	Modelo 1 - DRPCE . . . . .	69
5.2	Modelo 2 - Compensações financeiras . . . . .	74
5.3	Modelo 3 - Consumo . . . . .	76
5.4	Considerações gerais . . . . .	80
6	CONCLUSÕES . . . . .	83
	REFERÊNCIAS . . . . .	85

	<b>APÊNDICES</b>	<b>91</b>
	<b>APÊNDICE A – CÓDIGO STATA</b> . . . . .	<b>93</b>
<b>A.1</b>	<b>Preparação do modelo</b> . . . . .	<b>93</b>
<b>A.2</b>	<b>Modelo 1 - DRPCE</b> . . . . .	<b>93</b>
<b>A.3</b>	<b>Modelo 2 - Compensações financeiras</b> . . . . .	<b>94</b>
<b>A.4</b>	<b>Modelo 3 - Consumo</b> . . . . .	<b>95</b>

# 1 Introdução

A preocupação dos homens em relação ao acesso contínuo de insumos energéticos existe desde os anos pré-históricos com o descobrimento do fogo, que auxiliou com a preparação de alimentos e proteção contra o frio. Com o desenvolvimento da sociedade e suas tecnologias, o uso contínuo de energia elétrica passou a ser tão essencial quanto ao fogo nas ações humanas.

Atualmente, a energia elétrica é considerada um insumo essencial tanto para o conforto pessoal, quanto para o desenvolvimento econômico de um país. Segundo Vaccaro et al., o crescimento e desenvolvimento de um país são reflexos da disponibilidade de energia elétrica, uma vez que seu uso viabiliza melhorias nos âmbitos econômicos, sociais, ambientais e institucionais (VACCARO et al., 2011).

A energia elétrica é um produto de consumo que deve ter a confiabilidade e satisfação para o cliente assim como qualquer outro produto do mercado. Locais que sofrem com ausência do fornecimento de energia elétrica são desviados do progresso que pode ser proporcionado pela mesma (BARBOSA, 2018).

Em 2014, o Instituto Acende Brasil, analisando a qualidade do fornecimento de energia elétrica, apresentou uma importante consideração à respeito do PIB no Brasil. Ele apontou que apenas 2,2% do PIB do país são correspondentes ao setor elétrico. Entretanto, a falta de energia tem um impacto muito maior sobre essa porcentagem, uma vez que a eletricidade é um insumo básico para a produção de bens de consumo, sendo utilizada em praticamente todos os setores da economia (BRASIL, 2014).

Diante de tais fatos, destaca-se a importância da continuidade no fornecimento de energia elétrica, fato que recebe atenção da Organização das Nações Unidas (ONU), que de acordo com o Objetivo 7 dos Objetivos de Desenvolvimento Sustentável (ODS), referente à Energia Limpa e Acessível, estabelece como meta garantir acesso à energia barata, confiável, sustentável e renovável para todos (ONU, 2000).

Fora a ONU, tal fato também atrai atenção da Agência Nacional de Energia Elétrica (ANEEL), que no cenário atual, é o agente responsável por estabelecer às concessionárias limites baseados nos indicadores voltados à qualidade de energia, além de sanções, que representam compensações financeiras pagas aos clientes dada à violação de tais limites.

## 1.1 Histórico do cenário regulatório e reestruturação do setor elétrico brasileiro

No Brasil, o descobrimento e o início da utilização da energia elétrica ocorreu ainda no império brasileiro, em atividades restritas ao serviço público e industrial, muito

distantes do sistema interligado que existe atualmente (BARBOSA, 2018).

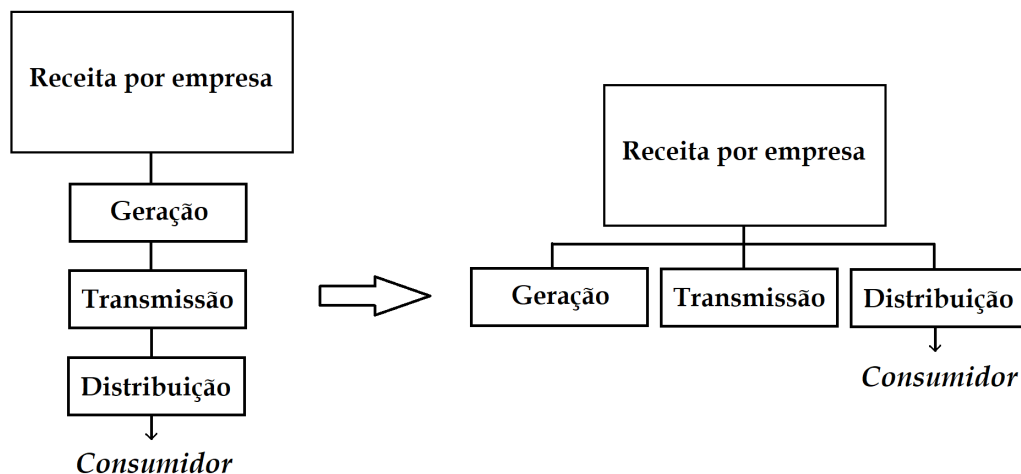
Até a década de 70, as empresas concessionárias empregavam apenas os dados históricos das falhas com o intuito de melhorar a qualidade de energia elétrica. Instrumentos legais para o estabelecimento de metas voltadas à qualidade de energia até então não existiam (FERREIRA, 2019).

Nesse período, se tratando da regulação da qualidade de energia, as portarias DNAEE nº 46 e nº 47 de 1978 foram pioneiras, introduzindo os primeiros limites voltados à qualidade do serviço e do produto (DNAEE, 1978a) (DNAEE, 1978b).

O processo de reestruturação do setor elétrico brasileiro teve seu início em 1993, com a publicação da lei de nº 8.631, que extinguiu a CRC (Conta de Resultados a Compensar) e desqualizou as tarifas em território nacional.

Após a eleição de Fernando Henrique Cardoso, o Ministério de Minas e Energia realizou o projeto RESEB (Reestruturação do Setor Elétrico Brasileiro), que propôs em seus relatórios uma série de ações com o intuito de fortalecer os mecanismos de mercado no setor. Dentre tais ações pode-se citar a introdução da competição, a desverticalização das atividades do setor elétrico, representada na figura 1.1, a privatização das estatais energéticas, além da criação de um órgão regulador, a ANEEL, um operador independente, o Operador Nacional do Sistema Elétrico (ONS), e um mercado livre, o Mercado Atacadista de Energia Elétrica (MAE) (PELEGRINI, 2003).

Figura 1.1 – Desverticalização das atividades do setor elétrico.



Fonte: Autor.

Sob regime de monopólio natural, a ausência de uma forma de regulação dos serviços poderia acarretar em condições desfavoráveis aos consumidores. A necessidade da atuação de órgãos reguladores nas definições de metas e sanções às concessionárias foi intensificada com a privatização de boa parte das concessionárias.

Em termos práticos, dada a ausência de penalidades em casos de transgressão das metas de continuidade, a melhoria na continuidade de fornecimento de energia elétrica com as portarias do DNAEE foi insuficiente. Além disso, devido as privatizações, intensificou-se também a necessidade da atualização dos contratos de concessão, visto que as metas de continuidade até então definidas não refletiam o novo cenário do setor elétrico de distribuição (PESSANHA et al., 2007).

A criação da ANEEL, uma autarquia em regime especial, ocorreu em 1996 através da lei de nº 9.427, vinculada ao Ministério de Minas e Energia. Em Janeiro de 2000, o Brasil, seguindo as tendências do setor elétrico mundial, definiu através da resolução de nº 24 que a ANEEL passa a ser responsável por essa regulação, definindo as metas de continuidade. Essa resolução, se aproveitando da experiência com as portarias do DNAEE, introduziu sanções em casos de não cumprimento das novas metas baseados nos limites estabelecidos (ANEEL, 2000).

Neste mesmo ano a ANEEL, com o intuito de se avaliar e monitorar os níveis de satisfação em relação às concessionárias, sob a perspectiva dos consumidores residenciais, propôs o índice ANEEL de satisfação do consumidor (IASC), que são pesquisas realizadas anualmente com consumidores de todas as distribuidoras, concessionárias e permissionárias brasileiras.

As distribuidoras mais bem avaliadas pelo IASC são premiadas com certificados e um selo, que pode ser apresentado nas faturas de energia elétrica. Os resultados das pesquisas IASC realizadas são utilizados no aprimoramento dos instrumentos regulatórios e na priorização das ações de fiscalização.

Em 2008, através de sua resolução normativa nº 345, a ANEEL elaborou os Procedimentos de Distribuição (PRODIST), dividido em onze módulos, que foram responsáveis pela normatização e padronização das atividades técnicas dos sistemas de distribuição, além de auxiliar na fiscalização das metas de continuidade (ANEEL, 2008b). Segundo da Silva et al., a inserção do PRODIST estimulou uma redução dos indicadores de continuidade (SILVA et al., 2014).

De acordo com a ABRADÉE, o setor de distribuição brasileiro é composto por empresas concessionárias que são controladas por entes públicos (municipal, estadual e federal) e privados. Em muitos casos, tais concessionárias não conseguem se adequar aos limites estabelecidos pela ANEEL, como consequência desta incapacidade, determinadas concessionárias têm sofrido com altos valores de compensações financeiras. A tabela 1.1 apresenta os valores das compensações financeiras pagas aos consumidores nos últimos 10 anos, das cinco concessionárias que mais foram penalizadas.

Tabela 1.1 – Compensações financeiras pagas no período de 2010 até 2019 (Dados disponibilizados pela ANEEL).

Concessionária	Valor das compensações financeiras pagas (R\$)
CELG	668.833.393,50
ELETROPAULO	504.539.154,00
CEMIG	403.295.208,80
COELBA	388.271.447,40
LIGHT	364.434.897,70

Atualmente, o cenário energético brasileiro é um sistema fortemente regulado, com um mercado justo e competitivo, responsável por coagir as empresas, principalmente as concessionárias de distribuição de energia elétrica, a promover melhorias em sua estrutura de negócios e de tomada de decisões à respeito dos investimentos. Em virtude disso, a busca por maiores níveis de confiabilidade aos sistemas é um dos principais objetivos das empresas distribuidoras (RODRIGUEZ, 2017).

Além disso, com a modernização dos equipamentos elétricos, alguns podem apresentar maiores sensibilidades, aumentando o grau de importância de discussões em relação à qualidade do fornecimento de energia (BARBOSA, 2018).

Portanto, a melhoria na qualidade do fornecimento de energia elétrica é um objetivo tanto da ANEEL e da ONU, quanto das concessionárias e dos consumidores, do setor industrial até o residencial, em virtude de todos serem diretamente afetados.

## 1.2 Objetivos do trabalho

De modo geral, o objetivo deste trabalho é modelar matematicamente, via regressão linear múltipla, o impacto na satisfação do consumidor residencial, avaliada pelo IASC, dos indicadores da distribuição, além de outras variáveis do setor, como o número de consumidores, consumo de energia elétrica e valor das compensações financeiras pagas pelas concessionárias, buscando auxiliar na melhoria da qualidade de energia elétrica, atendendo não somente as diretrizes apresentadas pela ANEEL, mas também a Meta 7 dos ODS.

De modo específico, os objetivos são:

- Compreender a metodologia aplicada no cálculo dos indicadores da distribuição, das compensações financeiras e do IASC;
- Avaliar a relação teórica entre as variáveis utilizadas no modelo, baseados nos PRODIST e nas resoluções normativas apresentadas pela ANEEL;

- Estimar os parâmetros dos modelos de RLM, com o auxílio do software STATA, e aplicar testes de hipóteses para avaliar os níveis de significância dos parâmetros individualmente e de forma conjunta;
- Verificar as hipóteses clássicas da RLM baseado no teorema de Gauss-Markov e as hipóteses para os modelos de dados em painel;
- Discutir o impacto dos diferentes indicadores do setor da distribuição na satisfação do consumidor residencial, baseado nos parâmetros estimados com os modelos, com o intuito de auxiliar as concessionárias na tomada de decisões;
- Comparar os resultados obtidos com os de trabalhos na literatura que avaliaram a melhoria da qualidade de energia sob outros indicadores ou metodologias.

### 1.3 Organização do trabalho

O capítulo 2 define os conceitos de qualidade de energia através de suas três dimensões: confiabilidade, conformidade e prestação no atendimento. Esse capítulo também apresenta as equações utilizadas para o cálculo dos indicadores da distribuição, assim como as compensações financeiras e a metodologia utilizada para o IASC.

O capítulo 3 apresenta a teoria econométrica utilizada como base para os modelos de RLM e as hipóteses para sua validação, além de uma revisão dos trabalhos que abordaram, através de ferramentas de regressão, a melhoria na qualidade de energia avaliada pelos indicadores de distribuição.

O capítulo 4 introduz a metodologia utilizada para se definir a combinação de regressores à serem avaliados nos modelos propostos e apresenta os testes utilizados para avaliar a significância estatística do modelo e dos regressores individualmente, assim como a teoria baseado nos critérios de seleção dos modelos. Por fim, apresenta as equações dos modelos de RLM propostos neste trabalho.

O capítulo 5 apresenta os parâmetros estimados para o modelo, junto das estatísticas de teste utilizadas para validar as hipóteses do modelo clássico e o de dados em painel. É apresentada também uma discussão dos parâmetros apontando o impacto que cada indicador da distribuição possui na satisfação do consumidor e possíveis ações visando a melhoria da qualidade de energia.

O capítulo 6 finaliza a discussão realizada no capítulo anterior e apresenta uma síntese dos resultados obtidos com os modelos, finalizando com observações pertinentes à realização do trabalho.





## 2 Qualidade da energia elétrica

Em processos de regulação, é de senso comum que a primeira idéia a ser cogitada é a definição de tarifas. Entretanto, o processo de regulação não se limita apenas a modicidade tarifária, a idéia da qualidade também é de extrema importância (BRASIL, 2014).

A qualidade no fornecimento de energia elétrica é conceituada à partir de três dimensões: confiabilidade, conformidade e prestação no atendimento comercial. O módulo 8 do PRODIST, referente à qualidade da energia elétrica, define a confiabilidade como qualidade do serviço, conformidade como qualidade do produto e prestação no atendimento comercial como qualidade do tratamento de reclamações.

A dimensão da confiabilidade diz respeito à capacidade do sistema de assegurar o equilíbrio estrutural entre oferta e demanda, além da robustez voltadas às ocorrências não previstas. A dimensão da conformidade se refere às propriedades físicas, que respeitam determinados padrões, garantindo que oscilações no fornecimento de energia elétrica não prejudiquem a operação de aparelhos elétricos ou até mesmo danificá-los. Já dimensão da prestação no atendimento comercial, mesmo que não esteja associada ao fornecimento físico de energia elétrica, também caracteriza a qualidade de energia de um sistema, considerando pontos como cortesia e prestação (ANEEL, 2008a).

A confiabilidade envolve dois conceitos, a adequabilidade, que está relacionada com o potencial do sistema em atender as demandas do consumidor, e a segurança, que se refere ao potencial do sistema de lidar com ocorrências de origens interna e externa (BRASIL, 2014).

A qualidade do serviço é um dos aspectos mais importantes na qualidade da energia, pois é essencial atender a demanda de forma contínua e sem interrupções. Os indicadores de continuidade podem ser mensurados com relativa facilidade e, através de metas e sanções, controlados pelos agentes reguladores. A regulação na continuidade de fornecimento pode ser realizada avaliando tais indicadores, comparando os valores apurados com seus limites, definidos como metas de continuidade.

De acordo Mamede Filho, o termo confiabilidade está relacionada à capacidade do elemento de proteção realizar com segurança e exatidão suas funções propostas. Um menor número de interrupções, tempos de reestabelecimento menores e, conseqüentemente, clientes mais satisfeitos são vantagens, no contexto da confiabilidade, que um sistema bem projetado e coordenado pode garantir (FILHO, 2005).

A qualidade do produto estabelece a terminologia, caracteriza os fenômenos e define os indicadores e seus limites relacionados à conformidade do nível de tensão em regime permanente e às perturbações que podem ocorrer na forma de onda da tensão. Em outras palavras, a qualidade do produto é a caracterização de sua forma de onda nos componentes de um sistema trifásico e abrange os fenômenos de variação de frequência, variação de

tensão de curta e longa duração, distorções harmônicas e desequilíbrios de tensão e corrente, além das flutuações de tensão (KAGAN et al., 2005).

Já a qualidade do tratamento de reclamações estabelece a metodologia para os cálculos dos indicadores de qualidade comercial (ANEEL, 2008a).

Em síntese, a confiabilidade é a dimensão da qualidade que é caracterizada pela sua falta, isto é, os indicadores estão relacionados a duração e a frequência das interrupções no fornecimento, definidos como indicadores de continuidade. A conformidade retrata a uniformidade do fornecimento de energia elétrica em função das amplitudes e frequências das tensões estabelecidas. Já a presteza no atendimento comercial retrata a agilidade e o cuidado no atendimento comercial ao consumidor (BRASIL, 2014).

De tal maneira, tais indicadores que abrangem as três dimensões da qualidade de energia elétrica no setor de distribuição podem ser definidos como *indicadores da distribuição*. Divergindo desta definição, em alguns trabalhos voltados a qualidade de energia elétrica, certos autores se referem aos indicadores de distribuição como aqueles voltados apenas à dimensão da confiabilidade. Neste trabalho, uma vez que um dos objetivos consiste na avaliação dos indicadores voltados às três dimensões, a definição a ser utilizada será indicadores da distribuição.

Os sistemas elétricos, até mesmo no contexto mundial, podem ser divididos em três setores: geração, transmissão e distribuição. No setor da geração, a energia elétrica é produzida em usinas, principalmente hidrelétricas no Brasil. A tensão elétrica gerada é elevada através de transformadores e conduzida pelo setor de transmissão, que é, posteriormente, reduzida na rede de subtransmissão. Por fim, a tensão elétrica é reduzida novamente e direcionada ao setor de distribuição.

Em termos de aplicações voltadas à confiabilidade, o setor da distribuição difere da geração e transmissão pelo fato de ser mais orientado ao ponto de carga do consumidor ao invés de sistema. A confiabilidade de um sistema de potência, idealmente falando, significa um fornecimento contínuo de energia para os setores de geração, transmissão e distribuição. Entretanto, os principais indicadores voltados à confiabilidade de um sistema de potência são a duração e a frequência das interrupções no ponto de carga, isto é, mais orientado ao ponto de utilização da energia elétrica.

A literatura relata que mais de 80% de todas as interrupções do cliente ocorrem devido a falhas no sistema de distribuição. Embora o reforço no sistema de distribuição seja relativamente barata quando comparada com a geração ou um esquema de melhoria de transmissão, concessionárias de energia elétrica normalmente gastam altas quantias de capitais e orçamentos de manutenção (CHOWDHURY; KOVAL, 2009).

Contando que os sistemas de distribuição sejam responsáveis por até 90% dos problemas voltados à confiabilidade do cliente, melhorar a confiabilidade na distribuição pode ser a chave para melhorar o fornecimento contínuo de energia elétrica para os clientes (BROWN, 2009).

## 2.1 Indicadores da distribuição

Pesquisas voltadas à área de confiabilidade nos setores de transmissão e distribuição, foco que até então era no setor de geração, começaram a surgir na década de 60. As primeiras contribuições foram de Gaver et al. e Todd que apresentaram os cálculos para importantes indicadores de confiabilidade (GAVER et al., 1964) (TODD, 1964).

