



ISEL – INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA

MEET – MESTRADO EM ENGENHARIA ELETRÓNICA E TELECOMUNICAÇÕES

DISSERTAÇÃO

Desenvolvimento de um Modelo de Consumo Energético Multi-Tecnologia e Multi-Vendor para Nós de Acesso Rádio em Redes Móveis

THAÍNA DE CASTRO SARAIVA

(Graduada em Engenharia Elétrica)

Dissertação para obtenção do grau de Mestre em Engenharia Eletrónica e Telecomunicações na área de especialização em Telecomunicações

Orientador: Doutor Pedro Manuel de Almeida Carvalho Vieira

Co-orientação: Doutora Iola Maria Silvério Pinto

Mestre David Eduardo Correia Duarte

Presidente: Doutor António João Nunes Serrador

Júri: Doutor Pedro Manuel de Almeida Carvalho Vieira

Doutor António José Castelo Branco Rodrigues

Outubro de 2019

Agradecimentos

Agradeço aos meus orientadores, Professor Doutor Pedro Vieira e Professora Doutora Iola Pinto, pela disponibilidade para orientar a presente tese e pelo rigor na revisão da mesma.

Deixo também um agradecimento muito especial à Celfinet pela oportunidade de trabalhar em um ótimo ambiente e disponibilizar todos os dados necessários, especialmente ao Eng. David Duarte, que não mediu esforços para esclarecer as minhas questões.

Finalmente, agradeço à minha família, especialmente aos meus pais, Ana e Dávio, que nunca deixaram de acreditar em mim e que tudo fizeram para que eu tivesse sempre as melhores condições possíveis.

Por último, gostaria de expressar um agradecimento muito especial à pessoa que mais me inspirou e que mais força me deu para realizar esta Dissertação. Muito obrigada Luís Elias.

Resumo

Durante as últimas décadas, as redes móveis apresentaram um crescimento constante, acompanhado do surgimento de novas tecnologias. Atualmente, esse crescimento tende a continuar, principalmente devido à introdução da próxima 5^a geração (5G) de comunicações móveis. Muitas vezes o consumo energético das redes móveis é uma questão frequentemente negligenciada em relação à de conectividade, avaliação do ritmo de transmissão, cobertura e qualidade de serviço. À medida que as redes crescem, o desafio surge em suportar o tráfego, diminuindo ou mantendo o consumo de energia. Além das responsabilidades corporativas na proteção ambiental, os operadores estão também conscientes dos seus gastos com o consumo energético, correspondendo entre 18% e 32% das despesas de *Operational Expenditure* (OPEX). Nesta perspectiva, a redução de custos e sustentabilidade ambiental podem ser vistos como objetivos convergentes.

A avaliação detalhada da quantidade de energia consumida pelos equipamentos da rede de acesso é um passo essencial para determinar soluções otimizadas. O trabalho desenvolvido nesta dissertação visa propor modelos de consumo energético para os equipamentos que constituem uma estação base, nomeadamente as *Remote Radio Units* (RRU's) e *Baseband Units* (BBU's), para as tecnologias atualmente massificadas no mercado de telecomunicações, 2G, 3G e 4G, e para os dois tipos de fabricantes *Ericsson* e *Huawei*.

A modelação dos dados baseia-se no método de inferência de regressão linear de efeitos mistos. De modo a avaliar os resultados, foram utilizadas métricas de *Akaike Information Criterion* (AIC), *Mean Absolute Percentage Error* (MAPE) e *Root Mean Square Error* (RMSE). Os resultados dos modelos atingiram um erro percentual máximo de 3.7% para os equipamento de rádio e de 8.49% para os equipamentos de banda base. Dada a evolução das gerações de comunicações móveis e o possível *switch-off* de tecnologias, apresenta-se uma estimativa da redução do consumo de energia ao desligar as tecnologias 2G e/ou 3G.

Palavras-chave: Redes móveis, Rede de Acesso Rádio, Regressão Linear de Efeitos Mistos.

Abstract

During the last decades, the mobile networks presented a constant growth, with the emergence of new technologies. Currently, this growth is likely to continue, mainly due to the introduction of the next fifth generation (5G) of mobile communications. Often the energy consumption of mobile networks is often neglected in favour of connectivity, high data throughput, coverage, and quality of service. As networks grow, the challenge lies in supporting traffic and reducing or maintaining power consumption. In addition to corporate responsibilities for environmental protection, operators are also aware of their energy consumption expenditures, accounting for 18% to 32% of *Operational Expenditure* (OPEX) expenditures. In this perspective, cost reduction and environmental sustainability can be seen as converging objectives.

Accurate assessment of the amount of power consumed by radio access network equipment is an essential step in determining optimized solutions. The work developed in this dissertation aims to propose models of energy consumption for the equipment constituting a base station, namely *Remote Radio Units* (RRU's) and *Baseband Units* (BBU's), for technologies currently massified in the telecommunications market, 2G, 3G and 4G, and for the two vendors, *Ericsson* and *Huawei*. The data modeling is based on the mixed-effects linear regression inference method. In order to evaluate the results, we used metrics of *Akaike Information Criterion* (AIC), *Mean Absolute Percentage Error* (MAPE) and *Root Mean Square Error* (RMSE). The model's results reached a maximum percentage error of 3.7 % for the radio equipment and 8.49% for the baseband equipment. Given the evolution of mobile communications generations and the possible switch-off of technologies, an estimate of the energy consumption reduction is presented by turn off 2G and/or 3G technologies.

Keywords: Mobile Networks, Radio Access Network, Mixed Effects Linear Regression.

Índice

Agradecimentos	i
Resumo	iii
Abstract	v
Lista de Figuras	ix
Lista de Tabelas	xiii
Acrónimos	xv
Lista de Símbolos	xvii
1 Introdução	1
1.1 Motivação	2
1.2 Objetivos	4
1.3 Estrutura do documento	4
1.4 Publicações	4
2 Estado da Arte	5
2.1 Redes de Acesso Rádio	6
2.2 2G - <i>Global System for Mobile Communication</i>	7
2.2.1 Arquitetura	8
2.2.2 Capacidade	9
2.3 3G - <i>Universal Mobile Telecommunication System</i>	10
2.3.1 Arquitetura	11
2.3.2 Capacidade	11
2.4 4G - <i>Long Term Evolution</i>	12
2.4.1 Arquitetura	12
2.4.2 Capacidade	13
2.5 Indicadores de desempenho	14
2.6 Soluções de Monitorização e Funcionalidades de Poupança Energética	14

3	Fundamentos Estatísticos	19
3.1	Estatística Descritiva e Inferencial	20
3.1.1	Análise de Componentes Principais	21
3.1.2	Regressão Linear de Efeitos Fixos	22
3.1.3	Regressão Linear de Efeitos Mistos	23
3.2	Metodologia	30
3.2.1	Software utilizado	32
4	Caracterização e Análise dos Dados de Consumo Energético	33
4.1	Equipamento de Medida	34
4.2	Caracterização das Estações Base	35
4.3	<i>Vendor 1 - Ericsson</i>	36
4.3.1	Equipamentos de Rádio 4G	36
4.3.2	Equipamentos de Rádio 3G	39
4.3.3	Equipamentos de Rádio 2G	41
4.3.4	Equipamentos de Rádio Partilhado 2G/3G	42
4.3.5	Equipamentos de Banda Base	43
4.4	<i>Vendor 2 - Huawei</i>	44
4.4.1	Equipamentos de Rádio 4G	44
4.4.2	Equipamentos de Rádio 3G	46
4.4.3	Equipamentos de Rádio Partilhado 2G/3G	47
4.4.4	Equipamentos de Banda Base	48
5	Desenvolvimento dos Modelos de Consumo Energético	49
5.1	<i>Vendor 1 - Ericsson</i>	50
5.1.1	Modelo RRU 4G	50
5.1.2	Modelo RRU 3G	57
5.1.3	Modelo RRU 2G	59
5.1.4	Modelo RRU Partilhado 2G/3G	60
5.1.5	Modelo BBU	63
5.2	<i>Vendor 2 - Huawei</i>	64
5.2.1	Modelo RRU 4G	64
5.2.2	Modelo RRU 3G	66
5.2.3	Modelo RRU Partilhado 2G/3G	68
5.3	Estimativas e Métricas dos Modelos	70
5.4	Extrapolação do Consumo Energético	71
6	Conclusões	73
6.1	Síntese	74
6.2	Trabalho Futuro	75

Lista de Figuras

1.1	Tráfego global de dados móveis e crescimento anual [1].	2
1.2	Consumo energético de uma estação base (adaptado de [6]).	3
2.1	Exemplo real de uma estação base [9].	6
2.2	Arquitetura do GSM (Adaptado de [12]).	8
2.3	Arquitetura do UMTS (adaptado de [13]).	11
2.4	Arquitetura do LTE (adaptado de [18]).	12
2.5	Alocação de recursos no LTE [19].	14
2.6	Consumo energético em <i>sites</i> 3G e 4G [24].	15
2.7	Exemplo de redução de potência para canais em <i>idle</i> e <i>active</i> (adaptado de [28]). . .	17
3.1	Processo de inferência estatística [30].	20
3.2	Exemplo de redução de dimensões em componentes principais [32].	21
3.3	Comparação de estimativas entre o modelo de regressão de efeitos fixos e mistos	25
4.1	Sensor de Efeito <i>Hall</i> [57].	34
4.2	Exemplo dos equipamentos que compõem uma estação base.	35
4.3	<i>Boxplot</i> da potência consumida pelos rádios 4G - <i>Ericsson</i>	36
4.4	Potência consumida pelo equipamento nº 30 - <i>Ericsson</i>	37
4.5	Poupança energética mensal devido a ativação da <i>feature Sleep</i> - <i>Ericsson</i>	37
4.6	Potência consumida pelo equipamento nº 07 - <i>Ericsson</i>	38
4.7	ACP das variáveis utilizadas no modelo para equipamentos de rádio 4G - <i>Ericsson</i> . . .	38
4.8	<i>Boxplot</i> da potência consumida pelos rádios 3G - <i>Ericsson</i>	39
4.9	ACP das variáveis utilizadas no modelo para equipamentos de rádio 3G - <i>Ericsson</i> . . .	40
4.10	Evolução do tráfego 3G nas células de 900 MHz e 2100 MHz.	40
4.11	<i>Boxplot</i> da potência consumida pelos rádios 2G - <i>Ericsson</i>	41
4.12	ACP das variáveis utilizadas no modelo para equipamentos de rádio 2G - <i>Ericsson</i> . . .	42
4.13	<i>Boxplot</i> da potência consumida pelos rádios partilhado 2G/3G - <i>Ericsson</i>	42
4.14	ACP das variáveis utilizadas no modelo para equipamentos de rádio partilhado 2G/3G - <i>Ericsson</i>	43
4.15	<i>Boxplot</i> da potência consumida pelas BBU's 4G - <i>Ericsson</i>	43
4.16	<i>Boxplot</i> da potência consumida pelas BBU's 3G - <i>Ericsson</i>	43
4.17	<i>Boxplot</i> da potência consumida pelas BBU's 2G - <i>Ericsson</i>	44