Os indicadores de continuidade nos sistemas de distribuição de energia elétrica já tem sido estudados tanto no setor elétrico mundial, quanto no brasileiro. No contexto internacional, para as interrupções e os tempos médios de restabelecimento de energia, o *Institute of Electrical and Electronics Engineers* (IEEE) definiu os indicadores *System Average Interruption Frequency Index* (SAIFI), *Customer Average Interruption Frequency Index* (CAIFI), *System Average Interruption Duration Index* (SAIDI) e o *Customer Average Interruption Duration Index* (CAIDI). Esses indicadores fazem referência à frequência e a duração das interrupções, voltados ao consumidor ou ao sistema (IEEE, 2012). Dentre tais indicadores citados, os mais analisados são o SAIFI e o SAIDI, correspondentes aos indicadores Frequência Equivalente de Interrupção por Unidade Consumidora (FEC) e Duração Equivalente de Interrupção por Unidade Consumidora (DEC), definidos pela ANEEL e utilizados no Brasil.

No setor elétrico brasileiro, baseado nos indicadores de continuidade internacionais, a ANEEL estabeleceu seus indicadores visando analisar a qualidade do serviço, do produto e do atendimento ofertados pelas concessionárias de energia elétrica. Prado afirma que a ANEEL possui autonomia gerencial, financeira e competência para normatizar as questões técnicas voltadas ao fornecimento de energia elétrica. Além disso, também possui autonomia decisória, garantidas por sua diretoria, para garantir a qualidade técnica e a neutralidade em sua tomada de decisões (PRADO, 2006).

Para qualquer concessionária, o sistema elétrico de distribuição se divide em conjuntos de unidades consumidoras, os denominados conjuntos elétricos, que podem ser formadas por uma ou mais subestações de distribuição. De acordo com a ANEEL, são definidos como agrupamentos de consumidores com seus níveis de qualidade do serviço especificados dadas às características físicas da rede e do mercado de energia (ANEEL, 2000).

Um conjunto elétrico pode ter áreas de abrangência variadas, podendo alguns abrangerem mais de um município ao mesmo tempo que um município pode conter mais de um conjunto elétrico. Caso uma subestação de distribuição possua um número menor ou igual a 1000 unidades consumidoras, ela pode ser agregada à outras subestações para formar um único conjunto elétrico (ANEEL, 2018).

A apuração dos indicadores da distribuição é de acordo com os conjuntos elétricos, que muitas vezes e à nível de comparação, é agrupado em concessionárias, como por exemplo os indicadores DEC e FEC que são apresentados para cada conjunto elétrico ou para concessionária.

O período de apuração dos indicadores é definido como o intervalo de tempo entre o início e o final da contabilização das interrupções, em outras palavras, constitui o espaço de observação da ocorrência das interrupções, e são apurados como períodos de definição civil mensal, trimestral ou anual. A avaliação dos valores apurados e estipulação das metas é realizada de acordo com a apuração mensal dos últimos doze meses (ANEEL, 2008a).

Na literatura, os trabalhos cuja abordagem foi a redução ou avaliação dos indicadores da distribuição seguiram diferentes metodologias.

Steiner et al. através de técnicas de pesquisa operacional salientam que o redimensionamento ótimo das equipes de manutenção pode reduzir os custos e, conseqüentemente, reduzir os indicadores de continuidade, melhorando assim a confiabilidade do sistema (STEINER et al., 2006).

Pessanha et al. abordaram a regulação dos indicadores de continuidade através de uma metodologia que combinava dois modelos de análise envoltória de dados, propondo uma nova forma de definição de metas de continuidade locais para os indicadores DEC e FEC, dado que o custo de interrupção é distinto entre diferentes classes do setor de distribuição, contribuindo com um modelo que confere maior transparência ao agente regulador (PESSANHA et al., 2007).

Rodriguez aborda o problema da gestão eficiente de tarefas de manutenção de redes elétricas de distribuição através de um modelo de otimização multi-objetivo não linear dinâmico misto, em que se busca minimizar os custos das tarefas de manutenção nos sistemas de distribuição, maximizando sua confiabilidade, sob as restrições dos indicadores individuais de continuidade Duração de Interrupção Individual por Unidade Consumidora ou por Ponto de Conexão (DIC) e Frequência de Interrupção Individual por Unidade Consumidora ou por Ponto de Conexão (FIC), além de outros fatores (RODRIGUEZ, 2017).

Já Boff et al. abordaram a redução dos indicadores da distribuição, baseado no histórico de falhas do sistema em alimentadores da concessionária Companhia Paulista de Força e Luz (CPFL), via aplicação de um algoritmo evolutivo multi-objetivo e o algoritmo genético de classificação não dominado elitista II (NSGA-II), apresentando soluções em custo, em continuidade e as dominantes, para o problema de otimização abordado (BOFF et al., 2017).

Também seguindo a abordagem de otimização, León e Leite utilizaram a técnica NSGA-II na obtenção de uma função ótima que represente os custos dos consumidores e das concessionárias, baseado no histórico de falhas devido ao crescimento de vegetação nas proximidades dos dispositivos do sistema e os indicadores da distribuição, apresentando uma quantidade aceitáveis de soluções (LEON; LEITE, 2020).

Mondani et al. analisaram a relação entre os indicadores de custo, consumo de energia elétrica e receita líquida da venda dos serviços, além dos indicadores de qualidade do serviço. Os autores utilizaram como metodologia a análise de cluster, além de testes de

hipóteses, e observaram que não existem comprovações estatísticas de uma relação entre tais indicadores (MONDANI et al., 2017).

Nunes et al. avaliaram o impacto, baseado em modelos markovianos, das redes de transmissão e subtransmissão nos indicadores de continuidade, incluindo no modelo as falhas dos principais equipamentos de subestações e linhas de transmissão (NUNES et al., 2018).

Oliveira et al. avaliaram a existência de correlação entre o desempenho em indicadores de qualidade e o desempenho financeiro de organizações do setor elétrico, utilizando como variáveis os indicadores de continuidade, financeiros, de liquidez e tempos de atendimento às ocorrências emergenciais. A metodologia utilizada consistiu em estatísticas não paramétricas e no uso do coeficiente de correlação de Spearman que, diferente do coeficiente de Pearson que é utilizado para avaliar a relação linear entre duas variáveis, pode ser utilizado para avaliar a relação monotônica entre duas variáveis ordinárias.

Os autores concluíram que a correlação entre alguns indicadores de continuidade e financeiros possuem certa correlação apenas em um intervalo do período total amostrado. Os autores também observaram que os indicadores associados aos tempos de atendimento às ocorrências emergenciais e à liquidez não apresentam correlação para nenhum intervalo e nenhum outro indicador (OLIVEIRA et al., 2019).

Torres et al. abordaram o problema de reconfiguração da rede elétrica através do algoritmo busca tabu, avaliando como uma das restrições operacionais o indicador de continuidade FIC. Através da metodologia proposta, os autores foram capazes de propor uma topologia radial eficiente para o problema (TORRES et al., 2019).

Vieira et al., baseados na teoria dos testes de hipóteses, procuraram verificar a existência de uma relação significativa entre o número de interrupções com o as ocorrências de chuvas. Os autores utilizaram como base de dados setenta municípios do estado do Ceará e observaram que, para esta amostra, há maiores interrupções no sistema de distribuição em dias de chuva (VIEIRA et al., 2020).

Paiva et al. realizaram uma estimação de estados trifásica, baseada em matriz de admitância, nos indicadores de conformidade de tensão em regime permanente. Os autores observaram que tal estimação pode contribuir na apuração dos indicadores Duração Relativa da Transgressão de Tensão Precária (DRP) e Duração Relativa da Transgressão de Tensão Crítica (DRC), possibilitando a identificação de manipulações nos valores de magnitude de tensão (PAIVA et al., 2021).

### 2.1.1 Confiabilidade no fornecimento de energia elétrica

As interrupções no fornecimento de energia elétrica podem ser entendidas como a ausência de energia elétrica. Em sua resolução de nº 24, a ANEEL define a interrupção como uma descontinuidade do neutro ou da tensão em qualquer uma das fases do circuito da unidade consumidora. O retorno, para o consumidor, do neutro e da tensão elétrica

em todas as fases pode ser chamado de *Reestabelecimento da Continuidade de Energia Elétrica* (ANEEL, 2000).

As interrupções no fornecimento de energia podem ser classificadas em programadas ou não programadas. As interrupções do tipo não programada têm sua abrangência, em uma linha de distribuição de energia elétrica, que pode ir desde a fonte de geração até o consumidor final. De acordo com a ELETROBRAS (1982) as causas que podem acarretar em interrupções do tipo não programadas podem ser divididos em (ELETROBRAS, 1982):

- *meio ambiente*: poluição, corrosão, fogo, inundação, erosão, animais, etc;
- *terceiros*: vandalismo, acidentes e empresas de serviços públicos ou contratadas;
- *falha humana*;
- *próprias do sistema*: subtensão, sobretensão ou características construtivas do equipamento;
- *falha de componente*;
- *outro órgão ou companhia*;
- *outras* ou *desconhecidas*.

Os indicadores de continuidade são apurados considerando apenas as interrupções de longa duração, que de acordo com a ANEEL, são todas interrupções do sistema elétrico com duração maior ou igual à três minutos (ANEEL, 2000).

Os indicadores de duração expressam o intervalo de tempo, de forma média, que cada unidade consumidora ou concessionária (de forma equivalente) ficaram desprovidos do fornecimento de energia elétrica, dado um período de apuração, enquanto que os indicadores de frequência expressam o número de interrupções que, em média, cada unidade consumidora ou concessionária sofreram dado um período de apuração, que pode ser mensal ou anual.

Os indicadores de duração e frequência de interrupção são apurados como coletivos e individuais, que por sua vez são dispostas para cada unidade consumidora e independente das unidades vizinhas. Os indicadores de continuidade coletivos são apresentados unicamente como parâmetros de qualidade para a verificação da performance da concessionária e definição de limites para os indicadores de continuidade individuais, que por sua vez são apresentados não só como parâmetros de qualidade do serviço, mas também como objetos de cálculo das compensações financeiras a serem creditadas na forma de desconto nas faturas de consumo, para todas as unidades consumidoras (ANEEL, 2008a).

É importante citar que esses indicadores coletivos não refletem a continuidade do fornecimento de cada unidade consumidora individualmente, uma vez que reflete o estado equivalente da continuidade do fornecimento do conjunto elétrico considerado, em outras

palavras, esses indicadores representam de forma média as unidades consumidoras individuais. Guder et al. (2006) salientam que o diferencial entre os indicadores de continuidade individuais e coletivos é a capacidade de oferecer ao consumidor garantias reais de serviço, uma vez que em grandes redes de distribuição podem existir conjuntos com qualidade do serviço distintos, características que os indicadores coletivos não podem agregar (GUDER et al., 2008).

Os indicadores individuais de continuidade *Duração de Interrupção Individual por Unidade Consumidora ou por Ponto de Conexão* (DIC) e *Frequência de Interrupção Individual por Unidade Consumidora ou por Ponto de Conexão* (FIC) são calculados pelas equações:

$$DIC = \sum_{i=1}^n t_i \quad (2.1)$$

e

$$FIC = n, \quad (2.2)$$

onde  $t_i$ , da equação 2.1, é a duração da interrupção expressa em horas e  $n$ , da equação 2.2, é o número de interrupções da unidade consumidora ou por ponto de conexão, consideradas no período de apuração.

A ANEEL, em sua resolução normativa de nº 482, define um dia crítico como aquele em que o número de ocorrências emergenciais superar a média acrescida de três desvios padrões dos valores diários, dado um certo conjunto de unidades consumidoras. Nessa mesma resolução, define-se também a interrupção em situação de emergência, como aquela motivada por um caso fortuito ou de força maior, que deve ser comprovada documentalmente pela distribuidora (ANEEL, 2012).

Outros indicadores individuais, também apurados em uma unidade consumidora ou por um ponto de conexão, são o *Duração da Interrupção Individual ocorrida em Dia Crítico* (DICRI) e o *Duração Máxima de Interrupção Contínua* (DMIC), que representa o tempo da máxima duração de interrupção contínua ocorrida em uma unidade consumidora ou ponto de conexão (ANEEL, 2008a).

Os indicadores coletivos de continuidade são expressos através da razão que considera o número de unidades consumidoras em cada conjunto, ou concessionária, atendidos em baixa tensão (BT) ou média tensão (MT). Tais indicadores são definidos como *Duração Equivalente de Interrupção por Unidade Consumidora* (DEC) e *Frequência Equivalente de Interrupção por Unidade Consumidora* (FEC), que são calculadas pelas seguintes equações:

$$DEC = \frac{\sum_{i=1}^N DIC_i}{N} \quad (2.3)$$

e

$$FEC = \frac{\sum_{i=1}^N FIC_i}{N} \quad (2.4)$$



em que  $N$ , das equações 2.3 e 2.4, é o número total de unidades consumidoras faturadas do conjunto elétrico no período de apuração (ANEEL, 2008a).

O Módulo 8 do PRODIST também apresenta os fatos geradores das interrupções do tipo programada e não programada e suas diferentes causas, indicados no quadro 2.1.

Tabela 2.1 – Lista dos fatores geradores.

Origem	Tipo	Causa	Detalhe
Interna	Programada	Alteração	Para melhoria
			Para ampliação
		Manutenção	Corretiva
			Preventiva
	Não-Programada	Meio ambiente	Poluição
			Corrosão
			Queima ou incêndio
			Inundação
			Erosão
			Árvore ou Vegetação
			Descarga Atmosférica
			Animais
			Vento
			Terceiros
		Abalroamento	
		Roubo	
		Acidente	
		Objeto na Rede	
		Defeito cliente afetando outros	
		Ligação clandestina	
		Empresas de serviços públicos ou suas contratadas	
		Defeito interno não afetando outras unidades consumidoras	
		Interferência de terceiros	
		Falha operacional	Erro de operação
	Serviço mal executado		
	Acidente		
Próprias do sistema	Subtensão		
	Sobretensão		
	Sobrecarga		
			Desligamento para manutenção emergencial

Em relação aos serviços prestados pelas empresas de distribuição de energia, os mesmos podem ser classificados como comerciais ou emergenciais. Os serviços comerciais abrangem atividades como ligações de unidades consumidoras, confirmações de leituras, cortes de energia de consumidores inadimplentes, aumentos de carga, além de outros. Já em relação aos serviços emergenciais, as atividades abordam cabos caídos, cortes indevidos, ramais derretidos, além de roubos de ramais de ligação.

As chamadas comerciais podem ser caracterizadas como determinísticas, enquanto as chamadas emergenciais são estocásticas e estratificadas. Diferente dos serviços comerciais,



que em casos de infortúnios ou imprevistos podem ser reagendados com compensação financeira por atrasos, os emergenciais, como o próprio nome sugere, devem ser executados no menor tempo possível, uma vez que podem acarretar em maiores prejuízos (STEINER et al., 2006).

Para a apuração dos indicadores associados às ocorrências emergenciais, as distribuidoras não devem considerar os casos de (ANEEL, 2008a):

- Solicitação de serviços voltados à iluminação pública;
- Serviços de caráter comercial;
- Reclamações voltadas ao nível de tensão;
- Reclamações à interrupções de energia em razão de manutenção programada previamente comunicada;
- Interrupções em situação de emergência.

Figura 2.1 – Estratificações das interrupções de longa duração.

	X		I
P	XP	XPC	IPC
NP	XN	XNC	IND INO INE

Fonte: (ANEEL, 2008a).

À partir de 2010, a ANEEL passou a apresentar as estratificações para os indicadores DEC e FEC, figura 2.1, onde as interrupções de energia elétrica são classificados da seguinte forma:

- XP : origem externa ao sistema de distribuição e programada;
- XN : origem externa ao sistema de distribuição e não-programada;
- IP : origem interna ao sistema de distribuição e programada;
- IND : origem interna ao sistema de distribuição, não-programada e não expurgável;
- INE: origem interna ao sistema de distribuição, não-programada e ocorrida em situação de emergência;

- INC: origem interna ao sistema de distribuição, não-programada e ocorrida em dia crítico;
- INO: origem interna ao sistema de distribuição, não-programada, vinculadas a programas de racionamento da União e provindas do Esquema Regional de Alívio de Carga da ONS;
- IPC: origem interna ao sistema de distribuição, programada e ocorrida em dia crítico;
- XPC: origem externa ao sistema de distribuição, programada e ocorrida em dia crítico;
- XNC: origem externa ao sistema de distribuição, não-programada e ocorrida em dia crítico.

Na composição dos indicadores DEC e FEC analisados neste trabalho, a ANEEL considera apenas as estratificações XP, XN, IP e IND, representados em branco na figura 2.1, segregando as interrupções ocorridas em dia crítico, em situação de emergência e programas de racionamento da União e ONS. No mesmo repositório em que ANEEL disponibiliza tais estratificações, ela apresenta também os indicadores DECTOT e FECTOT, que representam a soma para todas as estratificações de DEC e FEC (ANEEL, 2021a).

Ainda no contexto da qualidade do serviço, a ANEEL apresenta também os indicadores voltados aos tempos médios de atendimento às ocorrências emergenciais, definidos à seguir, sendo apurados para cada conjunto elétrico, mensalmente, pelas concessionárias.

O *Tempo Médio de Preparação* (TMP) é um indicador que avalia o quão eficiente são os meios de comunicação, o dimensionamento das equipes e os fluxos de informação dos centros de operação. Seu cálculo é realizada à partir da equação 2.5,

$$TMP = \frac{\sum_{i=1}^n TP_i}{n} \quad (2.5)$$

sendo  $i$  o índice relacionado à  $i$ -ésima interrupção das  $n$  ocorrências emergenciais verificadas avaliada para os conjuntos de unidades consumidoras, dentro do período de apuração considerado.

$$TMD = \frac{\sum_{i=1}^n TD_i}{n} \quad (2.6)$$

O *Tempo Médio de Deslocamento* (TMD), calculado pela equação 2.6, é um indicador que avalia o quão eficaz é a localização geográfica das equipes responsáveis pela manutenção e operação do sistema de distribuição.

$$TME = \frac{\sum_{i=1}^n TE_i}{n} \quad (2.7)$$

Já o *Tempo Médio de Execução* (TME), calculado pela equação 2.7, é o indicador que mede a eficácia do reestabelecimento do sistema de distribuição realizado pelas equipes.

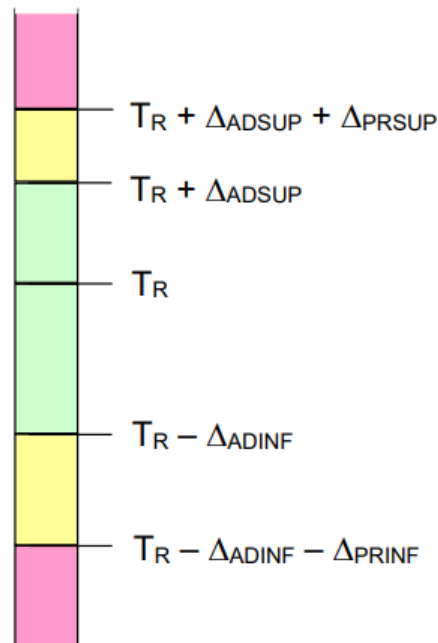
É o indicador que avalia o tempo médio despendido por uma equipe para o tratamento de alguma ocorrência emergencial, após a preparação e o deslocamento da equipe até o local da ocorrência (ANEEL, 2008a).