4.18	<i>Boxplot</i> da potência consumida pelos rádios 4G - <i>Huawei</i>	44
4.19	Poupança energética mensal com a ativação da <i>feature Sleep</i> - <i>Huawei</i>	45
4.20	Potência consumida pelo equipamento n° 27 com a <i>feature Radio Frequency Channel Shutdown</i> ativa - <i>Huawei</i>	46
4.21	ACP das variáveis utilizadas no modelo de rádio 4G - <i>Huawei</i>	46
4.22	<i>Boxplot</i> da potência consumida dos equipamentos de rádio 3G - <i>Huawei</i>	46
4.23	ACP das variáveis utilizadas no modelo de rádio 3G - <i>Huawei</i>	47
4.24	<i>Boxplot</i> da potência consumida nos equipamentos de rádio partilhado 2G/3G - <i>Huawei</i>	47
4.25	ACP das variáveis utilizadas no modelo de rádio partilhado 2G/3G - <i>Huawei</i>	48
4.26	<i>Boxplot</i> da potência consumida pelas BBU's 3G e 4G - <i>Huawei</i>	48
5.1	Reta de regressão linear ajustada do modelo 4G - <i>Ericsson</i>	50
5.2	Análise residual do modelo 4G com regressão linear de efeitos fixos - <i>Ericsson</i>	51
5.3	Estimativa dos coeficientes de regressão linear de efeitos fixos ajustados por equipamento no modelo 4G - <i>Ericsson</i>	52
5.4	Desvios dos coeficientes estimados do modelo 4G - <i>Ericsson</i>	54
5.5	Análise residual para verificação da homocedasticidade e normalidade dos erros no modelo 4G - <i>Ericsson</i>	55
5.6	Análise residual para verificação da linearidade e independência no modelo 4G - <i>Ericsson</i>	55
5.7	Aplicação do modelo 4G ao equipamento n° 07 da configuração 3 (1800 MHz) - <i>Ericsson</i>	56
5.8	Aplicação do modelo 4G ao equipamento n° 30 da configuração 6 (1800 MHz) - <i>Ericsson</i>	56
5.9	Análise residual para verificação da homocedasticidade e normalidade no modelo 3G - <i>Ericsson</i>	57
5.10	Análise residual para verificação da linearidade e independência no modelo 3G - <i>Ericsson</i>	58
5.11	Aplicação do modelo 3G ao equipamento n° 02 da configuração 1 (900 MHz) - <i>Ericsson</i>	58
5.12	Aplicação do modelo 3G ao equipamento n° 26 da configuração 2 (2100 MHz) - <i>Ericsson</i>	59
5.13	Análise residual para verificação da homocedasticidade no modelo 2G - <i>Ericsson</i>	59
5.14	Análise residual para verificação da linearidade e independência no modelo 2G - <i>Ericsson</i>	60
5.15	Aplicação do modelo 2G ao equipamento n° 06 - <i>Ericsson</i>	60
5.16	Análise residual para verificação da homocedasticidade no modelo partilhado 2G/3G - <i>Ericsson</i>	61
5.17	Análise residual para verificação da independência no modelo de rádio partilhado 2G/3G - <i>Ericsson</i>	61
5.18	Análise residual para verificação da homocedasticidade no modelo de rádio partilhado 2G/3G - <i>Ericsson</i>	62
5.19	Aplicação do modelo partilhado 2G/3G ao equipamento n° 02 - <i>Ericsson</i>	62
5.20	Análise residual para verificação da independência no modelo 4G - <i>Huawei</i>	64
5.21	Análise residual para verificação da homocedasticidade no modelo 4G - <i>Huawei</i>	65
5.22	Análise residual para verificação da normalidade, linearidade no modelo 4G - <i>Huawei</i>	65
5.24	Aplicação do modelo 4G ao equipamento n° 27 da configuração 3 (1800 MHz) - <i>Huawei</i>	65
5.23	Aplicação do modelo 4G ao equipamento n° 24 da configuração 3 (1800 MHz) - <i>Huawei</i>	66

5.25	Análise residual para verificação da independência no modelo 3G - <i>Huawei</i>	66
5.26	Análise residual para verificação de homocedasticidade no modelo 3G - <i>Huawei</i>	67
5.27	Análise residual para verificação da linearidade e normalidade no modelo 3G - <i>Huawei</i>	67
5.28	Aplicação do modelo 3G ao equipamento n° 02 - <i>Huawei</i>	67
5.29	Análise residual para verificação da independência no modelo 2G/3G - <i>Huawei</i>	68
5.30	Análise residual para verificação da homocedasticidade no modelo 2G/3G - <i>Huawei</i>	68
5.31	Análise residual para verificação da linearidade e normalidade no modelo 2G/3G - <i>Huawei</i>	69
5.32	Aplicação do modelo partilhado 2G/3G ao equipamento n° 08 da configuração 1 - <i>Huawei</i>	69

Lista de Tabelas

4.1	Quantidade de equipamentos por tecnologia e <i>vendor</i>	35
4.2	Configurações dos equipamentos de rádio 4G - <i>Ericsson</i>	36
4.3	Medidas estatísticas descritivas referentes às variáveis independentes 4G - <i>Ericsson</i>	37
4.4	Configurações dos equipamentos de rádio 3G - <i>Ericsson</i>	39
4.5	Medidas estatísticas descritivas referentes às variáveis independentes 3G - <i>Ericsson</i>	40
4.6	Medidas estatísticas descritivas referentes às variáveis independentes em 2G	41
4.7	Medidas estatísticas descritivas referentes às variáveis independentes 2G/3G - <i>Ericsson</i>	42
4.8	Configurações dos equipamentos de rádio utilizados no modelo 4G - <i>Huawei</i>	44
4.9	Medidas estatísticas descritivas referentes às variáveis independentes em 4G - <i>Huawei</i>	45
4.10	Medidas estatísticas descritivas referentes às variáveis independentes em 3G - <i>Huawei</i>	47
4.11	Configurações dos equipamentos de rádio utilizados no modelo partilhado 2G/3G - <i>Huawei</i>	47
4.12	Medidas estatísticas descritivas referentes às variáveis independentes em 2G/3G - <i>Huawei</i>	48
5.1	Comparação entre modelos de regressão linear de efeitos fixos 4G - <i>Ericsson</i>	51
5.2	Resumo do critério de seleção do efeito aleatório de agrupamento para o modelo 4G - <i>Ericsson</i>	53
5.3	Resumo do critério de seleção do efeito fixo para o modelo 4G - <i>Ericsson</i>	54
5.4	Resumo do critério de seleção do efeito aleatório para o modelo 4G - <i>Ericsson</i>	54
5.5	Estimativa dos intervalos de confiança a 95% para os parâmetros de efeito fixo do modelo 4G - <i>Ericsson</i>	56
5.6	Comparação entre efeitos fixos e efeitos misto no modelo 4G - <i>Ericsson</i>	57
5.7	Estimativas intervalares a 95% para os coeficientes do modelo 3G - <i>Ericsson</i>	58
5.8	Estimativas intervalares a 95% para os coeficientes do modelo 2G - <i>Ericsson</i>	60
5.9	Estimativas intervalares a 95% para os coeficientes do modelo 2G/3G - <i>Ericsson</i>	62
5.10	Estimativas intervalares a 95% para os coeficientes do modelo BBU 4G - <i>Ericsson</i>	63
5.11	Estimativas intervalares a 95% para os coeficientes do modelo BBU 3G - <i>Ericsson</i>	63
5.12	Estimativas intervalares a 95% para os coeficientes do modelo BBU 2G - <i>Ericsson</i>	63
5.13	Estimativas intervalares a 95% para os coeficientes do modelo 4G - <i>Huawei</i>	65
5.14	Estimativas intervalares a 95% para os coeficientes do modelo 3G - <i>Huawei</i>	67
5.15	Estimativas intervalares a 95% para os coeficientes do modelo 2G/3G - <i>Huawei</i>	69
5.16	Equação de regressão linear estimada dos modelos - <i>Ericsson</i>	70

5.17 Métricas dos modelos de Banda Base - <i>Ericsson</i>	70
5.18 Equação de regressão linear estimada dos modelos - <i>Huawei</i>	70
5.19 Extrapolação do consumo energético anual.	71
5.20 Impacto económico e pegada ecológica com o <i>switch-off</i> da rede 2G e/ou 3G.	72

Acrónimos

2G	<i>Segunda Geração</i>	EARTH	<i>Energy Aware Radio and Network Technologies</i>
3G	<i>Terceira Geração</i>	EDGE	<i>Enhanced Data Rates For GSM Evolution</i>
3GPP	<i>Third Generation Partnership Project</i>	EIR	<i>Equipment Identity Register</i>
4G	<i>Quarta Geração</i>	EPC	<i>Evolved Packet Core</i>
5G	<i>Quinta Geração</i>	ETSI	<i>European Telecommunications Standards Institute</i>
ACP	<i>Análise de Componentes Principais</i>	EPS	<i>Evolved Packet Switched System</i>
AIC	<i>Akaike Information Criterion</i>	E-UTRAN	<i>Evolved UTRAN</i>
AIR	<i>Antenna Integrated Radio</i>	FCCH	<i>Frequency Correction Channel</i>
ACF	<i>Autocorrelation Function</i>	GGSN	<i>Gateway GPRS Support Node</i>
AMR-WB	<i>Adaptive Multi-Rate Wideband</i>	GMSC	<i>Gateway MSC</i>
AGCH	<i>Access Grant Channel</i>	GPRS	<i>General Packet Radio Services</i>
AuC	<i>Authentication Center</i>	GSM	<i>Global System for Mobile Communications</i>
BBU	<i>Baseband Unit</i>	HLR	<i>Home Location Register</i>
BCCH	<i>Broadcast Channel</i>	HSCSD	<i>High Speed Circuit Switched Data Service</i>
BSC	<i>Base Station Controller</i>	HSDPA	<i>High Speed Downlink Packet Access</i>
BS	<i>Base Station</i>	HSPA	<i>High Speed Packet Access</i>
BTS	<i>Base Transceiver Station</i>	HSS	<i>Home Subscriber Server</i>
BSS	<i>Base Station Subsystem</i>	HSUPA	<i>High Speed Uplink Packet Access</i>
CCCH	<i>Common Control Channel</i>	IP	<i>Internet Protocol</i>
CE	<i>Channel Element</i>	ISDN	<i>Integrated Services Digital Network</i>
CN	<i>Core Network</i>	KPI	<i>Key Performance Indicators</i>
CPRI	<i>Common Public Radio Interfaces</i>		
CS	<i>Circuit Switched</i>		
CBCH	<i>Cell Broadcast Channel</i>		
eNodeB	<i>evolved NodeBs</i>		

LTE	<i>Long Term Evolution</i>		<i>Estimation</i>
MAPE	<i>Mean Absolute Percentage Error</i>	RRH	<i>Remote Radio Head</i>
ME	<i>Mobile Equipment</i>	RNC	<i>Radio Network Controller</i>
MS	<i>Mobile Station</i>	RRU	<i>Remote Radio Unit</i>
MIMO	<i>Multiple-Input Multiple-Output</i>	RU	<i>Radio Unit</i>
MME	<i>Mobility Management Entity</i>	RSS	<i>Residual Sum of Squares</i>
MLE	<i>Maximum Likelihood Estimation</i>	SC-FDMA	<i>Single Carrier - Frequency Division Multiple Access</i>
MSC	<i>Mobile Switching Center</i>	SCH	<i>Synchronization Channel</i>
NSS	<i>Network Switching Subsystem</i>	SGSN	<i>Serving GPRS Support Node</i>
OFDM	<i>Orthogonal Frequency Division Multiplexing</i>	S-GW	<i>Serving Gateway</i>
OFDMA	<i>Orthogonal Frequency Division Multiplexing Access</i>	SIM	<i>Subscriber Identity Module</i>
OSS	<i>Operations and Support System</i>	SMS	<i>Short Message Service</i>
OPEX	<i>Operational Expenditure</i>	SON	<i>Self-Organizing Networks</i>
PAPR	<i>Peak to Average Power Ratio</i>	SDCCH	<i>Stand-alone Dedicated Control Channel</i>
PCRF	<i>Policy Control and Charging Rules Function</i>	RSS	<i>Residual Sum of Squares</i>
P-GW	<i>Packet Data Network Gateway</i>	TCH	<i>Traffic Channel</i>
PCH	<i>Paging Channel</i>	TRV	<i>Teste de Razão de Verossimilhanças</i>
PS	<i>Packet Switched</i>	TAPS	<i>Traffic Aware Power Save</i>
PRB	<i>Physical Resource Block</i>	TDMA	<i>Time Division Multiple Access</i>
PRO	<i>Power Reduction Offset</i>	TRX	<i>Transceiver</i>
QoS	<i>Qualidade de Serviço</i>	UE	<i>User Equipment</i>
RAN	<i>Radio Access Network</i>	UMTS	<i>Universal Mobile Telecommunications System</i>
RACH	<i>Random Access Channel</i>	USIM	<i>UMTS Subscriber Identity Module</i>
RB	<i>Resource Block</i>	UTRAN	<i>UMTS Terrestrial RAN</i>
RE	<i>Resource Element</i>	VLR	<i>Visitor Location Register</i>
RMSE	<i>Root Mean Square Error</i>	VoLTE	<i>Voice over LTE</i>
RMLE	<i>Residual Maximum Likelihood</i>	WCDMA	<i>Wideband Code Division Multiple Access</i>