### 2.1.2 Conformidade do nível de tensão

Os indicadores de conformidade do nível de tensão, que avaliam a qualidade do produto, podem ser divididos em permanentes ou transitórios. Para este trabalho serão considerados apenas os indicadores em regime permanente, uma vez que são os apurados pelas concessionárias, de acordo com o PRODIST (ANEEL, 2008a).

A tensão de atendimento é classificada de acordo com uma tensão de referência  $T_R$  e a ANEEL define as faixas de tensão como adequada, precária e crítica. Suas faixas de variação são distintas para cada nível de tensão, sendo que as faixas de variação são menores para altos valores de tensão.

Figura 2.2 – Faixas de tensão em relação à de tensão de referência  $T_R$ .



Fonte: (ANEEL, 2008a).

De acordo com a figura 2.2, que apresenta os intervalos para as diferentes classificações das faixas de tensão, baseados na tensão de referência  $T_R$ , a região em verde caracteriza a faixa de tensão adequada, matematicamente representada por todas as tensões dentro dos limites estabelecidos por  $(T_R - \Delta_{ADINF}, T_R + \Delta_{ADSUP})$ , em que  $\Delta_{ADINF}$  e

$\Delta_{ADSUP}$  representam os limites para as variações de tensão inferiores e superiores, respectivamente, para a faixa adequada.

A faixa em amarelo caracteriza a faixa de tensão precária, representada por todas as tensões dentro dos limites estabelecidos por  $(T_R - \Delta_{ADINF} - \Delta_{PRINF}, T_R + \Delta_{ADSUP} + \Delta_{PRSUP})$ , em que  $\Delta_{PRINF}$  e  $\Delta_{PRSUP}$  representam os limites para as variações de tensão inferiores e superiores, respectivamente, para a faixa precária. Já a faixa em vermelho caracteriza a faixa de tensão crítica, que representa todas as tensões que não estejam dentro do intervalo para as faixas adequada e precária.

Fundamentada em tais definições para as classificações das faixas de tensão, a ANEEL define certos indicadores individuais e coletivos associados à conformidade do nível de tensão.

Os indicadores individuais *Duração Relativa da Transgressão de Tensão Precária* (DRP) e *Duração Relativa da Transgressão de Tensão Crítica* (DRC), associados a um mês civil, são calculados de acordo com as seguintes equações:

$$DRP = \frac{nlp}{1008} \cdot 100[\%] \quad (2.8)$$

e

$$DRC = \frac{nlc}{1008} \cdot 100[\%], \quad (2.9)$$

em que  $nlp$ , da equação 2.8, e  $nlc$ , da equação 2.9, representam o maior valor entre as fases do número de leituras situadas nas faixas precária e crítica, respectivamente. O parâmetro 1008, definido pelo PRODIST, representa o número total de leituras válidas associadas a um mês civil.

Já os indicadores coletivos de *Duração Relativa de Transgressão de Tensão Precária Equivalente* (DRPE), equação 2.10, e *Duração Relativa de Transgressão de Tensão Crítica Equivalente* (DRCE), equação 2.11, são definidos por

$$DRPE = \frac{\sum DRP_i}{N_L} [\%] \quad (2.10)$$

e

$$DRCE = \frac{\sum DRC_i}{N_L} [\%], \quad (2.11)$$

onde  $N_L$  é o total de unidades consumidoras avaliadas em determinada concessionária (ANEEL, 2008a).

### 2.1.3 Qualidade do tratamento de reclamações

Na avaliação da qualidade do tratamento de reclamações, a ANEEL define dois indicadores globais, o *Duração Equivalente de Reclamação* (DER) e *Frequência Equivalente de Reclamação a cada mil unidades consumidoras* (FER), regulamentado pelas *Condições Gerais de Fornecimento* (resolução normativa nº 414 de 2010) (ANEEL, 2010).

Na apuração dos indicadores voltados a qualidade do tratamento de reclamações, as concessionárias devem computar as reclamações efetuadas em todos seus meios disponibilizados, tais como centrais ou postos fixos de atendimento, internet e correspondências. Além disso devem apurar, mensalmente, as seguintes informações:

- quantidade de reclamações recebidas;
- quantidade de reclamações procedentes;
- quantidade de reclamações improcedentes;
- prazo médio de solução das reclamações procedentes.

O prazo de solução de uma reclamação, expresso em horas e centésimos de horas, é definida como o intervalo de tempo entre o recebimento e a sua solução, por parte da distribuidora. Em casos que os consumidores reclamarem reiteradas vezes sobre o mesmo problema, considera-se apenas a primeira reclamação para a apuração desses indicadores.

Com essas informações apuradas pelas distribuidoras, os indicadores anuais DER, equação 2.12, e FER, equação 2.13, são calculados por

$$DER = \frac{\sum_{i=1}^n \text{Reclamações Procedentes}_i \cdot PMS_i}{\sum_{i=1}^n \text{Reclamações Procedentes}_i} \quad (2.12)$$

e

$$FER = \frac{\sum_{i=1}^n \text{Reclamações Procedentes}_i}{N_L} \cdot 1000, \quad (2.13)$$

onde  $i$  indica o tipo de reclamação, dos  $n$  tipos definidos em anexo na resolução normativa de nº 414 da ANEEL, que basicamente são resumidos em informação, reclamação, solicitação de serviços e de urgência ou emergência.  $\text{Reclamações Procedentes}_i$  indica a quantidade de reclamações procedentes do consumidor do tipo  $i$  solucionadas pela distribuidora no período de apuração e  $PMS_i$  indica o prazo médio de solução das reclamações procedentes do tipo  $i$  (ANEEL, 2010).

#### 2.1.4 Resumo dos indicadores da distribuição

Os indicadores da distribuição que foram utilizados como regressores estão resumidos na tabela 2.2.

Tabela 2.2 – Resumo dos indicadores da distribuição aplicados no modelo.

Indicadores	Individuais	Coletivos
Continuidade do fornecimento	DIC, FIC, DMIC e DICRI	DEC e FEC
Conformidade do nível de tensão	DRP e DRC	DRPE e DRCE
Ocorrências emergenciais	-	TMP, TMD e TME
Tratamento de reclamações	-	DER e FER

## 2.2 Compensações financeiras pagas aos consumidores

O módulo 8 do PRODIST define as compensações financeiras à serem pagas pelas concessionárias aos consumidores em casos de violação dos limites para os indicadores individuais, sendo tais compensações um dos principais estímulos à melhoria na prestação de serviços de fornecimento de energia elétrica. Essas compensações financeiras representam descontos nos valores apurados das faturas de cada uma das unidades consumidoras que foram afetadas pelo serviço de fornecimento de energia abaixo do nível de qualidade estipulado.

No caso da violação dos limites de continuidade individuais DIC, FIC, DMIC e DICRI, as distribuidoras de energia devem calcular as compensações ao consumidor e efetuar o crédito na fatura ou deduzir débitos vencidos, que devem ser apresentados em até dois meses após o período de apuração.

$$COMP_{DIC} = \left( \frac{DIC_v}{DIC_p} - 1 \right) \cdot DIC_p \cdot \frac{EUSD_{médio}}{730} \cdot kei \quad (2.14)$$

$$COMP_{DMIC} = \left( \frac{DMIC_v}{DMIC_p} - 1 \right) \cdot DMIC_p \cdot \frac{EUSD_{médio}}{730} \cdot kei \quad (2.15)$$

$$COMP_{FIC} = \left( \frac{FIC_v}{FIC_p} - 1 \right) \cdot FIC_p \cdot \frac{EUSD_{médio}}{730} \cdot kei \quad (2.16)$$

$$COMP_{DICRI} = \left( \frac{DICRI_v}{DICRI_p} - 1 \right) \cdot DICRI_p \cdot \frac{EUSD_{médio}}{730} \cdot kei \quad (2.17)$$

Para o cálculo de tais compensações são utilizadas as fórmulas 2.14, 2.15, 2.16 e 2.17, para os indicadores DIC, DMIC, FIC e DICRI, respectivamente, onde os índices  $v$  e  $p$  apresentados representam o valor verificado e o limite estabelecido, respectivamente, para os indicadores em questão.

$EUSD_{médio}$  representa a média aritmética dos encargos do uso do sistema de distribuição correspondentes aos meses do período de apuração do indicador em questão. O parâmetro  $kei$  representa o coeficiente de majoração, e assume os valores de acordo

com os níveis de tensão da unidade consumidora ou ponto de conexão atendidos, para alta tensão (AT) fixa-se o valor 27, MT o valor 20 e para BT o valor 15, definidas pelo PRODIST. Já o parâmetro 730 representa o número médio de horas por mês (ANEEL, 2008a).

O período de apuração para os indicadores DIC, FIC e DMIC pode ser mensal, trimestral e anual, enquanto que para o DICRI a distribuidora deve efetuar uma compensação para cada interrupção ocorrida em dia crítico que supere os limites de DICRI.

$$COMP_S = COMP_{DIC} + COMP_{DMIC} + COMP_{FIC} \quad (2.18)$$

$$COMP_{SC} = COMP_{DICRI} \quad (2.19)$$

Na avaliação de tais compensações financeiras no modelo, resumiu-se tais compensações nas variáveis  $COMP_S$  e  $COMP_{SC}$ , calculadas por 2.18 e 2.19, para possibilitar a avaliação de forma individual do impacto das compensações financeiras voltadas aos indicadores individuais de continuidade ocorridas em dia crítico.

Além das compensações relacionadas à violação dos indicadores de continuidade, as concessionárias também devem apurar compensações devido à transgressão dos limites voltados aos indicadores de conformidade do nível de tensão individuais, o DRP e DRC.

De acordo com o PRODIST, o limite do indicador DRP é de 3%, enquanto que para o indicador DRC é de 0,5%. O valor das compensações voltados à qualidade do produto, equação 2.20, é calculado por

$$COMP_P = \left[ \left( \frac{DRP - DRP_{limite}}{100} \right) \cdot k_1 + \left( \frac{DRC - DRC_{limite}}{100} \right) \cdot k_2 \right] \cdot EUSD \quad (2.20)$$

onde a constante  $k_1$  é igual à zero se o valor apurado de DRP for menor ou igual que seu valor limite, caso contrário,  $k_1$  é igual a três (3). A constante  $k_2$  é igual à zero se o valor apurado de DRC for menor ou igual seu valor limite, caso contrário, é de 7, para consumidores em BT, de 5 para consumidores em MT e de 3 para consumidores em AT (ANEEL, 2008a).

## 2.3 Índice ANEEL de satisfação do consumidor

Nos últimos anos, satisfazer os clientes se tornou um dos pilares nas estratégias de mercado de empresas de diversos nichos, dada que a satisfação de um cliente está diretamente associada com a recompra do produto ou serviço (MITTAL; KAMAKURA, 2001).

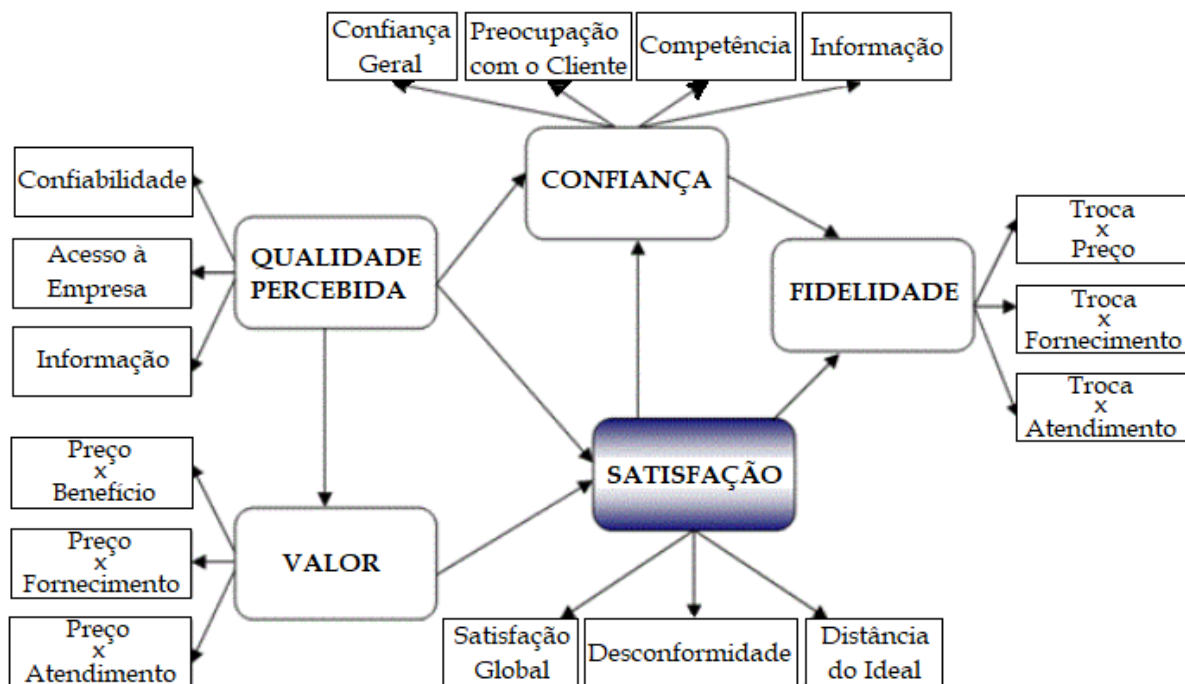
Em decorrência desse entendimento da importância da satisfação por parte do cliente, começaram a surgir métodos para mensurar-lá. Ao fim dos anos 80, na Suécia,

surgiu o *Swedish Customer Satisfaction Index* (SCSI), baseado nos 31 principais ramos das atividades econômicas. Assim como a Suécia, alguns outros países como Alemanha, Israel, Nova Zelândia e Taiwan, foram pioneiros nos trabalhos voltados à esses métodos. Entretanto, seu marco ocorreu em 1994 com o *American Customer Satisfaction Index* (ACSI), que posteriormente inspirou o *European Customer Satisfaction Index* (ECSI), introduzido em 2001 (LOPES et al., 2009).

Em relação aos serviços prestados pelas distribuidoras de energia elétrica brasileiras, a ANEEL apresenta o índice ANEEL de satisfação do consumidor, o IASC. Os resultados do IASC são aplicados no aprimoramento dos instrumentos regulatórios e na priorização das ações de fiscalização (ANEEL, 2015).

Para a metodologia a ser estabelecida para o cálculo do IASC foram consolidadas discussões entre os representantes da ANEEL, das agências estaduais conveniadas e das distribuidoras de energia elétrica e, à partir de uma pesquisa qualitativa realizada no ano de 2000, foi proposto o modelo representado na figura 2.3.

Figura 2.3 – Modelo para o cálculo do IASC.



Fonte: (ANEEL, 2015) adaptado.

O modelo é composto de cinco variáveis, avaliadas à partir de um questionário aplicada por meio de escalas de mensuração, sendo elas: qualidade percebida, valor percebido, satisfação, confiança no fornecedor e fidelidade (ANEEL, 2015). Dentre tais variáveis, utilizou-se a qualidade percebida, uma vez que esta abrange os itens que,



baseados na satisfação do consumidor, correspondem aos itens avaliados pelos indicadores da distribuição estudados neste trabalho.

A qualidade percebida é avaliada sob três *escores*: informações ao cliente, acesso à empresa e confiabilidade nos serviços. Uma vez que os *escores* informações ao cliente e acesso à empresa são compostos apenas por itens que são voltadas a presteza, acessibilidade e cordialidade no atendimento comercial aos consumidores, utilizou-se apenas o *escore* confiabilidade nos serviços, que abrange os seguintes itens (ANEEL, 2015):

- Fornecimento de energia sem interrupção;
- Fornecimento de energia sem variação na tensão;
- Avisos antecipados sobre o corte de energia devido à falta de pagamento;
- Confiabilidade nas soluções dadas;
- Rapidez na volta da energia quando há interrupção;
- Avisos antecipados sobre o desligamento de energia devido à manutenção.

Uma vez que a apuração anual do IASC é realizada à partir da metade do ano, podendo variar os meses para cada ano, nos modelos à serem propostos, será considerado o aumento em uma unidade de tempo (um ano) para o IASC, indicado por  $t+1$  nas equações nos capítulos de metodologia e resultados.



## 3 Teoria Econométrica

Neste capítulo são apresentados os conceitos, baseados na teoria econométrica, e a revisão bibliográfica de trabalhos relacionados, que foram utilizados como referência para o desenvolvimento dos modelos de RLM.

Primeiramente, são introduzidos os conceitos das diferentes estruturas de dados utilizadas como amostra. Em seguida, são apresentados as definições dos modelos de RLM e as hipóteses para a validação do teorema de *Gauss-Markov*.

Por fim, são apresentados os métodos aplicados na estimação dos parâmetros para os modelos de dados em painel, estrutura de dados que corresponde à séries temporais avaliadas para diferentes unidades de corte transversal, o que envolve uma estimação dos parâmetros mais complexa quando comparada a estimação para séries temporais ou cortes transversais.

### 3.1 Estrutura de dados

Diversos autores de livros sobre Econometria classificam as estruturas de dados em: corte transversal, série temporal, corte transversal agrupado e dados em painel ou longitudinal. Para cada uma das diferentes estruturas de dados, os testes realizados e análise dos efeitos possuem suas singularidades. Por exemplo, em modelos de séries temporais e dados em painel que avaliam a variável tempo nos modelos de regressão, deve-se considerar possíveis efeitos de correlação serial.

Um conjunto de dados de corte transversal basicamente é uma amostra de uma certa unidade de análise que é observada em um período de tempo específico, em outras palavras, no mesmo ponto de tempo. Essa estrutura de dados é fundamental para testar hipóteses, além de avaliar o efeito de políticas, sendo muito utilizada em economia e ciências sociais. Pode consistir de uma amostra de empresas, consumidores, cidades, indivíduos entre outras unidades, tomadas em um determinado ponto do tempo ([WOOLDRIDGE, 2010](#)).

Um conjunto de dados estruturado como uma série temporal é uma coleção de observações amostradas sequencialmente ao longo do tempo de forma que observações em períodos vizinhos são dependentes. Uma série temporal é um conjunto de observações ordenadas no tempo apresentando dependência entre seus diferentes instantes ([GUJARATI; PORTER, 2011](#)).

Como exemplos de uma série temporal temos o PIB de um país, o histórico do consumo energética de uma determinada consumidora, entre vários outros estudados na área de Econometria.

Em uma análise de regressão para dados em séries temporais, a identificação de padrões não aleatórios na série da variável de interesse é essencial, e sua observação em períodos passados permite realizar previsões sobre o futuro (JUNIOR, 2016).

Um conjunto de dados pode conter características tanto de uma série temporal quanto de um corte transversal, como por exemplo uma pesquisa feita com certas famílias à respeito do número de filhos e renda familiar realizada em 2010 e 2020. Com o intuito de se aumentar o tamanho da amostra, pode-se combinar os dois anos para formar o que se define como corte transversal agrupado, sendo eficiente para analisar os efeitos de novas políticas governamentais.

Um conjunto de dados em painel, também conhecido como dados longitudinais, basicamente são séries temporais para cada um dos indivíduos do corte transversal do conjunto de dados. Um exemplo muito discutido no cenário político é o PIB para um grupo de países ao longo de anos (GUJARATI; PORTER, 2011).

Assim como os cortes transversais agrupados, os dados em painel representam uma boa estratégia para o aumento da amostra. Dados em painel são, em certos casos, mais difíceis de se obter, uma vez que dados em painel consistem em observações das mesmas unidades de corte transversal. Entretanto, os economistas reconhecem que os dados em painel apresentam determinadas vantagens que são essenciais em algumas questões, como o controle sobre as características não observáveis dos indivíduos, avaliar as defasagens do comportamento ou na tomada de decisões (WOOLDRIDGE, 2010).

Uma vez que o conjunto de dados utilizados neste trabalho consiste em variáveis associadas às diferentes concessionárias do Brasil apuradas para diferentes anos, foi imprescindível o estudo das diferentes estruturas de dados em um modelo de regressão, assim como as hipóteses para validação de tais modelos.

## 3.2 Análise de regressão linear múltipla

Muitos problemas atuais de diversas linhas de pesquisa envolvem a exploração das relações entre duas ou mais variáveis, dada a limitação da análise entre apenas um regressando e um regressor. Através de medições ou coletas, faz-se uso das informações obtidas através de relações empíricas. Avaliar o comportamento de saída, com a variação de uma ou mais variáveis de entrada, viabilizando inferências sobre a variável de saída assim como a predição do seu comportamento, é um dos objetivos na utilização de modelos de RLM (MONTGOMERY; RUNGER, 2009).

A regressão é uma das ferramentas mais utilizadas em análise de dados, encontrando aplicações em diversas áreas do conhecimento, como engenharia, ciências biológicas, administração, dentre outras.

Diferente da análise de regressão linear simples (RLS), a análise de RLM permite que vários fatores observados afetem a variável dependente. Em uma análise de RLM o

pesquisador procura determinar as variáveis explicativas que melhor realizam previsões da variável dependente (HAIR et al., 2005).

De maneira geral, o modelo de RLM pode ser definido através da função de regressão populacional (FRP), definida por

$$y = \beta_0 + \beta_1.x_1 + \beta_2.x_2 + \dots + \beta_k.x_k + \mu, \quad (3.1)$$

onde  $\beta_0$  é denominado intercepto, que pode ser interpretado como o valor da variável dependente condicionado à  $x_1 = x_2 = x_3 = \dots = x_k = 0$ , sendo  $k$  o número de variáveis explicativas (também conhecidos como regressores) no modelo. Já os  $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ , também denominados parâmetros de inclinação, são os parâmetros associados às variáveis explicativas  $x_1, x_2, x_3, \dots, x_k$ , respectivamente.

O termo  $\mu$  da equação 3.1, denominado termo de erro ou perturbação é inserido no modelo por conter fatores, além dos fatores explicados por  $x_1, x_2, x_3, \dots, x_k$ , que possam influenciar na variável dependente.

Em outras palavras, por mais que um modelo possa ter muitas variáveis independentes, sempre existe a possibilidade de que certos fatores, desconhecidos do pesquisador, da variável dependente ainda não são explicadas por tais variáveis independentes. Assim, considera-se que esses fatores não observados estarão contidos, coletivamente, nesse termo de erro (WOOLDRIDGE, 2010).

Os primeiros trabalhos relacionados a análise de regressão linear surgiram em 1885, com o trabalho desenvolvido por Galton, em que o mesmo constatou que a altura média dos filhos com pais de uma certa altura tendiam a diminuir (GALTON, 1885).

No setor elétrico, alguns trabalhos já utilizaram modelos de RLM para verificar a relação entre variáveis do setor. Moghram e Rahman aplicaram a RLM nas previsões de carga elétrica por hora durante as estações de pico de inverno e verão. Fora a RLM, os autores aplicaram mais quatro outras técnicas voltadas à previsão de carga de curto prazo para, posteriormente, realizar uma comparação entre elas (MOGHRAM; RAHMAN, 1989).

Hong et al. aplicaram a RLM para desenvolver um modelo da demanda de carga elétrica por hora e investigar a causalidade do consumo de energia elétrica, aplicado a previsão de demanda de energia de uma concessionária dos Estados Unidos (HONG et al., 2010).

Araújo Jr. se baseou em um modelo de RLM como ferramenta para verificar a influência dos modos de falha em um sistema de distribuição, priorizando os detectados por termografia, nos valores dos indicadores de continuidade DEC e FEC (JUNIOR, 2016).

Magalhães realizou uma comparação entre as previsões dos indicadores de continuidade à partir dos modos de falhas, obtidas por Redes Neurais Artificiais e RLM (MAGALHAES, 2017).

Maestri e Andrade procuraram identificar a existência da relação entre os custos operacionais, investimentos, satisfação do consumidor e qualidade de energia elétrica, utilizando como abordagem modelos econométricos baseados em dados em painel estáticos e dinâmicos, *GMM* e *GMM System*. Os autores concluíram que a taxa de variação dos indicadores DEC e FEC carregam informações do período anterior, característica que da continuidade na melhoria da qualidade de energia elétrica. Os autores também concluem que um aumento nos investimentos na rede elétrica proporciona uma redução da frequência das interrupções (FEC) e das reclamações (FER), além de que o IASC possui relação negativa com a variação de DEC, DER e FER (MAESTRI; ANDRADE, 2018).

Ferreira et al. procuraram avaliar o impacto dos investimentos em manutenção nos indicadores de qualidade, para as linhas aéreas da concessionária brasileira *Light*, através de duas metodologias, *Machine Learning* e regressão com dados em painel. Os autores concluíram que os resultados obtidos com os modelos de regressão foram efetivos em razão dos pequenos erros nas estimativas dos indicadores DEC, FEC e compensações financeiras (FERREIRA et al., 2020).

### 3.3 Hipóteses clássicas do modelo de Regressão Linear Múltipla

Uma das etapas do desenvolvimento de um modelo de RLM consiste em aplicar certos testes de hipótese para a adequação dos modelos desenvolvidos (MONTGOMERY; RUNGER, 2009).

Como uma das hipóteses essenciais para o modelo de RLM deve-se considerar a esperança do termo de erro condicionada às explicativas igual a zero, representado por

$$E(u|x_1, x_2, x_3, \dots, x_k) = 0. \quad (3.2)$$

Em outras palavras, todos os fatores no termo de erro não podem ser correlacionados com as variáveis explicativas, caso contrário, a equação 3.2 é diferente de zero. Essa hipótese essencial implica que o método dos mínimos quadrados ordinários (MQO) aplicado na estimação dos parâmetros é não viesado, problema que pode surgir devido a omissão de alguma variável explicativa (WOOLDRIDGE, 2010).

Para a obtenção das estimativas, outro pressuposto é a existência de linearidade no modelo de regressão. O termo linearidade se descreve pela relação linear entre os parâmetros, isto é, os parâmetros são elevados apenas à primeira potência, podendo ou não ser linear nas explicativas (GUJARATI; PORTER, 2011).

Uma outra hipótese que incorre entre as variáveis de um modelo é a colinearidade não perfeita entre as variáveis explicativas. A existência da relação linear entre duas ou

mais variáveis explicativas é denominada multicolinearidade e é um problema que pode prejudicar na estimação dos parâmetros.

$$VIF_j = \frac{1}{1 - R_j^2} \quad (3.3)$$

A multicolinearidade pode ser facilmente identificada via fatores de inflação da variância (VIF), definido pela equação 3.3, onde  $R_j^2$  é o coeficiente de determinação da regressão auxiliar de  $X_j$  sobre as demais explicativas (GUJARATI; PORTER, 2011).

Esses fatores avaliam o aumento da variância de um parâmetro estimado da regressão quando suas explicativas estão correlacionados. Dessa forma, quanto maior o valor de VIF, maior a correlação entre as explicativas. Assim como muitos outros autores, recomenda-se que o VIF deve ser inferior a 10, caso contrário, a multicolinearidade poderá influenciar nos coeficientes da regressão (PETRINI et al., 2012).

Considerando as hipóteses de que os dados foram amostrados de forma aleatória, a esperança do termo de erro condicionada às explicativas é nula, a hipótese da linearidade dos parâmetros e a de colinearidade não perfeita entre as explicativas, os estimadores obtidos pelo MQO são lineares e não enviesados (WOOLDRIDGE, 2010).

Em uma análise de RLM, a condição em que a variância do erro é constante é denominada homocedasticidade. Esse pressuposto garante uma melhor eficiência dos estimadores do MQO, em outras palavras, na presença de erros homocedásticos é possível obter estimativas com menores variâncias do que as obtidas na presença de heterocedasticidade.

A detecção da heterocedasticidade em um modelo pode seguir abordagens distintas. Uma maneira informal de se avaliar a presença de heterocedasticidade seria a comparação, através de um teste de hipótese baseado na variância das estimações, de um modelo com e outro sem a presença de heterocedasticidade.

Além disso, uma técnica vastamente empregada para o tratamento da heterocedasticidade é o uso de erros-padrões robustos. Essa estatística segue a seguinte idéia, caso o quadrado dos resíduos estimados pelos mínimos quadrados for constante para todas as unidades de corte transversal  $i$ , os erros-padrões robustos seriam os mesmos obtidos por MQO. Lembrando que essa estatística segue comportamento assintótico, isto é, é válido apenas para amostras relativamente grandes (WOOLDRIDGE, 2010).

Adicionando a hipótese de erros homocedásticos ao modelo clássico de regressão linear, o teorema de *Gauss-Markov* garante que os estimadores obtidos são BLUE (*Best Linear Unbiased Estimators*), isto é, são os estimadores lineares não viesados de variância mínima (GUJARATI; PORTER, 2011).

A existência de autocorrelação entre os dados analisados, que significa que uma observação está relacionada com as observações antecessoras ou sucessoras dela, assim como a existência de heterocedasticidade, não faz com que os estimadores de MQO deixem de ser lineares e não enviesados, porém, deixam de ser eficientes, em outras palavras, deixam de ter variância mínima (WOOLDRIDGE, 2010).

Quando se procura compreender a relação quantitativa entre duas variáveis, pode-se recorrer a análise de correlação. Larson e Faber definem a correlação como sendo a relação entre dois conjuntos, onde um representa o conjunto de variáveis dependentes e o outro o conjunto de independentes. Quando a relação está próxima de uma reta, isto é, o aumento de uma unidade na variável independente gera um aumento de  $\beta$  unidades na variável dependente, ela é definida como correlação linear (LARSON; FARBER, 2015).

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.4)$$

O coeficiente de correlação, equação 3.4, é um coeficiente que mede a relação entre a variável resposta e a explicativa. Seu valor está contido entre -1 e +1, onde 0 indica uma ausência de correlação, em outras palavras a correlação não é significativa, e 1 uma correlação total, assumindo valores negativos em casos de correlação negativa e valores positivos em casos de correlação positiva. Em síntese, esse coeficiente mede o grau de associação da relação linear entre duas variáveis (GUJARATI; PORTER, 2011).

Para a inferência estatística, é necessário conhecer a distribuição do termo de erro. Nos modelos lineares, outra suposição é a de que o termo de erro, condicionado às explicativas, seja normalmente distribuído (WOOLDRIDGE, 2010). Duas ferramentas gráficas que podem ser utilizadas para identificar a normalidade é o histograma de frequência dos resíduos ou o gráfico de probabilidade normal dos resíduos.

Além disso, também pode-se aplicar testes de hipóteses, baseados na detecção de desvios gaussianos na forma de assimetria e excesso de curtose nos termos de erro da regressão, tais como o teste de *Jarque-Bera* (ALEJO et al., 2015).

Além da presença de padrões, a análise de resíduos exige outro diagnóstico que é em relação aos *outliers*. Segundo Pires e Branco, *outliers* são definidos como observações com valores extremos, que se encontram afastados da maioria dos dados (PIRES; BRANCO, 2007).

A natureza dos *outliers* é um fator importante a ser considerado. Nascimento et al. afirmam que a natureza dos *outliers* tem capacidade de influenciar na análise dos dados de medição. Assim, a compreensão da série em análise é essencial, no sentido de que a constatação de um *outlier* pode ser sinal de algum fator afetando a série. (NASCIMENTO et al., 2012).

### 3.4 Modelos de Regressão Linear Múltipla com dados em painel

A utilização de um banco de dados que consiste das concessionárias brasileiras avaliadas ao longo dos anos, diferente da análise de apenas uma concessionária pode



apresentar certas vantagens, como o controle de características não observadas no tempo, que em certos casos pode estar correlacionado com as variáveis explicativas.

Em uma análise de RLM, onde o principal objetivo é determinar estimações para os parâmetros de inclinação e o intercepto, é essencial entender as relações entre as variáveis explicativas e a explicada para que seja possível interpretar corretamente as estimações obtidas.

$$y_{it} = \beta_1 \cdot x_{1it} + \beta_2 \cdot x_{2it} + \dots + \beta_k \cdot x_{kit} + \alpha_i + u_{it} \quad (3.5)$$

Em um modelo de regressão aplicado à uma estrutura de dados em painel, descrito pela equação 3.5,  $k$  representa o número de parâmetros à serem estimados,  $u_{it}$  é o *erro idiossincrático*,  $\alpha_i$  é o efeito não observado e os  $\beta_1, \beta_2, \dots, \beta_k$  são os parâmetros à serem estimados (WOOLDRIDGE, 2010).

Na estimação de modelos de efeitos não observados com dados em painel, a análise do efeito não observado  $\alpha_i$  é imprescindível na escolha entre a estimação por efeitos fixos ou efeitos aleatórios. O efeito não observado  $\alpha_i$ , também denominado efeito fixo ou heterogeneidade do indivíduo uma vez que é fixo ao longo do tempo, capta todos os fatores não observados que afetam a variável dependente.

A partir da **transformação de efeitos fixos**, aplicada à equação 3.5, o efeito não observado  $\alpha_i$  é eliminado, apontando o método dos mínimos quadrados ordinários agrupado para estimação da equação transformada (WOOLDRIDGE, 2010).

$$\hat{y}_{it} = \beta_1 \cdot \hat{x}_{1it} + \beta_2 \cdot \hat{x}_{2it} + \dots + \beta_k \cdot \hat{x}_{kit} + \hat{u}_{it} \quad (3.6)$$

Na consideração desse modelo, uma das hipóteses é a de que existe uma correlação arbitrária entre o efeito não-observado e as variáveis explicativas do modelo. Dessa forma, após a transformação, qualquer explicativa que seja constante ao decorrer do tempo é eliminada do modelo, conforme a equação 3.6.

A estimação por efeitos fixos é vastamente empregada em painéis desbalanceados, porém é importante assumir que os períodos de tempo faltando não são sistematicamente relacionados com o erro idiossincrático.

Por outro lado, caso o entendimento seja de que o efeito não observado não se correlaciona com as variáveis explicativas ao longo dos períodos de tempo, isto é,  $Corr(x_{itj}, \alpha_i) = 0$ , as estimativas por efeitos fixos serão ineficientes, logo, as estimativas por efeitos aleatórios é mais adequada (WOOLDRIDGE, 2010).

Dessa forma, tal condição é traduzida pela hipótese para o modelo de efeitos aleatórios de que o valor esperado do efeito não observado  $\alpha_i$ , condicionado às explicativas é constante, isto é,  $E(\alpha_i, x_i) = \beta_0$ . Essa hipótese é a principal diferença entre os dois estimadores.

$$y_{it} = (\alpha_i + \beta_0) + X'_{it} \cdot \beta + u_{it} \quad (3.7)$$

$$y_{it} = \beta_0 + X'_{it} \cdot \beta + (\alpha_i + u_{it}) \quad (3.8)$$

Na estimação por efeitos fixos, o efeito não observado é componente do intercepto, representado pela equação 3.7, enquanto que para o modelo de efeitos aleatórios, representado pela equação 3.8, o efeito não observado compõe o que se define como termo de erro composto ( $v_{it} = \alpha_i + u_{it}$ ) (PARK, 2011).

Uma vez que o efeito não observado está incluído no termo de erro composto  $v_{it}$ , a correlação serial ao longo do tempo pode ser abordada pela estimação através de mínimos quadrados generalizados (WOOLDRIDGE, 2010).

Para os efeitos fixos, o intercepto  $\alpha_i + \beta_0$  difere entre os indivíduos, uma vez que cada unidade  $i$  de corte transversal possui suas características específicas. Já nos efeitos aleatórios, é assumido que o intercepto é extraído aleatoriamente de uma população relativamente maior e com um valor médio constante  $\beta_0$ .

Na escolha entre os modelos, a literatura apresenta alguns testes estatísticos para verificar qual forma de estimação é mais adequada. Na escolha entre o modelo de efeitos fixos e aleatórios pode-se utilizar o teste de Hausman, cuja hipótese nula é de que os estimadores de efeito fixo e os de efeito aleatório não diferem substancialmente, isto é,

$$H_0 : \beta_{FE} = \beta_{RE}, \quad (3.9)$$

onde a estatística de teste utilizada segue distribuição  $\chi^2$  com  $k$  graus de liberdade, sendo  $k$  o número de variáveis explicativas do modelo (GUJARATI; PORTER, 2011).

De forma geral, os estimadores de efeito fixo são mais consistentes que os de efeito aleatório, porém são menos eficientes (maior variância dos estimadores) assim, dada a hipótese nula 3.9, caso ela seja rejeitada os estimadores de efeito fixo é mais adequada, porém, caso 3.9 não seja rejeitada, ambos os estimadores podem ser utilizados, entretanto o de efeito aleatório é mais adequado uma vez que são mais eficientes.

Além do teste de Hausman, uma abordagem informal para escolha do modelo de efeitos fixos ou aleatórios, seria analisar o número de unidades de corte transversal  $N$  e o número de períodos da série temporal  $T$ . Caso  $T \gg N$ , os resultados irão diferir muito pouco, logo, a escolha é baseada em conveniência computacional, entretanto caso  $N \gg T$ , definido como painel curto, e as unidades de corte transversal são extrações aleatórias de uma população muito maior, os estimadores de efeitos aleatórios serão mais eficientes (GUJARATI; PORTER, 2011).

Por fim, outra consideração importante à respeito dos dados em painel é a normalidade dos termos de erro. Uma vez que em modelos para dados em painel, em sua composição, existe o efeito não observado  $\alpha_i$  e o erro idiossincrático  $\mu_{it}$ , na avaliação de sua distribuição, existe uma preocupação em identificar qual destes componentes de erro apresenta fontes de não normalidades. Dessa forma, os testes para normalidade foram

baseados nos trabalhos de Alejo et al., que apresentaram um teste que permite a avaliação separada ou conjunta dos termos de erro ([ALEJO et al., 2015](#)).



## 4 Metodologia

O capítulo à seguir começa apresentando os testes aplicados para avaliar os níveis de significância estatísticos do modelo e dos regressores. Em seguida, apresenta uma discussão à respeito dos critérios de desempenho de modelos de RLM e finaliza definindo as equações para os modelos propostos neste trabalho.

O banco de dados utilizados neste trabalho consistiu nos valores para diferentes variáveis do setor da distribuição apurados pelas concessionárias e apresentadas pela ANEEL assim como os resultados para a pesquisa IASC. Todas as variáveis utilizadas neste trabalho foram retidas do site da ANEEL ([ANEEL, 2021a](#)) ([ANEEL, 2021c](#)) ([ANEEL, 2021b](#)).

Esses dados foram armazenados com o auxílio do software Excel, e organizados por concessionária e por ano, correspondendo às unidades de corte transversal e os períodos das séries temporais, respectivamente.

### 4.1 Significância dos parâmetros de regressão e do modelo

Para se avaliar os ajustes dos modelos de RLM, certas estatísticas podem ser utilizadas, dentre elas pode-se citar o coeficiente de determinação ou coeficiente de ajuste  $R^2$ , que mede o percentual da variação total de uma variável dependente explicada pelo modelo com os regressores propostos ([GUJARATI; PORTER, 2011](#)).