Lista de Símbolos

ϵ	Resíduos do modelo
b_i	Coefficientes de efeito aleatório
β_i	Coefficientes de efeito fixo
σ_e	Desvio padrão residual
F_{Micro_Sleep}	<i>Feature Micro-Sleep</i> de poupança energética 4G
F_{RF_OFF}	<i>Feature RF Shutdown</i> de poupança energética 4G
F_{Sleep}	<i>Feature Sleep Mode</i> de poupança energética
N_{Cells}	Número de células ativas 4G
N_{CE}	Número de <i>Channel Elements</i> 3G
P_{4G}	Potência consumida pelos equipamentos 4G
P_{3G}	Potência consumida pelos equipamentos 3G
P_{2G}	Potência consumida pelos equipamentos 2G
$P_{2G/3G}$	Potência consumida pelos equipamentos 2G/3G
T_{4G}	Tráfego de dados 4G
T_{CS_3G}	Tráfego de voz 3G
T_{PS_R99}	Tráfego de dados 3G
T_{HSDPA}	Tráfego de dados HSDPA
T_{CS_2G}	Tráfego de voz 2G
T_{PS_2G}	Tráfego de dados 2G

1

Introdução

1.1	Motivação	2
1.2	Objetivos	4
1.3	Estrutura do documento	4
1.4	Publicações	4

1.1 Motivação

O principal objetivo do planeamento e otimização de redes móveis é fornecer a máxima *Qualidade de Serviço* (QoS) e cobertura da rede aos utilizadores. Atualmente, o crescimento do tráfego está sendo impulsionado pelo aumento do número de *smartphones* e do volume médio de dados por utilizador, principalmente devido a serviços de *streaming*. A Figura 1.1 mostra que no primeiro trimestre de 2019, o tráfego mundial de dados móveis cresceu 82% em comparação com o mesmo período de 2018. O tráfego total de dados móveis continua a aumentar globalmente, prevendo-se que atinja 131 Exabytes por mês até ao final de 2024. Isto significa uma taxa de crescimento anual de 30% de 2018 a 2024 [1].

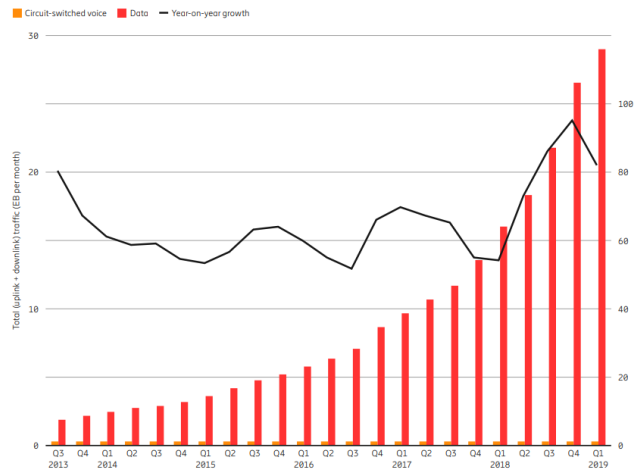


Figura 1.1: Tráfego global de dados móveis e crescimento anual [1].

O desafio surge em suportar o aumento do tráfego da rede, diminuindo ou mantendo o consumo de energia. Recentemente, tem existido o aumento do interesse em desenvolver redes de comunicação móveis mais eficientes energeticamente. Há fortes factores motivadores que impulsionam mais investigação e desenvolvimento na área das *Green Networks* [2]. É crucial minimizar o impacto ambiental deste setor devido a mudanças climáticas, causadas pelo aumento de emissões de CO_2 e outros níveis de concentração de gases de efeito estufa na atmosfera, isto devido ao uso de combustíveis fósseis como fonte primária para a produção de energia elétrica. Além das responsabilidades corporativas na proteção ambiental, os operadores de redes móveis também estão conscientes dos seus gastos com o consumo energético, que correspondem cerca de 18% a 32% das despesas de *Operational Expenditure* (OPEX) [3]. Nesta perspetiva, a redução de custos e sustentabilidade ambiental podem ser vistos como objetivos convergentes.

Os nós de acesso rádio são particularmente consumidores de energia, e podem chegar a cerca de 58% do consumo energético de um operador móvel [4]. A nível global, consomem cerca de 60 TWh/ano, e o aumento é impulsionado pelo crescimento exponencial do consumo de dados móveis [5]. Na Figura 1.2 observa-se que uma parcela significativa do consumo de energia nas estações base é devido aos equipamentos de rádio, e consumindo aproximadamente 65% do total. Entre os outros componentes da estação base, destacam-se ainda as unidades de ar condicionado (17.5%), unidades de processamento de sinal (10%) e fontes de alimentação (7%).

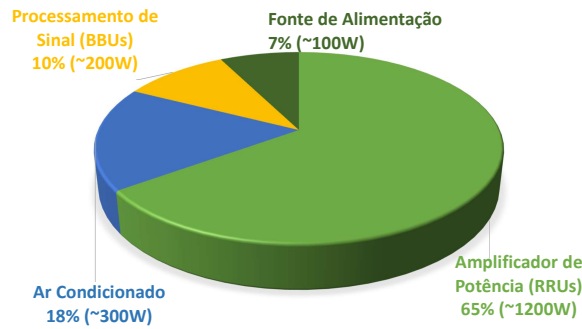


Figura 1.2: Consumo energético de uma estação base (adaptado de [6]).

A carga de tráfego nas redes móveis não é equitativamente distribuída ao longo do tempo e pelas diferentes células, ocorrendo um desperdício excessivo de energia em situações de baixo tráfego. Isto motivou vários estudos na área de eficiência energética de redes móveis, com o objetivo de projetar protocolos e propor estratégias de otimização que incluem eficiência energética na operação diária da rede.