O coeficiente de determinação pode ser definido através da soma quadrática total (SQT), da regressão (SQE) e dos resíduos (SQR), de acordo com a seguinte equação:

$$R^2 = \frac{SQ_E}{SQ_T} = 1 - \frac{SQ_R}{SQ_T}, \quad (4.1)$$

onde

$$SQ_T = SQ_E + SQ_R, \quad (4.2)$$

$$SQ_E = \sum_{i=1}^p (\hat{y}_i - \bar{y}_i)^2, \quad (4.3)$$

$$SQ_R = \sum_{i=1}^p (\hat{y}_i - y_i)^2, \quad (4.4)$$

e

$$SQ_T = \sum_{i=1}^p (y_i - \bar{y}_i)^2. \quad (4.5)$$

Por exemplo, se o coeficiente de determinação for igual à 0,8, significa que 80% da variação na variável dependente pode ser explicada pelo modelo que o relaciona com as explicativas. O restante dos  $(1 - R^2)$  não são explicados e podem ser consequências de outros fatores, como variáveis não consideradas (LARSON; FARBER, 2015).

Alguns autores não recomendam analisar apenas o coeficiente  $R^2$ , uma vez que altos valores de  $R^2$  não confirmam, por si só, que o modelo está bem ajustado, podendo acarretar em interpretações errôneas.

Conforme Montgomery e Runger, a adição de variáveis independentes e sem correlação podem aumentar o valor de  $R^2$ , inflando de forma artificial o modelo. Dessa forma, o  $R^2$  é um bom indicador para a qualidade do ajuste, isto é, o grau que as variáveis independentes explicam a dependente, entretanto, não é um bom comparador de modelos (MONTGOMERY; RUNGER, 2009).

Além disso, o coeficiente de determinação não indica se o modelo ou os regressores individualmente são estatisticamente significativos, assim são necessários certos testes de hipóteses.

Um teste de hipótese basicamente é um processo em que os resultados baseados em uma amostra são aplicados na verificação de veracidade de uma determinada hipótese. Sua idéia consiste em um estimador, também conhecido como estatística de teste ou teste estatístico, e a distribuição amostral desse estimador sob uma hipótese nula, sendo a decisão de aceitar ou rejeitar tal hipótese baseado no valor desse estimador (GUJARATI; PORTER, 2011).

O teste de hipótese para avaliação da significância estatística do modelo pode indicar se a equação que modela a regressão é significativa, em outras palavras, se a relação funcional entre a variável dependente e suas explicativas é expressiva. O teste pode-se basear em uma comparação entre os resultados com o modelo com as restrições lineares (restritos) e sem as restrições lineares (irrestrito).

Na realização dos testes estatísticos para avaliação da significância para os modelos com dados em painel, segue-se abordagens distintas de acordo com a técnica de estimação. Para a estimação por efeitos fixos, uma alternativa para o teste de significância do modelo é baseado na estatística de teste tradicional F, enquanto que para o modelo de efeitos aleatórios o teste é baseado na estatística Wald com distribuição  $\chi^2$ .

Para o modelo de efeitos fixos, a estimação pode ser feita por MQO enquanto que para o modelo de efeitos aleatórios pode ser feita por MQG, uma vez que os termos de erros de uma mesma unidade de corte transversal para diferentes períodos de tempo podem estar correlacionadas, tal correlação pode fazer que com as estimações por MQO sejam ineficientes (GUJARATI; PORTER, 2011).

As hipóteses para os testes são realizadas sobre um ou mais parâmetros do modelo de regressão com  $k$  variáveis. A hipótese nula para o teste de significância é de  $\beta_j = 0$  para todos os  $j = 1, 2, 3, \dots, k$ . Desta forma, a rejeição da hipótese nula implica que ao menos

uma das explicativas contribuem significativamente para o modelo (MONTGOMERY; RUNGER, 2009).

Assim como teste F, o teste Wald também é baseado em restrições lineares impostas sobre os parâmetros do modelo. Definindo  $R_R^2$  como coeficiente de determinação do modelo sujeito às restrições lineares e  $SQR_R$  a soma quadrática dos resíduos sujeito às restrições lineares, as estatísticas para os testes de significância para o modelo são baseados nas seguintes equações (GUJARATI; PORTER, 2011):

$$F = \frac{(R^2 - R_R^2)/m}{(1 - R^2)/(n - k)} \quad (4.6)$$

e

$$\text{Wald } \chi^2 = \frac{(n - k) \cdot (SQR - SQR_R)}{SQR_R}. \quad (4.7)$$

Caso o  $p$ -valor para as estatísticas apresentadas em 4.6 e 4.7 forem inferiores à um determinado nível de significância, usualmente 5%, rejeita-se a hipótese nula, concluindo que o modelo com os regressores propostos é, conjuntamente, significativo.

Por fim, além da avaliação da significância estatística do modelo com a combinação de regressores propostos, é necessário também uma avaliação da significância estatística dos regressores individualmente. Para tal avaliação, a abordagem irá seguir os testes baseados nas estatísticas  $t$  e  $z$  (GUJARATI; PORTER, 2011).

Sob a hipótese de normalidade, a estatística de teste  $t$  é calculada por

$$t = \frac{\hat{\beta}_i - \beta_i}{\sigma_{\hat{\beta}_i}}, \quad (4.8)$$

onde  $\hat{\beta}_i$  é o parâmetro estimado para a  $i$ -ésima explicativa do modelo e  $\sigma_{\hat{\beta}_i}$  é o desvio padrão de tal estimação.

Já a estatística de teste  $z$ , também utilizada para realizar inferências sobre uma população, é calculada pela equação

$$z = \frac{(\bar{x} - \mu_x)}{\frac{\sigma_x}{\sqrt{n}}}, \quad (4.9)$$

onde  $\mu_x$  indica o valor de  $x$  a ser testado como hipótese,  $\sigma_x$  o desvio padrão da variável  $x$  e  $n$  é o número de observações. Diferente da estatística  $t$ , que segue a distribuição  $t$  de Student, a estatística  $z$  segue uma distribuição normal (GUJARATI; PORTER, 2011).

No software utilizado, a execução do comando para estimação dos parâmetros computa automaticamente tais testes, para o modelo de efeitos fixos computou-se o teste  $t$ , já para o modelo de efeitos aleatórios computou-se o teste  $z$  na avaliação da significância estatística individual das explicativas (STATA CORP, 2021).

## 4.2 Critérios para seleção de modelos

Além dos coeficientes  $R^2$  e  $R_{ajustado}^2$ , existem outros coeficientes utilizados para selecionar o modelo estimado que, condicionado às variáveis explicativas, melhor descreve o modelo real, que em teoria é desconhecido. Dentre os diversos métodos avaliados na literatura que, em geral, buscam a minimização a soma dos quadrados dos resíduos (SQR) ou aumentar o  $R^2$ , pode-se citar o *Akaike Information Criterion* e o *Bayesian Information Criterion*, cuja idéia foi utilizada na avaliação dos modelos de regressão.

Akaike (1973) demonstrou que o viés é dado assintoticamente pelo número de parâmetros à serem estimados no modelo e definiu seu critério de informação de Akaike (AIC), cuja essência consiste em impor uma medida corretiva pelo acréscimo de regressores em um determinado modelo. Seu coeficiente é calculado por (AKAIKE, 1973)

$$AIC = 2k - 2.ln(\hat{L}), \quad (4.10)$$

onde  $\hat{L}$  representa o valor máximo da função de verossimilhança e  $k$  o número de parâmetros do modelo.

Em modelos de regressão, a equação 4.10 pode ser aproximada por

$$ln(AIC) = \frac{2k}{n} + ln\left(\frac{SQ_R}{n}\right), \quad (4.11)$$

sendo  $n$  o número de observações do modelo de regressão (AKAIKE, 1973).

Como observado na equação 4.11, o AIC estabelece uma medida corretiva mais rígida que o  $R^2$  em função do acréscimo de regressores. Na comparação de dois ou mais modelos, o modelo com o menor valor de AIC é preferível. Além disso, o AIC possibilita previsões do desempenho de um modelo, mesmo em problemas fora da amostra (GUJARATI; PORTER, 2011).

Similar ao AIC, porém impondo medidas corretivas mais duras, Schwarz definiu o critério de informação de Schwarz (SIC), ou critério de informação Bayesiano (BIC). Para modelos de regressão, o SIC é calculado por

$$SIC = k.ln(n) - 2.ln(\hat{L}), \quad (4.12)$$

onde os parâmetros da equação são os mesmos apresentados para o AIC.

$$ln(SIC) = \frac{k}{n}.ln(n) + ln\left(\frac{SQ_R}{n}\right) \quad (4.13)$$

Para modelos de regressão, a equação 4.12 pode ser aproximada pela equação 4.13 (SCHWARZ, 1978).

No contexto da regressão, o AIC é assintoticamente ideal para selecionar o modelo com o menor erro quadrático médio, sob a suposição de que o modelo real não esteja no conjunto de candidatos. Já o SIC não é assintoticamente ideal sob o pressuposto (YANG, 2005).



Em síntese, os critérios AIC e SIC quantificam a idéia de que, caso dois modelos com números de parâmetros distintos possuem a mesma capacidade preditiva, sempre será preferível o modelo com o menor número de parâmetros. Baseado nisso e no fato de que certos indicadores apurados pelas concessionárias apresentam diversas estratificações, resumiu-se tais estratificações em apenas uma variável, com o intuito de diminuir o número de parâmetros associados ao modelo, procurando manter a mesma capacidade dos regressores preverem a variável dependente.

### 4.3 Modelos de Regressão Linear Múltipla propostos

Modelos de RLM são vastamente utilizados como modelos empíricos, onde a equação que representa o fenômeno que relaciona as variáveis de entrada e de saída é desconhecida e, através de testes estatísticos de hipóteses, obtém-se uma equação que representa o modelo desconhecido dado um nível de significância.

Inicialmente, antes de se propor as equações para os modelos de RLM, realizou-se o levantamento dos dados. O banco de dados utilizado neste trabalho consistiu nos indicadores da distribuição, dados de consumo, número de consumidores e o IASC, para diferentes concessionárias distribuidoras de energia elétrica atuantes no cenário brasileiro.

Durante o levantamento, separou-se os dados por conjunto elétrico ou por concessionária, entretanto, como os indicadores apurados por conjunto elétrico se resumiam apenas aos indicadores individuais e coletivos (e suas extratificações) de continuidade e os tempos de atendimento às ocorrências emergenciais, optou-se pelo desenvolvimento de modelos com dados apurados apenas por concessionária.

Desejando um painel equilibrado, descartou-se todas as concessionárias que estavam inadimplentes em relação a apuração dos indicadores, além disso, à partir de uma inspeção visual dos dados, observou-se que determinadas concessionárias apresentavam excelentes indicadores, como por exemplo vários zeros voltados às compensações financeiras, desta forma, tais concessárias foram descartadas por representarem outliers. A amostra total foi constituída por 24 concessionárias analisadas em um período de 8 anos, de 2011 a 2019 com a exceção de 2010 cujos dados não foram divulgados pela ANEEL pois a pesquisa realizada pela empresa que venceu a licitação em 2011 não foi validada, totalizando assim 192 observações para cada variável do modelo.

Como regressores dos modelos de regressão propostos, procurou-se avaliar: os indicadores voltados as três dimensões da qualidade de energia; as compensações voltadas à transgressão dos limite estabelecidos; dados de consumo de energia elétrica e receita apuradas pelas concessionárias além do número de consumidores de determinada concessionária.

Como variável à ser explicada utilizou-se o IASC, indicador mais próximo da perspectiva do consumidor. Em específico, utilizou-se o *escore* confiabilidade, da dimensão

qualidade percebida do IASC, conforme apresentado pela figura 2.3.

No desenvolvimento dos modelos, baseado na ideia essencial dos critérios de seleção AIC e SIC de que, entre dois modelos com a mesma capacidade preditiva, o modelo com o menor de número de regressores é preferível, utilizou-se duas variáveis que resumissem os indicadores associados aos tempos médios de atendimento às ocorrências emergenciais (TMA) e os associados à duração relativa da tensão nas faixas precária e crítica (DRPCE), de acordo com as equações

$$TMA = TMP + TMD + TME \quad (4.14)$$

e

$$DRPCE = DRPE + DRCE = \frac{(\sum_{DRP_i} + \sum_{DRC_i})}{N_L}, \quad (4.15)$$

onde  $TMA$ , da equação 4.14, representa a soma dos tempos médios de preparação, de deslocamento e de execução, apresentados pelas equações 2.5, 2.6 e 2.7, e  $DRPCE$  a duração relativa de transgressão de tensão precária e crítica de um determinado conjunto elétrico, isto é, a duração relativa da tensão fora da faixa adequada.

Na seleção das variáveis, avaliando a correlação teórica entre os indicadores dispostos, observou-se a inviabilidade de utilizar mais de um indicador para cada dimensão da qualidade previamente apresentada, desta forma, para cada um dos modelos utilizou-se um indicador de continuidade, um de conformidade, um de tratamento de reclamações, um voltado às ocorrências emergenciais e um voltado ao número de consumidores ou consumo, uma vez que também estão teoricamente correlacionadas, isto é, existe uma relação entre o número de consumidores e consumo.

Por mais que os indicadores associados às ocorrências emergenciais constituam indicadores de continuidade e sua correlação com os indicadores DEC e FEC não sejam nulas, optou-se por utilizá-las conjuntamente no modelo, uma vez que as estratificações associados às situações de emergência e dia crítico compõem apenas os indicadores DECTOT e FECTOT, como citado no capítulo 2.

Assim, o primeiro modelo proposto foi o seguinte:

$$IASC_{it+1} = \beta_0 + \beta_1.FEC_{it} + \beta_2.DRPCE_{it} + \beta_3.DERK_{it} + \beta_4.TMAH_{it} + \beta_5.NCONSM_{it}, \quad (4.16)$$

onde  $NCONSM_{it}$  é o número de unidades consumidoras da concessionária  $i$  no ano  $t$ , expresso na escala de milhões,  $DERK$  é a duração equivalente de reclamação expressa na escala mil e  $TMAH$  é o tempo médio de atendimento às ocorrências emergenciais, expresso em horas.

Em seguida, com a experiência do modelo proposto em 4.16, foi proposto a avaliação das compensações financeiras ( $COMP_{PMit}$ ), expresso em milhões de reais, voltadas à conformidade do nível de tensão ao invés do indicador DRPCE, descrito pela equação

$$\begin{aligned}
IASC_{it+1} = & \beta_0 + \beta_1.FEC_{it} + \beta_2.COMP_{PMit} + \\
& \beta_3.DERK_{it} + \beta_4.TMAH_{it} + \beta_5.NCONSM_{it}.
\end{aligned}
\tag{4.17}$$

Por fim, com o intuito de avaliar a relação entre o consumo e a satisfação do consumidor, foi necessário um outro modelo, sem considerar o número de consumidores, o que acarretaria em problemas na estimação dos parâmetros, desta forma, foi proposto a seguinte equação:

$$\begin{aligned}
IASC_{it+1} = & \beta_0 + \beta_1.FEC_{it} + \beta_2.COMP_{PMit} + \\
& \beta_3.DERK_{it} + \beta_4.TMAH_{it} + \beta_5.CONSUMOM_{it},
\end{aligned}
\tag{4.18}$$

onde  $CONSUMOM_{it}$  é o consumo de energia elétrica, expressa em TWh, da concessionária  $i$  no ano  $t$ .

Além disso, dado a forte correlação entre os indicadores DEC e FEC não foi possível avaliar os dois indicadores no mesmo modelo, uma vez que sem FEC não há DEC. Avaliando o coeficiente de correlação de Pearson, observou-se que existe uma correlação muito forte,  $\rho = -0,8900$ , entre DEC e FEC, em acordo com as conclusões de Mondani et al., que afirmaram existir uma forte correlação entre DEC e FEC (MONDANI et al., 2017), sugerindo assim que ambas sejam avaliadas em modelos diferentes. Uma vez que diversos trabalhos da literatura especializada trazem comparações entre o DEC e FEC, procurou-se avaliá-los em modelos separados para realizar tal comparação. Dessa forma, propõe-se também o modelo descrito pela equação 4.19.

$$\begin{aligned}
IASC_{it+1} = & \beta_0 + \beta_1.DEC_{it} + \beta_2.COMP_{PMit} + \\
& \beta_3.DERK_{it} + \beta_4.TMAH_{it} + \beta_5.CONSUMOM_{it}.
\end{aligned}
\tag{4.19}$$

Na implementação dos modelos de regressão no STATA escreveu-se um código, apresentado no apêndice A, para a estimação dos parâmetros, avaliação da heterogeneidade do indivíduo para dados em painel e validação das hipóteses clássicas de modelos de RLM.

A corporação STATA apresenta uma série de bibliotecas para auxiliar na implementação dos modelos de RLM e os testes estatísticos, além de manuais de referência. Neste trabalho, utilizou-se o *Longitudinal/panel data reference manual* como principal referência na implementação dos testes (STATACORP, 2021).



## 5 Resultados e discussões

Atualmente, diversos softwares estatísticos estão disponíveis para se trabalhar com modelos de RLM. Dentre as opções disponíveis, utilizou-se o software STATA visto que ele possui diversas ferramentas de regressão, além de vastas bibliotecas que contêm testes estatísticos importantes para a validação de um modelo de RLM.

Em testes preliminares, foi proposto um modelo de regressão para avaliação dos dados em séries temporais, iniciando o estudo pela concessionária Companhia Paranaense de Energia (COPEL), apurados anualmente no período de 2010 à 2019, e repetindo o processo para outras concessionárias.

Com os dados em séries temporais, não foi possível propor um modelo estatisticamente significativo com os regressores disponíveis, dado um nível de significância de 5%, levantando a suspeita de que o tamanho da amostra era insuficiente. Assim, tais testes foram agregados apenas como experiência para os modelos de dados em painel.

Em seguida, para o aumento do tamanho da amostra, uma opção foi mudar o foco para dados em painel, que diferente das séries temporais que captam apenas variações no tempo, a estrutura de dados em painel capta variações no tempo e entre os indivíduos, que para os modelos propostos são os anos e as concessionárias.

Na avaliação das possíveis combinações das variáveis para explicar o IASC, não foi possível a computação dos critérios AIC e SIC, uma vez que após os testes de especificação observou-se que os modelos com efeitos aleatórios seriam os mais adequados e a ferramenta computacional utilizada para a estimação por efeitos aleatórios não computa a função de verossimilhança, entretanto, aplicou-se a idéia essencial de tais critérios nos modelos.

Desta forma, com a experiência dos modelos para séries temporais e aplicando a idéia dos critérios AIC e SIC, de que parâmetros não devem ser adicionados ao modelo de forma inconsequente, a busca pelo modelo que realize as melhores estimações para os dados em painel foi baseada na avaliação de variáveis das diferentes dimensões do setor de distribuição, nos testes de significância dos regressores e do modelo, além dos testes de hipótese para validação das hipóteses clássicas de um modelo de RLM.

### 5.1 Modelo 1 - DRPCE

Em um primeiro teste realizado aos modelos para dados em painel, realizou-se a RLM para a equação 4.16. Avaliando a hipótese nula do *teste de Hausman*, observou-se que a estatística de teste  $\chi^2$  foi de 54,46 com  $Prob > \chi^2 = 0,0000 < 0,01$ , isto é, com 1% de significância pode-se rejeitar a hipótese nula de que os parâmetros estimados para os modelos de efeitos fixos e efeitos aleatórios não diferem substancialmente, logo, o modelo de efeitos fixos é o mais adequado, devido a sua maior consistência.

Para o modelo de efeitos fixos, o teste F de significância para o modelo indica que  $Prob > F = 0,0034$ , isto é, dado um nível de significância de 1%, o modelo com os regressores propostos em 4.16 é conjuntamente significativo, entretanto, os valores para os coeficientes de determinação foram extremamente baixos. O  $R^2_{WITHIN}$ , que indica o quanto o modelo se ajusta dentro das unidades, foi igual à 0,1019, enquanto o  $R^2_{BETWEEN}$ , que indica o ajuste entre as unidades, foi igual à 0,1862. Já o coeficiente de determinação geral  $R^2_{OVERALL}$  foi igual à 0,0623.

Por mais que o teste de significância indique que o modelo com os regressores proposto é coletivamente significativo, não foi possível afirmar o mesmo para as explicativas individualmente.

Tabela 5.1 – Parâmetros estimados e testes de significância para o modelo 4.16.

Explicativa	$\beta$	Desvio padrão	t	P >  t
FEC	-0,4849162	0,2156213	-2,25	0,026 **
DRPCE	0,7333461	0,3646142	2,01	0,046 **
DERK	2,7284	1,439006	1,90	0,060 *
TMAH	-0,055244	4,664325	-0,12	0,906
NCONSM	7,180449	4,065995	1,77	0,079 *
Intercepto	57,63748	7,128873	8,09	0,000 ***

\* p < 0,10; \*\* p < 0,05; \*\*\* p < 0,01

De acordo com as estatísticas dos testes  $t$ , apresentadas na tabela 5.1, observou-se que apenas duas explicativas possuíam significância estatística com o modelo, dado um nível de significância de 5%, levantando a suspeita de algum problema entre as variáveis. Uma vez que a correlação pode afetar a eficiência dos estimadores e, conseqüentemente, seus desvios padrões presumiu-se que talvez isto pudesse estar afetando o modelo.

Na avaliação da correlação entre as variáveis explicativas do modelo 4.16, dois valores se destacaram:  $Corr(DERK, DRPCE) = 0,3470$  e  $Corr(FEC, TMAH) = -0,3934$ .

Uma das preocupações da ANEEL é na fiscalização dos indicadores de transgressão de tensão, uma vez que são baseadas em medições realizadas nas unidades consumidoras e enviados pela própria distribuidora, sendo passíveis de erros e manipulações, conforme reforçado por Paiva et al., que apresentaram uma estimação de estados para abordar tais erros e manipulações nos indicadores DRP e DRC (PAIVA et al., 2021). Dessa forma, uma primeira tentativa para verificar o problema foi a regressão da equação 4.16 sem o indicador DRPCE.

Tabela 5.2 – Parâmetros estimados e testes de significância para o modelo 4.16 sem DRPCE.

Explicativa	$\beta$	Desvio padrão	t	$P >  t $
FEC	-0,509796	0,2172556	-2,35	0,020 **
DERK	2,652004	1,451799	1,83	0,070 *
TMAH	-0,0943962	0,470333	-0,20	0,841
NCONSM	7,376099	4,102398	1,80	0,074 *
Intercepto	58,90436	7,166617	8,22	0,000 ***

\* p < 0,10; \*\* p < 0,05; \*\*\* p < 0,01

Novamente, para a explicativa TMAH foi possível concluir que, na estimação por efeitos fixos, a mesma não possui significância estatística com o modelo para avaliação do IASC, conforme tabela 5.2. Assim, realizou-se a regressão apenas com as explicativas *FEC*, *DERK* e *NCONSM*, conforme a seguinte equação:

$$IASC_{it+1} = \beta_0 + \beta_1.FEC_{it} + \beta_2.DERK_{it} + \beta_3.NCONSM_{it}. \quad (5.1)$$

Para a estatística  $\chi^2$  do teste de *Hausman*, aplicado ao modelo da equação 5.1, foi obtido o valor de 6,68, o que indica que  $Prob > \chi^2 = 0,0828$ , logo, dado um nível de significância de 5%, o modelo de efeitos aleatórios é o mais adequado para este caso.

Tabela 5.3 – Resumo dos testes aplicados no modelo representado pela equação 5.1.

Teste	Hipótese nula	Estatística de teste
Autocorrelação (Wooldridge)	Sem correlação serial de 1ª ordem	$Prob > F = 0,1214$
Homocedasticidade (Poi & Wiggins)	Erros homocedásticos	$Prob > \chi^2 = 0,8487$
Normalidade (Skewness & Kurtosis)	Normalidade em $\alpha$	$Prob > \chi^2 = 0,9117$
	Normalidade em $\mu$	$Prob > \chi^2 = 0,2295$

Para a validação do modelo, foi necessário a aplicação de certos testes estatísticos para verificação das hipóteses do modelo de RLM, apresentadas na seção anterior. O resumo dos testes aplicados estão apresentados na tabela 5.3.

De acordo com o teste de Wooldridge para autocorrelação, a estatística de teste indicou a não rejeição da hipótese nula de correlação serial de 1ª ordem entre as explicativas, isto é, as variáveis explicativas não estão correlacionadas em função do tempo.

Fora o teste de homocedasticidade de Poi & Wiggins, que indica a não rejeição da hipótese de erros homocedásticos, para avaliação do problema da heterocedasticidade que

ocorre quando a variância não é constante para algum dos indivíduos, uma estratégia informal seria analisar o efeito da implementação de erros padrões robustos.

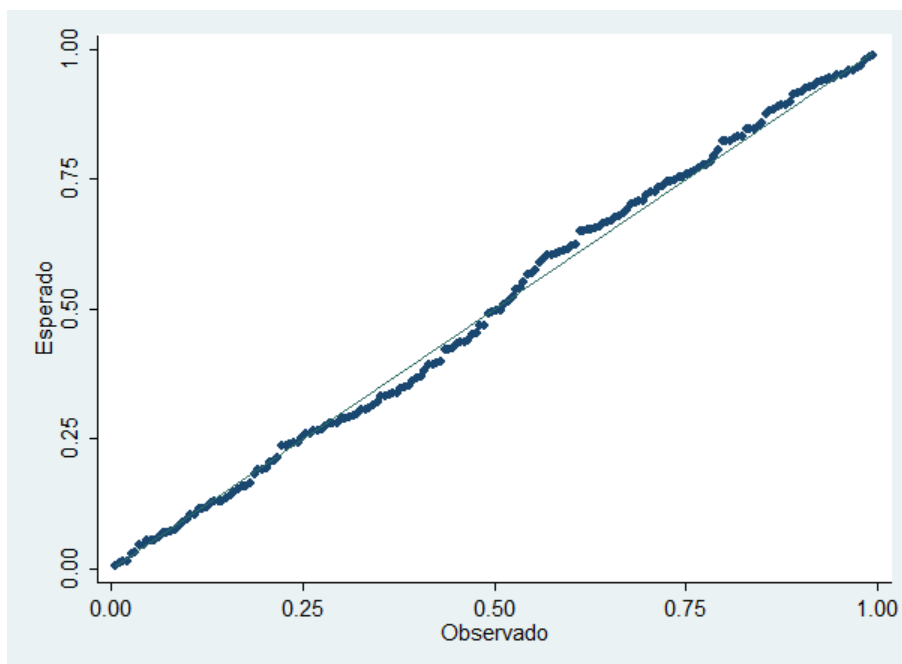
Tabela 5.4 – Comparação erros padrões e erros padrões robustos para o modelo da equação 5.1.

<b>Explicativa</b>	<b>Erro padrão</b>	<b>Erro padrão robusto</b>
FEC	0,1612615	0,2134215
DERK	1,399035	0,8581723
NCONSM	0,4543228	0,5569603
Intercepto	1,77611	2,129913

Rodando o mesmo modelo de efeitos aleatórios com os erros padrões robustos, observou-se que os erros padrões não aumentaram significativamente e, uma vez que a heterocedasticidade é um problema que afeta o desvio padrão das estimações, pode-se verificar os resultados para o teste anterior.

As hipóteses de normalidade dos termos de erro  $\mu_{it}$  e  $\alpha_i$  da equação 3.8, baseados nas avaliações de assimetria e curtose, também não foram rejeitadas, indicando que tais termos são normalmente distribuídos.

Figura 5.1 – Gráfico de probabilidade normal para o termo de erro composto referente ao modelo da equação 5.1.

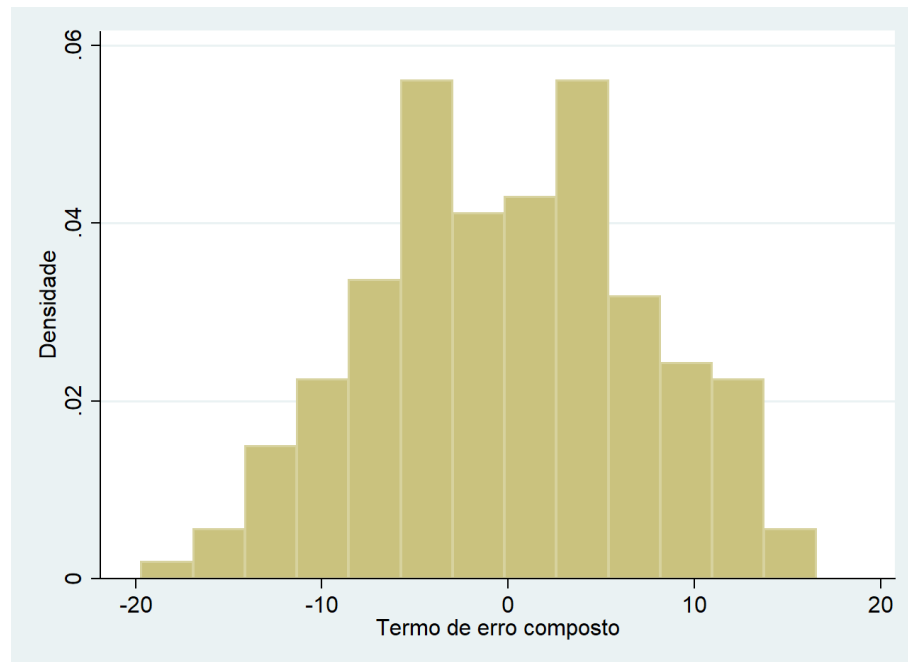


Fonte: Elaborado pelo autor (2021).

Além do teste de hipótese aplicado, verificou-se também a normalidade dos termos de erro através do gráfico de probabilidade normal, representada pelo bom ajuste à uma reta figura 5.1, e o histograma dos resíduos, figura 5.2, cujo formato se aproxima de uma distribuição normal.



Figura 5.2 – Histograma do termo de erro composto referente ao modelo da equação 5.1.



Fonte: Elaborado pelo autor (2021).

Tabela 5.5 – Matriz de coeficientes de correlação para o modelo da equação 4.16.

Explicativa	FEC	DERK	NCONSM
<b>FEC</b>	1,0000		
<b>DERK</b>	-0,0807	1,0000	
<b>NCONSM</b>	0,0487	-0,0345	1,0000

Fonte: Elaborada pelo autor (2021)

Além disso, avaliando a correlação entre as explicativas, apresentados na tabela 5.5, observou que a correlação entre as três explicativas que permaneceram no modelo foi pequena, uma vez que as explicativas *DRPCE* e *TMAH* foram retiradas.

Dessa forma, o modelo final obtido foi:

$$IASC_{it+1} = (75,26914 \pm 2,129913) + (-0,721526 \pm 0,2134215).FEC_{it} + (1,8743 \pm 0,8581723).DERK_{it} + (-1,343285 \pm 0,5569603).NCONSM_{it}, \quad (5.2)$$

com os testes indicando um nível de significância estatística de 5% para as explicativas *FEC*, *DERK* e *NCONSM*.

Com o modelo da equação 5.2, foi possível obter coeficientes de determinação maiores quando comparado ao modelo da equação 4.16. O  $R^2_{WITHIN}$  foi igual à 0,0520,  $R^2_{BETWEEN}$  foi igual à 0,4703 e  $R^2_{OVERALL}$  foi igual à 0,2316, em outras palavras, o modelo explica melhor as variações entre os indivíduos quando comparada às variações dentro de um indivíduo.

Da análise dos parâmetros estimados, pode-se concluir que: o aumento em uma unidade do FEC iria reduzir o IASC em  $(0,721526 \pm 0,2134215)$ ; o aumento em mil horas do DER iria aumentar o IASC em  $(1,8743 \pm 0,8581723)$  e um aumento de um milhão no número de consumidores de uma concessionária iria reduzir o IASC em  $(1,343285 \pm 0,5569603)$ .

O parâmetro estimado para a explicativa *FEC* foi esperado, entretanto, os resultados para *DERK* e *NCONSM* foram interessantes. Uma vez que o indicador *DER* carrega informações da quantidade de reclamações e o prazo médio de solução, em teoria, seria esperada uma relação negativa entre *DERK* e *IASC*, entretanto, ao se avaliar o incremento de apenas uma unidade em *DER*, ao invés de mil unidades, se observa que o *IASC* aumenta em aproximadamente 0,0018743, parcela pequena quando comparada ao parâmetro para a explicativa *FEC*, praticamente duas ordens de grandeza maior.

Além disso, o aumento em mil unidades do *DER* representa um aumento muito grande quando comparado com o que ocorre na prática, uma vez que para amostra em questão a explicativa *DER* possui o desvio padrão igual à 368,2579 horas, logo, por mais que o parâmetro estimado seja positivo em si, seu módulo é relativamente pequeno, indicando uma variação muito menor na satisfação do consumidor, quando comparada com o impacto do *FEC*.

Já em relação ao número de consumidores, foi obtido um resultado que pode ser extremamente prático do ponto de vista da concessionária. A relação negativa entre a satisfação do consumidor indica que, para um menor número de consumidores, existe uma melhor gestão das condições de satisfação, entretanto, reduzir o número de consumidores não se configura uma ação de planejamento, uma vez que representa um retrocesso às concessionárias. Assim, uma vez que a divisão proposta do *IASC* considera somente as concessionárias com mais de 400 mil consumidoras, atribuir um peso para sua classificação pode ser uma estratégia à ser utilizada.

## 5.2 Modelo 2 - Compensações financeiras

Para a segunda proposta de modelo, realizou-se uma RLM baseada na equação 4.17, que avalia as compensações financeiras voltadas à qualidade do produto ao invés de DRPCE, conforme equação 2.20. Diferente do ocorrido nos primeiros testes do modelo 1, o teste de *Hausman* obteve a estatística de teste  $\chi^2$  igual à 10,73 com  $Prob > \chi^2 = 0,0569$ , isto é, com 5% de significância, não é possível rejeitar a hipótese nula, logo, o modelo de efeitos aleatórios é mais adequado por sua maior eficiência.

Para o teste de significância do modelo utilizou-se o teste de *Wald*, obtendo uma estatística  $\chi^2$  igual à 60,24 com  $Prob > \chi^2 = 0,0000$ , que indica um modelo estatisticamente significativo a um nível de 1%. Os coeficientes de determinação melhoraram de forma expressiva, sendo o  $R^2_{WITHIN}$  igual à 0,0347, o  $R^2_{BETWEEN}$  igual à 0,7470 e o  $R^2_{OVERALL}$  igual à 0,3422.

Tabela 5.6 – Resumo dos testes aplicados no modelo representado pela equação 4.17.

Teste	Hipótese nula	Estatística de teste
Autocorrelação (Wooldridge)	Sem correlação serial de 1ª ordem	Prob $>F = 0,1240$
Homocedasticidade (Poi & Wiggins)	Erros homocedásticos	Prob $>\chi^2 = 0,6438$
Normalidade (Skewness & Kurtosis)	Normalidade em $\alpha$	Prob $>\chi^2 = 0,5227$
	Normalidade em $\mu$	Prob $>\chi^2 = 0,5422$

Os testes para verificação das hipóteses dos modelos de RLM estão apresentadas na tabela 5.6. Assim como no modelo proposto por 4.16, os testes de significância indicaram a não rejeição das hipóteses de autocorrelação, homocedasticidade e normalidade dos resíduos, reforçando que tais fatores não prejudicaram na estimação dos parâmetros.

Tabela 5.7 – Matriz de coeficientes de correlação para o modelo da equação 4.17.

Explicativa	FEC	TMAH	$COMP_{PM}$	DERK	NCONSM
<b>FEC</b>	1,0000				
<b>TMAH</b>	-0,4665	1,0000			
$COMP_{PM}$	-0,0657	0,1195	1,0000		
<b>DERK</b>	-0,1601	0,3010	0,0831	1,0000	
<b>NCONSM</b>	0,2153	-0,4517	-0,6030	-0,1117	1,0000

Fonte: Elaborada pelo autor (2021)

Os coeficientes de correlação entre as explicativas do modelo estão apresentadas no quadro 5.7. Por mais que as correlações entre  $FEC$  e  $TMAH$  e entre  $COMP$  e  $NCONSM$  sejam relativamente altas, elas não prejudicaram, ao nível de invalidação, na eficiência das estimações, como aconteceu com  $TMAH$  no modelo da equação 4.16, uma vez que no modelo de efeitos aleatórios todas as variáveis alcançaram significância estatística à um nível de 5%, como pode ser observado na tabela 5.8.

Tabela 5.8 – Parâmetros estimados e testes de significância para o modelo 4.17.

Explicativa	$\beta$	Erro padrão robusto	t	$P >  t $
FEC	-0,7494799	0,1771442	-4,23	0,000 ***
$COMP_{PM}$	1,687464	0,6858733	2,46	0,046 **
DERK	1,82375	0,8920557	2,04	0,041 **
TMAH	-0,8523816	0,1842367	-4,63	0,000 ***
NCONSM	-1,234805	0,3732171	-3,31	0,001 **
Intercepto	78,72889	0,3732171	50,14	0,000 ***

\*  $p < 0,10$ ; \*\*  $p < 0,05$ ; \*\*\*  $p < 0,01$

$$\begin{aligned}
IASC_{it+1} = & (78,72889 \pm 0,3732171) + (-0,7494799 \pm 0,1771442).FEC_{it} \\
& + (1,687464 \pm 0,6858733).COMP_{PMit} + (1,82375 \pm 0,8920557).DERK_{it} \quad (5.3) \\
& + (-0,8523816 \pm 0,1842367).TMAH_{it} + (-1,234805 \pm 0,3732171).NCONSM_{it},
\end{aligned}$$

No modelo da equação 5.3, diferente do ocorrido no modelo proposto em 4.16, a variável associada à conformidade do nível de tensão assume significância estatística com o modelo, que possivelmente pode ser explicado pelo problema na apuração dos indicadores associados à transgressão de tensão DRP e DRC, reforçado por Paiva et al. (PAIVA et al., 2021).

Com a variável  $COMP_P$  foi possível estimar seu parâmetro do modelo, uma vez que é uma variável cuja percepção por parte dos consumidores é maior quando comparada aos indicadores DRP e DRC, uma vez que tais compensações representam descontos na fatura de energia elétrica.

O aumento em um milhão de reais nas compensações financeiras voltadas à qualidade do produto acarretaria em um aumento do IASC em  $(1,687464 \pm 0,6858733)$ . Uma vez que as compensações financeiras estão diretamente relacionadas aos indicadores DIC, FIC, DMIC, DICRI, DRP e DRC, conforme as equações apresentadas no capítulo 2, seria esperado uma relação negativa com o IASC. Entretanto, tais políticas de compensações tendem a beneficiar o consumidor e cercear as concessionárias e, desta forma, uma relação positiva entre  $COMP_P$  e  $IASC$  não incondiz com a teoria.