É notório os resultados já alcançados pelos *Third Generation Partnership Project* (3GPP), como por exemplo desenvolver especificações que resultaram em baixa sinalização em períodos de baixa carga [7]. As células que apresentam maior variação na carga tem grandes oportunidades para economizar energia, tais como métodos de *switch-off* das portadoras. Melhorias podem ser alcançadas adaptando-se ao aumento real de tráfego, para isso as *Self-Organizing Networks* (SON) suportam a seleção adequada do mecanismo de poupança de energia apropriado e a reconfiguração automática colaborativa dos parâmetros da célula com as células vizinhas, cumprindo os requisitos de QoS da rede.

Embora o consumo anual de energia, devido à instalação de novas estações base, diminua cerca de 8% em média devido a avanços tecnológicos em comparação com o equipamento instalado no ano anterior, a quantidade de estações base continuam a aumentar. Em 2007 estavam a operar cerca de 3.3 milhões de estações base, estimando-se 11.2 milhões em 2020 [6]. Como resultado, o consumo de energia das estações base continuará a aumentar, principalmente com a contínua implantação das tecnologias *Quarta Geração* (4G) e *Quinta Geração* (5G) em todo o mundo [8].

Portanto, é necessário um esforço para reduzir o consumo de energia da rede de acesso rádio, e em simultâneo fornecer a qualidade de serviço esperada, levando em conta o custo associado. Enquanto soluções começam a ser introduzidas em produtos comerciais pelos fabricantes, é crucial que os operadores identifiquem o seu impacto potencial na rede. De facto, a poupança real alcançável por métodos de economização de energia depende muito do cenário de rede específico, em termos do próprio *design*, configurações e estatísticas de tráfego. Assim, é de fundamental importância para uma poupança de energia bem sucedida a monitorização das estações bases em intervalos regulares que permitem medir o consumo de energia real da rede, estimando os efeitos de variações de tráfego.

1.2 Objetivos

O objetivo desta dissertação focou-se no desenvolvimento de modelos para estimar o consumo energético das *Remote Radio Unit* (RRU)s e *Baseband Unit* (BBU)s que constituem as estações base, operando nas tecnologias *Segunda Geração* (2G), *Terceira Geração* (3G) e *Quarta Geração* (4G). Visto que a natureza do tráfego é variável nas redes móveis, a avaliação precisa sobre quais os factores que influenciam o consumo energético é o primeiro passo para determinar soluções otimizadas. Tem-se como ponto de partida um conjunto de dados reais com medições de consumo energético, obtidos por meio de sensores instalados em diferentes tipos de equipamentos de dois fornecedores. Assim, o modelo proposto para as RRUs relaciona estatísticas de tráfego coletadas da rede com o consumo energético, enquanto o modelo para as BBUs utiliza características de *hardware*. Os resultados esperados após a execução das tarefas do projeto são os seguintes:

- Apresentar um modelo de consumo energético dos equipamentos das estações base, nomeadamente RRUs e BBUs, agnóstico à tecnologia e ao fornecedor de equipamentos;
- Extrapolar o consumo energético da rede para avaliar os impactos a nível económico e ambiental, e realizar simulações com o *switch-off* das tecnologias 2G e/ou 3G.

O presente estudo é apresentado em parceria com a CELFINET, empresa portuguesa de consultoria em telecomunicações e o Instituto de Telecomunicações (IT).

1.3 Estrutura do documento

A dissertação encontra-se estruturada em seis capítulos. No Capítulo 1 realiza-se um enquadramento sobre o tema da dissertação, a sua motivação e objetivos. O Capítulo 2 tem como foco partilhar o estado da arte acerca das comunicações móveis e a caracterização dos equipamentos que constituem a estação base. No Capítulo 3 são apresentados os fundamentos estatísticos e a metodologia utilizada. O capítulo 4 aborda a caracterização e análise das variáveis utilizadas nos modelos. No Capítulo 5 são demonstrados os modelos desenvolvidos, bem como a extrapolação do modelo para a rede móvel. O capítulo 6 pretende descrever algumas conclusões do trabalho desenvolvido.

1.4 Publicações

Durante o desenvolvimento desta dissertação foram elaborados e submetidos dois artigos científicos para uma conferência nacional e uma internacional.

- Saraiva, T., Duarte, D., Pinto, I., Vieira, P., "An Enhanced Power Consumption Prediction for LTE Remote Radio Unit based on Mixed Models". The 11th Conference on Telecommunications (ConfTele), Lisboa, Portugal, Junho de 2019 (Apresentado).
- Saraiva, T., Duarte, D., Pinto, I., Vieira, P., "An Improved BBU/RRU Energy Consumption Predictor for 4G and Legacy Mobile Networks using Mixed Statistical Models". International Workshop on Computing, Networking and Communication 2020 (CNC 2020), Hawaii, USA, Fevereiro de 2020 (Aceite).

2

Estado da Arte

2.1	Redes de Acesso Rádio	6
2.2	2G - <i>Global System for Mobile Communication</i>	7
2.2.1	Arquitetura	8
2.2.2	Capacidade	9
2.3	3G - <i>Universal Mobile Telecommunication System</i>	10
2.3.1	Arquitetura	11
2.3.2	Capacidade	11
2.4	4G - <i>Long Term Evolution</i>	12
2.4.1	Arquitetura	12
2.4.2	Capacidade	13
2.5	Indicadores de desempenho	14
2.6	Soluções de Monitorização e Funcionalidades de Poupança Energética	14

Este capítulo apresenta alguns fundamentos de redes de comunicações móveis. Numa primeira parte incide-se sobre a rede de acesso rádio e as três tecnologias atualmente em operação. Posteriormente, o conceito de indicadores de desempenho da rede. Por fim, as principais soluções de monitorização e técnicas de poupança energética.

2.1 Redes de Acesso Rádio

A rede de acesso rádio é uma parte dos sistemas de telecomunicações constituída por estações base que permitem estabelecer comunicação entre os dispositivos móveis e a rede de transmissão por meio de ligações de rádio. As *macro* estações base são geralmente instaladas em torres ou no topo de edifícios. A estrutura da estação base acomoda diversos equipamentos, por exemplo, antenas, processador de rádio e banda base, baterias, ar-condicionados entre outros.

A antena é o dispositivo passivo responsável por propagar as ondas de rádio, e está localizada no topo da torre ou edifício. A unidade de processamento rádio é responsável pela codificação e decodificação, conversão analógico-digital, amplificação de potência no *downlink*, controlo de rádio, sincronização da rede e interface *Internet Protocol* (IP).

Em estações base mais recentes têm uma arquitetura distribuída, no qual a unidade de rádio, para reduzir perdas de potência dos cabos, está localizada o mais próximo possível da antena. Neste caso são chamadas de RRU ou *Remote Radio Head* (RRH), e são interligadas com a BBU via fibra óptica por padrão *Common Public Radio Interfaces* (CPRI). A Figura 2.1 apresenta os equipamentos dentro do *cabinet* e o topo da torre com as antenas e RRUs.



(a) *Cabinet* de uma estação base contendo BBU, baterias e controlo climático. (b) Três antenas com as respetivas RRUs no topo da torre.

Figura 2.1: Exemplo real de uma estação base [9].

2.2 2G - *Global System for Mobile Communication*

Ao longo da evolução das telecomunicações, vários sistemas foram desenvolvidos sem especificações padronizadas. Isto resultou em muitos problemas de compatibilidade. A grande diferença relativamente às redes de primeira geração compreende-se principalmente na evolução do sistema analógico para o digital. Da necessidade de sistemas digitais com maior capacidade, surgiu a tecnologia 2G, que teve vantagens sobre os sistemas analógicos: melhor codificação digital de voz, maior eficiência espectral, melhor qualidade de voz, transferência de dados, suporte de *roaming* internacional, compatibilidade com o *Integrated Services Digital Network* (ISDN) e a criptografia.

O *Global System for Mobile Communications* (GSM) é o padrão predominante nas redes de 2G, que foi estabelecido por um grupo de trabalho em 1982 para criar um padrão Europeu comum de telefonia móvel que formularia as especificações para um sistema celular móvel. Em 1989, o *European Telecommunications Standards Institute* (ETSI) definiu o GSM como o padrão de telefonia móvel digital internacionalmente aceite e, em 1991 surgiu o lançamento comercial na Finlândia. A tecnologia tornou-se um padrão globalmente aceite. Em 2019, o número de assinantes móveis que utilizam a tecnologia 2G em todo o mundo é cerca de 1.74 mil milhões [10].

Originalmente projetado para transportar tráfego de voz com uma taxa de transmissão de 9.6 kbps, a tecnologia utiliza *Circuit Switched* (CS), ou seja, atribui um canal físico permanente à ligação durante a duração da chamada. Além do serviço de voz, também foi introduzido o serviço de dados para mensagem de texto, chamado *Short Message Service* (SMS). Diante da necessidade de suportar as aplicações com acesso à Internet, o sistema foi aperfeiçoado. Numa primeira fase foi desenvolvido o *High Speed Circuit Switched Data Service* (HSCSD) com uma taxa de transferência de até 57.6 kbps.

Para resolver a ineficiente de alocação de recursos, o *General Packet Radio Services* (GPRS) foi desenvolvido na segunda fase do GSM, também conhecido como tecnologia 2.5 G. O GPRS consiste na utilização de técnicas de comutação de pacotes (*Packet Switched* (PS)) em vez da tradicional comutação de circuitos que permite libertar o canal para outros utilizadores caso não estejam a ser utilizados, mesmo que a conexão continue ativa. Com esse princípio, vários utilizadores podem partilhar um canal físico e alcançar uma taxa de transmissão em situações ideais de 115 kbps [11].

Na última fase do GSM foi desenvolvido o *Enhanced Data Rates For GSM Evolution* (EDGE), que consiste na implementação de novas técnicas de codificação de canal e modulação digital, permitindo altas taxas de débito de 384 kbps.

Enquanto a maioria das redes GSM opera na banda de frequência dos 900 MHz, também existem redes na faixa dos 1800 MHz e 1900 MHz (usado nos Estados Unidos e no Canadá). Para permitir uma maior interoperabilidade entre sistemas semelhantes, mas que funcionam a diferentes bandas de frequência, utilizam-se terminais *Dual-Band* (900 MHz e 1800 MHz) ou *Tri-Band* que inclui a banda de frequência dos 1900 MHz [11]. Entretanto, alguns países optaram por atribuir a banda de frequência dos 1800 MHz e 1900 MHz para que os operadores possam fazer um *refarming* de espectro e reutilizá-las para tecnologias mais recentes, como o *Long Term Evolution* (LTE).

2.2.1 Arquitetura

A arquitetura da rede GSM foi desenvolvida com o objetivo de ser simples e eficaz para fornecer os serviços de voz e dados. A Figura 2.2 mostra a arquitetura dividida em quatro partes principais: *Mobile Station* (MS), *Base Station Subsystem* (BSS), *Network Switching Subsystem* (NSS) e *Operations and Support System* (OSS).

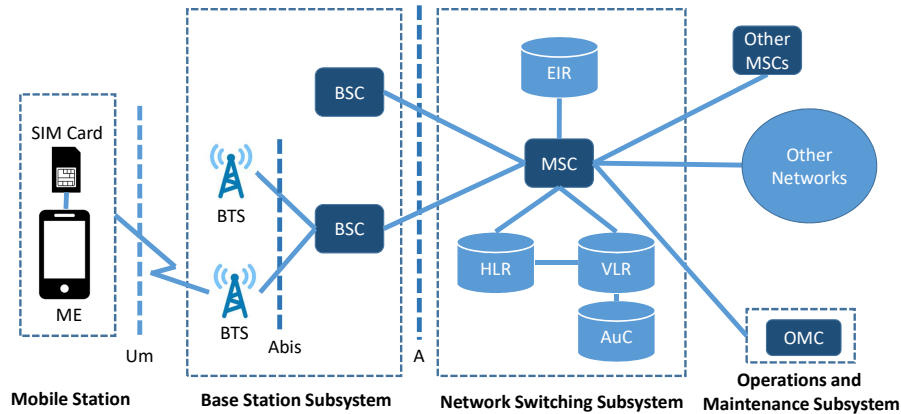


Figura 2.2: Arquitetura do GSM (Adaptado de [12]).