Além disso, o indicador  $TMAH$  também passa a ter significância estatística para o modelo de efeitos aleatórios, sugerindo que a combinação das variáveis incluindo as compensações voltadas à qualidade do produto descrevem melhor o modelo para avaliação do IASC. O aumento em uma hora no tempo médio de atendimento às ocorrências emergenciais iria reduzir o IASC em  $(-0,8523816 \pm 0,1842367)$ , condizente com o esperado teoricamente.

Por fim, para o último modelo proposto e baseado nas correlações apresentadas para a explicativa associada ao número de consumidores, almejando um modelo com uma maior capacidade preditiva, isto é, um maior coeficiente de determinação, procurou-se avaliar a variável de consumo de energia elétrica ao invés do número de consumidores, uma vez que as duas, em teoria, são altamente correlacionadas.

### 5.3 Modelo 3 - Consumo

Em testes preliminares para a avaliação das variáveis associadas ao consumo e receita de energia elétrica por uma concessionária, foi observado que a variável associada a receita de energia elétrica possuía níveis de significância estatísticos piores quando comparados com a variável associada ao consumo. Dessa forma, uma vez que ambas

também são altamente correlacionadas, avaliou-se nos modelos propostos apenas a variável associada ao consumo de energia elétrica.

Como último modelo a ser proposto, realizou-se a RLM baseado na equação 4.18. Para esse modelo, o teste de *Hausman* apresentou uma estatística  $\chi^2$  igual à 13,89 com  $Prob > \chi^2 = 0,0163$ , ou seja, à um nível de significância de 1%, não é possível rejeitar a hipótese nula, o que indica que o modelo de efeitos aleatórios, para este nível de significância, é o mais adequado.

Para esse último modelo, o nível de significância para o teste de *Hausman* foi de 1%, uma vez que à um nível de 5%, o modelo de efeitos fixos seria o mais adequado, entretanto, para o modelo da equação 4.18, tais regressores não possuem significância estatística através da estimação por efeitos fixos, o que motivou a avaliação de tal teste sob um nível de significância diferente, visto que os parâmetros estimados por efeitos fixos não seriam válidos, além de que teoricamente, a amostra consiste em uma extração aleatória e o número de unidades de cortes transversais é relativamente maior que os períodos das séries temporais, o que reforça a idéia de que, para os modelos propostos, os estimadores de efeitos aleatórios são mais eficientes (GUJARATI; PORTER, 2011).

Para a avaliação da significância do modelo, a estatística  $\chi^2$  do teste de *Wald* foi de 112,04 com  $Prob > \chi^2 = 0,0000$ , indicando que, à um nível de significância de 1%, o modelo com os regressores propostos da equação 4.18 é conjuntamente significativo.

Tabela 5.9 – Resumo dos testes aplicados no modelo representado pela equação 4.18.

Teste	Hipótese nula	Estatística de teste
Autocorrelação (Wooldridge)	Sem correlação serial de 1ª ordem	$Prob > F = 0,1649$
Homocedasticidade (Poi & Wiggins)	Erros homocedásticos	$Prob > \chi^2 = 0,5336$
Normalidade (Skewness & Kurtosis)	Normalidade em $\alpha$	$Prob > \chi^2 = 0,7181$
	Normalidade em $\mu$	$Prob > \chi^2 = 0,5697$

Assim como nos primeiros dois modelos, foi possível verificar as hipóteses de homocedasticidade, autocorrelação e normalidade dos termos de erro. Tais testes estão resumidos na tabela 5.9.

A substituição da explicativa associada ao número de consumidores pela variável associada ao consumo de energia elétrica gerou alguns efeitos no modelo. Primeiro, pode-se observar na matriz de correlação apresentada em 5.10, que as correlações entre *TMAH* e *CONSUMOM* e entre *COMP<sub>PM</sub>* e *CONSUMOM* forem menores quando comparadas com as correlações entre *TMAH* e *NCONSM* e entre *COMP<sub>PM</sub>* e *NCONSM*. Comparando as tabelas 5.8 e 5.11, foi observado que os desvios padrões das explicativas

Tabela 5.10 – Matriz de coeficientes de correlação para o modelo da equação 4.18.

Explicativa	FEC	TMAH	COMP <sub>PM</sub>	DERK	CONSUMOM
<b>FEC</b>	1,0000				
<b>TMAH</b>	-0,4990	1,0000			
<b>COMP<sub>PM</sub></b>	-0,0201	-0,0780	1,0000		
<b>DERK</b>	-0,1854	0,2735	0,1726	1,0000	
<b>CONSUMOM</b>	0,2600	-0,3406	-0,2759	-0,1964	1,0000

Fonte: Elaborada pelo autor (2021)

*FEC*, *COMP<sub>PM</sub>* e *DERK* sofreram uma leve redução e, conseqüentemente, faz com as explicativas sejam estatisticamente significativas à melhores níveis de significância.

Além disso, para os coeficientes de determinação observou-se que  $R_{WITHIN}^2$  foi igual à 0,0545, o  $R_{BETWEEN}^2$  foi igual à 0,7372 e o  $R_{OVERALL}^2$  foi igual à 0,3527, o que representa um aumento percentual de 3,068% em relação ao coeficiente de determinação geral, sendo que a possível explicação para tais melhorias está associada às correlações apresentadas pela explicativa *NCONSM*, apresentadas na tabela 5.7.

Tabela 5.11 – Parâmetros estimados e testes de significância para o modelo 4.18.

Explicativa	$\beta$	Erro padrão robusto	t	P >  t
FEC	-0,7481334	0,1705597	-4,39	0,000 ***
COMP <sub>PM</sub>	1,336175	0,5189307	2,57	0,010 **
DERK	2,006904	0,8382895	2,39	0,017 **
TMAH	-0,7700361	0,1894636	-4,06	0,000 ***
CONSUMOM	-0,3580954	0,0744948	-4,81	0,000 ***
Intercepto	78,8284	1,1511492	52,15	0,000 ***

\* p < 0,10; \*\* p < 0,05; \*\*\* p < 0,01

Analisando os parâmetros estimados, apresentados na tabela 5.11, pode-se observar que os valores são bem próximos quando comparados ao modelo da equação 4.17, com exceção da explicativa associada ao consumo de energia. Para o parâmetro estimado de *CONSUMOM*, observa-se que o incremento em um TWh no consumo de energia elétrica acarretaria em uma redução de  $(0,3580954 \pm 0,0744948)$  no IASC, sendo assim outra variável, assim como o número de consumidores, que pode auxiliar na metodologia do cálculo do IASC, uma vez que a variável que indica o consumo de energia elétrica por uma concessionária não é utilizada no cálculo do IASC e, como observado na tabela 5.11, possui uma relação negativa com o mesmo.

Por fim, com o intuito de comparar o impacto entre os indicadores DEC e FEC no IASC, realizou-se a RLM no modelo apresentado pela equação 4.19. O teste de Hausman indicou  $\chi^2$  igual à 11,11 com  $Prob > \chi^2$  igual à 0,0492, isto é, à um nível de significância de 1%, não é possível rejeitar a hipótese nula, em outras palavras, o modelo de efeitos aleatórios é o mais adequado para este modelo. A estatística  $\chi^2$  para o teste de Wald

também indicou, à um nível de 1% de significância, que o modelo com os regressores propostos em 4.19 é conjuntamente significativo.

Tabela 5.12 – Resumo dos testes aplicados no modelo representado pela equação 4.19.

Teste	Hipótese nula	Estatística de teste
Autocorrelação (Wooldridge)	Sem correlação serial de 1ª ordem	Prob >F = 0,1249
Homocedasticidade (Poi & Wiggins)	Erros homocedásticos	Prob > $\chi^2$ = 0,6221
Normalidade (Skewness & Kurtosis)	Normalidade em $\alpha$	Prob > $\chi^2$ = 0,8004
	Normalidade em $\mu$	Prob > $\chi^2$ = 0,9594

Assim como no modelo para o FEC, foi possível validar as hipóteses para o modelo de RLM, cujos testes estão resumidos na tabela 5.12. Além disso, os coeficientes de determinação foram parecidos, com  $R^2_{WITHIN}$  igual à 0,0567,  $R^2_{BETWEEN}$  igual à 0,7776 e  $R^2_{OVERALL}$  igual à 0,3724, representando um aumento percentual de 5,585% quando comparado ao modelo com o FEC.

Tabela 5.13 – Matriz de coeficientes de correlação para o modelo da equação 4.19.

Explicativa	DEC	TMAH	COMP <sub>PM</sub>	DERK	CONSUMOM
<b>DEC</b>	1,0000				
<b>TMAH</b>	-0,2919	1,0000			
<b>COMP<sub>PM</sub></b>	-0,0445	-0,1668	1,0000		
<b>DERK</b>	-0,1235	0,3064	0,1521	1,0000	
<b>CONSUMOM</b>	-0,0173	-0,3549	-0,2814	-0,2253	1,0000

Fonte: Elaborada pelo autor (2021)

Uma das possibilidades para este pequeno aumento nos coeficientes de determinação é a redução da correlação entre as explicativas. Como pode-se observar em 5.13, a correlação entre DEC e as outras explicativas é menor quando comparado ao FEC e as outras explicativas, apresentado em 5.10.

Em relação aos parâmetros estimados, os resultados foram bem similares, entretanto, em uma comparação entre o DEC e FEC avaliados em modelos distintos, o FEC possui um impacto sutilmente maior. Em números, o aumento em uma unidade no FEC iria acarretar em uma redução do IASC em  $(0,7481334 \pm 0,1705597)$  enquanto que o aumento em uma unidade no DEC iria acarretar em uma redução do IASC em  $(0,5012589 \pm 0,0982789)$ .

Para este trabalho, uma vez que os valores para essas estimações são muito próximos, é inviável, uma definição objetiva de qual, entre DEC e FEC, têm o maior impacto sobre

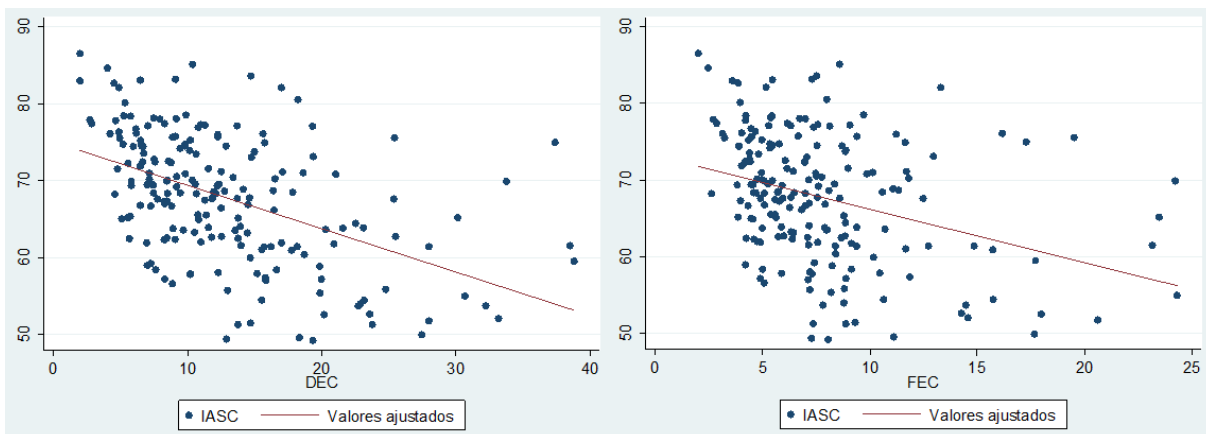
Tabela 5.14 – Parâmetros estimados e testes de significância para o modelo 4.19.

Explicativa	$\beta$	Erro padrão robusto	t	P >  t
DEC	-0,5012589	0,0982789	-5,10	0,000 ***
COMP <sub>PM</sub>	1,262446	0,5696787	2,22	0,027 **
DERK	2,113031	0,973771	2,17	0,030 **
TMAH	-0,4862249	0,1629974	-2,98	0,003 ***
CONSUMOM	-0,3250187	0,0698635	-4,65	0,000 ***
Intercepto	77,85549	1,426824	54,57	0,000 ***

\* p < 0,10; \*\* p < 0,05; \*\*\* p < 0,01

IASC, cabendo a concessionária tal avaliação, entretanto, com os modelos das equações 4.18 e 4.19 é possível se ter ao menos este parâmetro.

Figura 5.3 – IASC x DEC e IASC x FEC para o banco de dados utilizado.



Fonte: Autor.

Os gráficos apresentados na figura 5.3 ilustram a semelhança entre os indicadores DEC e FEC.

## 5.4 Considerações gerais

A partir dos modelos propostos pelas equações 4.18, 5.1 e 4.17, foi possível obter os parâmetros estimados, além de conclusões significativas para as variáveis explicativas avaliadas.

Foi possível a validação das hipóteses dos modelos de RLM, baseada nos testes de significância, para os três modelos. Em relação a heterocedasticidade, foi possível verificar que a mesma não prejudicou as estimações do modelo através dos erros padrões que, para os três modelos, não apresentou um aumento significativo quando comparado com os erros padrões robustos.



Com os dois primeiros modelos, foi possível verificar o problema salientado por Paiva et al., na apuração dos indicadores DRP e DRC, uma vez que ambos e até uma combinação deles não foram significativos com os modelos propostos (PAIVA et al., 2021).

Em relação ao número de consumidores atendidos por determinada concessionária, foi possível observar que o mesmo impacta negativamente o IASC e, uma vez que a metodologia para o cálculo do IASC não considera tal variável, este impacto pode ser um fator a ser considerado em novas reformulações do IASC.

Em paralelo à isso, a avaliação do modelo considerando o consumo de energia elétrica apresentou resultados bem similares, podendo ser outro fator importante em tal reformulação do IASC.

Além dos resultados apresentados para o número de consumidores e os indicadores DRP e DRC, outra contribuição importante deste trabalho foi em relação ao tempo médio de atendimento às ocorrências emergenciais.

Com os modelos de efeitos aleatórios, tanto para o modelo 2 quanto para o modelo 3, o parâmetro estimado foi negativo. Do parâmetro estimado para o modelo 2, por exemplo, tem-se que o aumento em uma hora no TMAH iria acarretar em uma redução de  $(0,8523816 \pm 0,1842367)$ , o que aparentemente é maior que a redução proporcionada pelo FEC que foi  $(0,749799 \pm 0,1771442)$ , entretanto, em termos práticos, o aumento em uma hora em TMAH representa uma variação muito maior quando comparada ao aumento em uma unidade em FEC.

Empregando como exemplo a concessionária Companhia Energética de Brasília (CEB), avaliando no intervalo de 2012 à 2019, a amplitude dos valores para o TMAH foi de 3,598 horas, sendo o valor máximo 7,025 horas e o mínimo 3,427 horas. Já para o indicador FEC, a amplitude foi de 11,94 unidades, sendo o máximo 17,98 e o mínimo 6,04. Assim, em termos práticos, pode-se observar que a variação do indicador FEC dentro do intervalo citado é, no mínimo, três vezes maior que a variação do indicador TMAH.

Entretanto, por mais que tal parâmetro para o TMAH não seja tão impactante no IASC, quando comparado ao FEC ou DEC, este resultado é uma nova contribuição, uma vez que na literatura especializada, não se encontrou trabalhos que avaliassem esse indicador em modelos econométricos, com exceção de Oliveira et. al., que observaram que não existe correlação entre TMAH e os indicadores de desempenho financeiro de uma concessionária (OLIVEIRA et al., 2019).

Para os três modelos propostos, os valores para os coeficientes de ajuste dentro das unidades, isto é, a variação entre observações de um mesmo indivíduo, foram baixos, indicando que os modelos propostos não se ajustam bem à variações dentro de um indivíduo. Entretanto, os valores dos coeficientes de ajuste entre os indivíduos apresentaram resultados significativos, indicando que o modelo proposto capta bem as variações entre os indivíduos, que no caso, as concessionárias de energia. Em outras palavras, os modelos propostos pelas equações captaram melhor as variações entre as concessionárias do que as variações em

relação aos anos dentro de uma concessionária.

Da avaliação do impacto das diferentes variáveis empregadas como regressores nos modelos propostos, destaca-se os parâmetros estimados para os indicadores DEC e FEC, uma vez que a melhoria na satisfação do consumidor proporcionada pela redução desses indicadores seria maior quando comparada à com os outros indicadores. Além disso, considerando as observações apresentadas por Maestri e Andrade (2018) de que a taxa de variação dos indicadores DEC e FEC carregam informações do período anterior, em outras palavras, uma redução dessa taxa de variação influencia uma redução no próximo período, proporcionar uma redução nos indicadores DEC e FEC pode ser a chave na melhoria ininterrupta da qualidade de energia ([MAESTRI; ANDRADE, 2018](#)).

## 6 Conclusões

O presente trabalho apresentou três modelos de RLM para avaliação do impacto dos indicadores da distribuição, além de outras variáveis do setor, na satisfação do consumidor residencial, avaliada pelo IASC. Os modelos propostos utilizaram como base de dados diferentes concessionárias avaliadas ao decorrer dos anos e desconsiderou as concessionárias que faltavam com alguns dados ou cujos dados representam *outliers*.

Baseado na teoria econométrica, foram empregados diversos testes de hipóteses para avaliar a significância estatística do modelo e dos regressores individualmente, além dos testes para validação do teorema de *Gauss-Markov* e as hipóteses acrescentadas na abordagem com dados em painel.

No modelo para avaliação do impacto dos indicadores DRPE, DRCE e DRPCE na satisfação do consumidor, não foi possível observar significância estatística para essas variáveis, entretanto, ao se avaliar as compensações financeiras voltadas à qualidade do produto, tal significância pode ser observada, o que reforça a dificuldade na apuração dos indicadores DRP e DRC, como levantado pela ANEEL.

Observou-se também que as concessionárias de energia elétrica com uma menor quantidade de consumidores possuem um melhor gerenciamento das condições de satisfação, isto é, o número de consumidores de uma certa concessionária impacta negativamente no IASC.

Por mais que esse resultado não represente uma alternativa de planejamento, uma vez que seria um retrocesso para a empresa, ele pode ser considerado na classificação do IASC, atribuindo um peso de acordo com o número de consumidores atendidos, uma vez que a divisão proposta atualmente considera somente as concessionárias com mais ou menos de 400 mil consumidores.

Para os dois modelos que avaliaram separadamente os indicadores DEC e FEC, os resultados foram bem similares, sendo que o modelo proposto com o DEC apresentou um aumento percentual de 5,585% em relação ao coeficiente de determinação geral. Os parâmetros estimados das variáveis que se repetiram nos dois modelos não divergiram muito, com o indicador FEC levando uma leve vantagem neste ponto.

Portanto, de acordo com os parâmetros estimados para os modelos e os testes de hipóteses realizados, priorizar ações que reduzam tanto a duração quanto a frequência das interrupções no fornecimento de energia elétrica pode ser a estratégia mais eficaz no aumento da satisfação do consumidor residencial. Além disso, propõe-se uma revisão na metodologia do IASC de forma a contemplar faixas menores de variação no número de consumidores atendidos.



# Referências

- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. b. petrov and f. csake, eds. *2nd International Symposium on Information Theory*, 1973. Cited on page 64.
- ALEJO, J.; ROJAS, G. M.; GALVAO, A.; ESCUDEIRO, W. S. Tests for normality in linear panel-data models. *The STATA Journal*, v. 15, n. 3, p. 822–832, 2015. Cited 2 times on page(s) 56 e 59.
- ANEEL. *Resolução ANEEL n<sup>o</sup> 24, de 27 de Janeiro de 2000*. [S.l.], 2000. Cited 3 times on page(s) 29, 35 e 38.
- ANEEL. *Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional - PRODIST. Módulo 8 - Qualidade de Energia Elétrica*. [S.l.], 2008. Cited 10 times on page(s) 33, 34, 36, 38, 39, 40, 41, 43, 44 e 47.
- ANEEL. *Resolução Normativa ANEEL n<sup>o</sup> 456, de 16 de Dezembro de 2008*. [S.l.], 2008. Cited on page 29.
- ANEEL. *Resolução Normativa ANEEL n<sup>o</sup> 414, de 9 de Setembro de 2010*. [S.l.], 2010. Cited 2 times on page(s) 44 e 45.
- ANEEL. *Resolução Normativa ANEEL n<sup>o</sup> 482, de 17 de Abril de 2012*. [S.l.], 2012. Cited on page 39.
- ANEEL. *Índice ANEEL de Satisfação do Consumidor (Iasc)*. 2015. Available on: <https://www.aneel.gov.br/iasc>. Cited 2 times on page(s) 48 e 49.
- ANEEL. *Resolução Normativa ANEEL n<sup>o</sup> 842, de 18 de Dezembro de 2018*. [S.l.], 2018. Cited on page 35.
- ANEEL. *Indicadores da Distribuição*. 2021. Available on: <https://www.aneel.gov.br/indicadores-da-distribuicao>. Cited 2 times on page(s) 42 e 61.
- ANEEL. *Relatórios de Consumo e Receita de Distribuição*. 2021. Available on: <https://www.aneel.gov.br/relatorios-de-consumo-e-receita>. Cited on page 61.
- ANEEL. *Resultados da pesquisa IASC*. 2021. Available on: <https://www.aneel.gov.br/pesquisas-iasc>. Cited on page 61.
- BARBOSA, L. M. Análise econômica das compensações financeiras pagas por transgressões do limite dos indicadores de continuidade na distribuição de energia elétrica o caso celg. 2018. Cited 3 times on page(s) 27, 28 e 30.
- BOFF, L. D.; CASTOLDI, M. F.; SILVA, M.; SOUZA, S. A.; SILVA, C. B. S. Melhoria da confiabilidade do fornecimento de energia elétrica em sistemas de distribuição utilizando alocação de chaves e algoritmo evolutivo multi-objetivo. *The 12th Latin-American Congress on Electricity Generation and Transmission*, p. 1–8, 2017. Cited on page 36.
- BRASIL, I. A. *Qualidade do fornecimento de energia elétrica: Confiabilidade, Conformidade e Presteza*. **white paper**. São Paulo, 2014. 36 p. Cited 3 times on page(s) 27, 33 e 34.

- BROWN, R. E. *Electric Power Distribution Reliability: Second Edition*. [S.l.]: CRC Press, 2009. Cited on page 34.
- CHOWDHURY, A.; KOVAL, D. O. *Power Distribution System Reliability: Practical Methods and Applications*. [S.l.]: John Wiley Sons Ltd, 2009. Cited on page 34.
- DNAEE. *Portaria DNAEE n<sup>o</sup> 046, de 17 de Abril de 1978*. [S.l.], 1978. Cited on page 28.
- DNAEE. *Portaria DNAEE n<sup>o</sup> 047, de 17 de Abril de 1978*. [S.l.], 1978. Cited on page 28.
- ELETROBRAS. *Desempenho de Sistemas de Distribuição*. [S.l.], 1982. v. 3. Cited on page 38.
- FERREIRA, M. I. J. *Planejamento da Manutenção de Sistemas de Distribuição Baseado em Índices de Confiabilidade*. Rio de Janeiro: [s.n.], 2019. Cited on page 28.
- FERREIRA, V. H.; OLIVEIRA, L. B.; PINHO, A. C.; HENRIQUES, H. O.; FORTES, M. Z.; NUNES, F. A.; POSE, A. C. A.; OLIVEIRA, R. B. de. Análise do impacto das ações de manutenção nos indicadores de continuidade em redes de distribuição utilizando machine learning e regressão com dados em painel. *Simpósio Brasileiro de Sistemas Elétricos*, v. 1, n. 1, 2020. Cited on page 54.
- FILHO, J. M. *Manual de Equipamentos Elétricos*. 3. ed. [S.l.]: LTC, 2005. Cited on page 33.
- GALTON, F. Regressão para mediocridade em estatura hereditária. *O Jornal do Instituto Antropológico do Grã-Bretanha e Irlanda*, v. 15, 1885. Cited on page 53.
- GAVER, D. P.; MONTMEAT, F. E.; PATTON, A. D. Power system reliability: I - measures and methods of calculation. *IEEE Trans. on Power Apparatus and Systems*, v. 83, n. 7, p. 727–737, 1964. Cited on page 35.
- GUDER, R.; COELHO, J.; ARAÚJO, W. P. P.; WILGES, B.; MATEUS, G.; STUMM, J.; FLORES, C.; MAZZOLA, V. B.; NASAR, S. M.; SARTORI, V. Um sistema de controle de qualidade de serviço de energia elétrica baseado nos indicadores de continuidade dic e fic. *VIII Conferência Internacional de Aplicações Industriais*, Poços de Caldas, 2008. Cited on page 39.
- GUJARATI, D. N.; PORTER, D. C. *Econometria básica: Quinta edição*. [S.l.]: McGrawHill Bookman Company, 2011. Cited 11 times on page(s) 51, 52, 54, 55, 56, 58, 61, 62, 63, 64 e 77.
- HAIR, J. F. J.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. *Análise Multivariada de Dados*. 5. ed. [S.l.]: Bookman, 2005. Cited on page 53.
- HONG, T.; GUI, M.; BARAN, M. E.; WILLS, H. L. Modeling and forecasting hourly electric load by multiple linear regression with interactions. *IEEE PES General Meeting*, p. 1–8, 2010. Cited on page 53.
- IEEE. *IEEE Guide for Electric Power Distribution Reliability Indices*. [S.l.], 2012. 1–43 p. Cited on page 35.

- JUNIOR, J. A. de A. *Modelo para Predição de Indicadores de Continuidade em um Sistema de Distribuição de Energia Elétrica, uma Aplicação à Gestão de Manutenção com a Perspectiva do uso da Termografia*. Salvador: [s.n.], 2016. Cited 2 times on page(s) 52 e 53.
- KAGAN, N.; OLIVEIRA, C. C. B.; ROBBA, E. J. *Introdução aos Sistemas de Distribuição de Energia Elétrica*. 2. ed. São Paulo: Blucher, 2005. Cited on page 34.
- LARSON, R.; FARBER, B. *Estatística aplicada - 6 edição*. [S.l.]: Pearson, 2015. Cited 2 times on page(s) 56 e 62.
- LEON, B. D. J.; LEITE, J. B. Otimização multiobjetivo para planejamento da manutenção preventiva em redes de distribuição considerando índices de confiabilidade e crescimento da vegetação. *LII Simpósio Brasileiro de Pesquisa Operacional*, 2020. Cited on page 36.
- LOPES, H. E. G.; PEREIRA, C. C. de P.; VIEIRA, A. F. S. Comparação entre os modelos norte-americano (acsi) e europeu (ecsi) de satisfação do cliente: um estudo no setor de serviços. *Revista de administração Mackenzie*, v. 10, n. 1, 2009. Cited on page 48.
- MAESTRI, C. O. N. M.; ANDRADE, M. E. M. C. Indicadores de qualidade do fornecimento de energia no brasil. *Encontro de Gestão e Negócios*, p. 380–394, 2018. Cited 2 times on page(s) 54 e 82.
- MAGALHAES, E. F. A. *Modelagem e simulação de indicadores de continuidade: ferramenta auxiliar para a manutenção em redes de distribuição de energia elétrica*. Salvador: [s.n.], 2017. Cited on page 53.
- MITTAL, V.; KAMAKURA, A. Satisfaction repurchase intent, repurchase behavior: Investigating the moderating effect of customer characteristics. *Journal of Marketing Research*, v. 38, p. 131–142, 2001. Cited on page 47.
- MOGHRAM, I.; RAHMAN, S. Analysis and evaluation of five short-term load forecasting techniques. *IEEE Transactions on power systems*, v. 4, n. 4, p. 1484–1491, 1989. Cited on page 53.
- MONDANI, V. U.; ALBUQUERQUE, A. A. de; LIZARELLI, F. L. Análise da relação entre os custos e a qualidade nas empresas de distribuição de energia elétrica brasileiras de capital aberto. *Gestão de Serviços*, v. 2, p. 37–45, 2017. Cited 2 times on page(s) 37 e 67.
- MONTGOMERY, D. C.; RUNGER, G. C. (Ed.). *Estatística aplicada e probabilidade para engenheiros*. 4d. [S.l.]: Gen LTC, 2009. Cited 4 times on page(s) 52, 54, 62 e 63.
- NASCIMENTO, R. M. do; OENING, A. P.; MARCILIO, D. C.; AOKI, A. R.; JUNIOR, E. P. R.; SCHIOCHET, J. M. *Algoritmo de Detecção e Correção de Outliers para Previsão de Carga*. [S.l.], 2012. Cited on page 56.
- NUNES, J. S.; SILVA, A. M. L. da; COSTA, J. G. de C. Impact of transmission and subtransmission failures on distribution reliability indices. *IEEE International Conference on Probabilistic Methods Applied to Power Systems*, p. 1, 2018. Cited on page 37.
- OLIVEIRA, A. N.; ALBUQUERQUE, A. A. de; LIZARELLI, F. L. Desempenho financeiro e qualidade: um estudo empírico no setor de distribuição de energia elétrica brasileiro. *XXVI Congresso de Iniciação Científica e XI Congresso de Iniciação em Desenvolvimento Tecnológico e Inovação*, 2019. Cited 2 times on page(s) 37 e 81.

- ONU. *Objetivo 7 - Energia Limpa e Acessível*. 2000. Available on: (<https://odsbrasil.gov.br/objetivo/objetivo?n=7>). Cited on page 27.
- PAIVA, T. V. de; TRINDADE, F. C. L.; FREITAS, W. Estudo do uso da estimação de estados para a apuração de indicadores de conformidade do nível de tensão em regime permanente em sistemas de distribuição de energia elétrica. *Simpósio Brasileiro de Sistemas Elétricos*, v. 1, n. 1, 2021. Cited 4 times on page(s) 37, 70, 76 e 81.
- PARK, H. M. Practical guides to panel data modeling: A step by step analysis using stata. tutorial working paper. *Public Management and Policy Analysis Program*, 2011. Cited on page 58.
- PELEGRINI, M. A. *A regulação das cooperativas de eletrificação rural*. São Paulo: [s.n.], 2003. Cited on page 28.
- PESSANHA, J. F. M.; SOUZA, R. C.; LAURENCELL, L. C. Um modelo de análise de envoltória de dados para o estabelecimento de metas de continuidade de fornecimento de energia elétrica. *Pesquisa Operacional*, v. 27, n. 1, p. 51–83, 2007. Cited 2 times on page(s) 29 e 36.
- PETRINI, J.; DIAS, R. A. P.; PERTILE, S. F. N.; ELER, J. P.; FERRAZ, J. B. S.; AO, G. B. M. Degree of multicollinearity and variables involved in linear dependence in additive-dominant models. *Pesquisa Agropecuária Brasileira*, v. 47, n. 12, p. 1743–1750, 2012. Cited on page 55.
- PIRES, A. M.; BRANCO, J. A. *Introdução aos métodos estatísticos robustos*. Lisboa: Sociedade Portuguesa de Estatística, 2007. Cited on page 56.
- PRADO, O. Agências reguladoras e transparência: a disponibilização de informações pela aneel. *Revista de Administração Pública*, v. 40, n. 4, p. 631–646, 2006. ISSN 0034-7612. Cited on page 35.
- RODRIGUEZ, A. M. *Planejamento de Tarefas de Manutenções Centradas em Confiabilidade para Redes de Distribuição de Energia Elétrica*. Faculdade de Engenharia do Campos de Ilha Solteira: [s.n.], 2017. Cited 2 times on page(s) 30 e 36.
- SCHWARZ, G. Estimating the dimension of a model. *Annals of Statistics*, v. 6, p. 461–464, 1978. Cited on page 64.
- SILVA, M. P. C. da; LEBORGNE, R. C.; ROSSINI, E. A influência da metodologia de regulação nos indicadores de continuidade dec e fec. *V Simpósio Brasileiro de Sistemas Elétricos*, p. 1–6, 2014. Cited on page 29.
- STATA CORP. *Longitudinal-data/Panel-data reference manual*. [S.l.], 2021. v. 17. Cited 2 times on page(s) 63 e 67.
- STEINER, M. T. A.; COSTA, C. E. da S.; COSTA, D. M. B.; FILHO, E. A.; ZAMBEDE-NEDETTI, V. C. Técnicas de pesquisa operacional aplicadas à logística de atendimento aos usuários de uma rede de distribuição de energia elétrica. *Revista Eletrônica Sistemas Gestão*, v. 1, n. 3, p. 229–243, 2006. Cited 2 times on page(s) 36 e 41.
- TODD, Z. G. A probability method for transmission and distribution outage calculations. *IEEE Trans. on Power Apparatus and Systems*, v. 83, n. 7, p. 696–701, 1964. Cited on page 35.



- TORRES, V. R. C.; FARDIN, J. F.; ENCARNACAO, L. F.; DONADEL, C. B.; FIOROTTI, R.; PEREIRA, A. L.; MEDINA, A. C. R. Reconfiguração inteligente de redes de distribuição de energia elétrica considerando o indicador fic por meio do algoritmo busca tabu. *Revista Ifes Ciência*, v. 5, n. 2, p. 24–38, 2019. ISSN 2359-4799. Cited on page [37](#).
- VACCARO, G. L. R.; MARTINS, J. C.; MENEZES, T. M. Análise estatística da qualidade de níveis de tensão em sistemas de distribuição de energia elétrica. *Produção SciELO Brasil*, v. 21, n. 3, p. 539–552, 2011. ISSN 0103-6513. Cited on page [27](#).
- VIEIRA, O. L.; CASTELO, C. M.; SILVA, F. J. R. da; COSTA, D. A. C.; SOARES, F. S.; BEZERRA, J. R.; BARROSO, G. C. Uso de testes de hipóteses para verificar a relação entre chuvas e interrupções no fornecimento de energia elétrica no sistema de distribuição. *Revista Tecnologia*, v. 41, n. 1, p. 1–11, 2020. Cited on page [37](#).
- WOOLDRIDGE, J. M. *Introdução à Econometria: uma abordagem moderna*. [S.l.]: Cengage Learning, 2010. Cited 8 times on page(s) [51](#), [52](#), [53](#), [54](#), [55](#), [56](#), [57](#) e [58](#).
- YANG, Y. Can the strenghts of aic and bic be shared? *Biometrika*, v. 92, p. 937–950, 2005. Cited on page [64](#).



# Apêndices



# APÊNDICE A – Código STATA

À seguir é apresentado os códigos utilizados no software STATA para a implementação dos modelos de regressão propostos.

## A.1 Preparação do modelo

```
* Tempo Médio de Atendimento às Ocorrências Emergenciais, expresso em horas
generate TMAE = TMP + TMD + TME
generate TMAEH = TMAH/60
* Transformação para escala de milhão
generate NCONSM = NCONS/1000000
generate CONSUMOM = CONSUMO/1000000
generate COMP_PM = COMP_P/1000000
* Transformação para escala de mil
generate DERK = DER/1000
* Define a estrutura de dados como dados em painel
* CONCESSIONARIA_ID como as unidades de corte transversal e ANO como unidades
de série temporal
tsset CONCESSIONARIA_ID ANO
* Comportamento das variáveis
xtsum CONCESSIONARIA_ID ANO IASC FEC DRPCE DERK TMAEH NCONSM
```

## A.2 Modelo 1 - DRPCE

```
* Regressão - modelo de efeitos fixos
xtreg IASC FEC DRPCE DERK TMAEH NCONSM, fe
est store FE_
* Regressão - modelo de efeitos aleatórios
xtreg IASC FEC DRPCE DERK TMAEH NCONSM, re
est store RE_
* Teste de Hausman (Especificação do modelo)
hausman FE_ RE_
* Erros padrões robustos
xtreg IASC FEC DRPCE DERK TMAEH NCONSM, re vce(robust)
```

```

* Teste de Multicolinearidade
estat vce, corr
* Teste de Homocedasticidade (Poi Wiggins)
* Avaliação da razão de verossimilhança de modelos com e sem a presença de
heterocedasticidade
xtgls IASC FEC DRPCE DERK TMAEH NCONSM, igls panels(heteroskedastic)
estimates store hetero
xtgls IASC FEC DRPCE DERK TMAEH NCONSM, igls
estimates store homosk
local df = e(N-g) -1
lrtest hetero homosk, df(`df')
* Teste de Autocorrelação
xtserial IASC FEC DRPCE DERK TMAEH NCONSM
* Teste de Normalidade
xtsktest IASC FEC DRPCE DERK TMAEH NCONSM

```

### A.3 Modelo 2 - Compensações financeiras

```

* Regressão - modelo de efeitos fixos
xtreg IASC FEC COMP_PM DERK TMAEH NCONSM, fe
est store FE_
* Regressão - modelo de efeitos aleatórios
xtreg IASC FEC COMP_PM DERK TMAEH NCONSM, re
est store RE_
* Teste de Hausman (Especificação do modelo)
hausman FE_ RE_
* Erros padrões robustos
xtreg IASC FEC COMP_PM DERK TMAEH NCONSM, re vce(robust)
* Teste de Multicolinearidade
estat vce, corr
* Teste de Homocedasticidade (Poi Wiggins)
* Avaliação da razão de verossimilhança de modelos com e sem a presença de
heterocedasticidade
xtgls IASC FEC COMP_PM DERK TMAEH NCONSM, igls panels(heteroskedastic)
estimates store hetero
xtgls IASC FEC COMP_PM DERK TMAEH NCONSM, igls
estimates store homosk
local df = e(N-g) -1

```

```
lrtest hetero homosk, df(df')
* Teste de Autocorrelação
xtserial IASC FEC COMP_PM DERK TMAEH NCONSM
* Teste de Normalidade
xtsktest IASC FEC COMP_PM DERK TMAEH NCONSM
```

## A.4 Modelo 3 - Consumo

```
* Regressão - modelo de efeitos fixos
xtreg IASC FEC COMP_PM DERK TMAEH CONSUMOM, fe
est store FE_

* Regressão - modelo de efeitos aleatórios
xtreg IASC FEC COMP_PM DERK TMAEH CONSUMOM, re
est store RE_

* Teste de Hausman (Especificação do modelo)
hausman FE_ RE_

* Erros padrões robustos
xtreg IASC FEC COMP_PM DERK TMAEH CONSUMOM, re vce(robust)

* Teste de Multicolinearidade
estat vce, corr

* Teste de Homocedasticidade (Poi Wiggins)
* Avaliação da razão de verossimilhança de modelos com e sem a presença de
heterocedasticidade
xtgls IASC FEC COMP_PM DERK TMAEH CONSUMOM, igls panels(heteroskedastic)
estimates store hetero
xtgls IASC FEC COMP_PM DERK TMAEH CONSUMOM, igls
estimates store homosk
local df = e(N-g) -1
lrtest hetero homosk, df(df')
* Teste de Autocorrelação
xtserial IASC FEC COMP_PM DERK TMAEH CONSUMOM
* Teste de Normalidade
xtsktest IASC FEC COMP_PM DERK TMAEH CONSUMOM
```