A comunicação entre o dispositivo móvel e a rede é estabelecida via interface aérea. Esses elementos são representados pelo *Mobile Station* (MS) ou *User Equipment* (UE) que contém o *Mobile Equipment* (ME) e o cartão *Subscriber Identity Module* (SIM), e o BSS com *Base Transceiver Station* (BTS) e *Base Station Controller* (BSC). A BTS suporta diversos equipamentos como *transceivers*, antenas e cabos de alimentação. O controlo e a inteligência de protocolo residem nos BSCs que são responsáveis, por exemplo, por funções de alocação de canal de rádio, configuração de canal e a gestão de *handovers*¹. Normalmente, um BSC controla várias BTSs, a conexão para a transmissão de informações de tráfego e sinalização entre eles são feitas através da interface *Abis* [12].

O *Mobile Switching Center* (MSC) é a parte central do *Network Switching Subsystem* (NSS). É responsável pela comutação de chamadas entre os utilizadores (entre diferentes BSCs ou para outros MSCs) e entre utilizadores da rede móvel e de rede fixa. Também incluem bases de dados necessária para armazenar informações sobre os utilizadores e gerir sua mobilidade, de modo a registrar informações, tais como localização atual e serviços dos assinantes pertencentes à área de cobertura de um MSC. O *Visitor Location Register* (VLR) mantém temporariamente informações que estão atualmente localizados em uma determinada área geográfica sob seu controlo.

As duas outras bases de dados são usadas para fins de autenticação e segurança. O *Equipment Identity Register* (EIR) é uma base de dados que contém uma lista de todos os ME válidos na rede. Já o *Authentication Center* (AuC) é uma base de dados protegida que contém uma cópia da chave secreta armazenada no cartão SIM de cada assinante, para autenticação e encriptação através do canal de rádio.

Por fim, o OSS é a entidade funcional para monitorização e controlo do sistema. O objetivo do OSS é oferecer ao cliente um suporte económico para atividades operacionais e de manutenção centralizadas, regionais e locais.

¹Processo de passagem da conexão do MS de uma célula para outra.

A técnica *Time Division Multiple Access* (TDMA) consiste na atribuição de frequência para uso exclusivo da estação móvel durante a duração de um *time slot*, sincronizados entre o transmissor e receptor. Assim, os canais físicos são definidos por uma sequência TDMA. Os canais lógicos são transmitidos nos *time slots* dos canais físicos. Os canais lógicos são subdivididos em canais de tráfego e de controlo.

Os canais de tráfego, *Traffic Channel* (TCH), são usados para a transmissão de voz e dados, em modo de comutação de circuitos. O TCH pode transmitir no modo *full rate* ou *half rate*. Esses dois tipos de canais transportam dados com débitos de 13 kbps e 6.5 kbps, respetivamente. Os canais de controlo podem ser dedicados ou partilhados pelos utilizadores. Estes são divididos em:

- *Broadcast Channel* (BCCH): é um dos canais de controlo de *broadcast*, é transmitido constantemente num canal fixo, geralmente está no *time slot* zero do primeiro *Transceiver* (TRX). Tem função de transmitir informações necessárias para os MSs fazerem seu registo na rede;
- *Frequency Correction Channel* (FCCH): transporta a sinalização para correção de frequência;
- *Synchronization Channel* (SCH): sinalização de sincronização;
- *Common Control Channel* (CCCH): são usados para transmitir informações da rede de modo a fornecer acesso ao MS. Incluem os canais *Paging Channel* (PCH) para funções de *paging*, *Random Access Channel* (RACH) para solicitar acesso ao sistema, *Access Grant Channel* (AGCH) para atribuição de um canal de controlo dedicado e o *Cell Broadcast Channel* (CBCH) para transmitir mensagens que são transmitidas para todos os MS dentro de uma célula;
- *Stand-alone Dedicated Control Channel* (SDCCH): é o canal utilizado para troca de mensagens de sinalização ou transporte de tráfego.

2.2.2 Capacidade

Cada célula GSM deve ser configurada de acordo com a finalidade de lidar com o tráfego durante a *busy hour*. E assim, a capacidade da célula é calculada de acordo com a probabilidade de bloqueio e os circuitos disponíveis. A probabilidade de bloqueio é a probabilidade de todos os canais estarem ocupados numa ligação, geralmente definindo o valor de bloqueio de 2%. A densidade de tráfego é indicada pela função de *Erlang*, em que o valor de 1 Erl indica que um canal físico é ocupado por uma hora. Por exemplo, um canal rádio que esteja ocupado 30 minutos durante uma hora transporta 0.5 Erl de tráfego.

2.3 3G - *Universal Mobile Telecommunication System*

O *Universal Mobile Telecommunications System* (UMTS) é um dos padrões mais importantes e populares do sistema de comunicação móveis de terceira geração e foi especificado pelo 3GPP. A tecnologia foi concebida para ser a evolução do GSM e possibilitar um uso efetivo da banda larga e acesso de alta velocidade para uma grande variedade de serviços, incluindo voz, acesso à Internet, videoconferência e aplicações gerais de multimídia, oferecendo uma experiência aprimorada ao utilizador final [13]. Cada serviço tem as suas próprias características de tráfego e requisitos específicos de QoS a serem atendidos pela rede, por exemplo, tempo de latência, perda de pacotes e taxa de bloqueio.

À semelhança do GSM, também suporta diversos débitos de dados e serviços comutados por circuitos. O UMTS também foi projetado para usar o *core* da rede GSM, o que garante a compatibilidade retroativa dos serviços e permite o *handover* entre as duas tecnologias. Isto proporciona aos operadores flexibilidade na prestação de serviços em diferentes áreas de cobertura.

O *Wideband Code Division Multiple Access* (WCDMA) é a tecnologia de rádio do UMTS, a qual consiste na técnica de modulação por espalhamento espectral em que usam canais cuja largura de banda é muito maior que a dos dados a serem transferidos. Assim, é possível aperfeiçoar a utilização do espectro, permitindo que vários dispositivos compartilhem ligações de canais comuns em *downlink* e alcancem taxas de até 2 Mbps para utilizadores com pouca mobilidade, e até 144 kbps em situações de alta mobilidade.

Desde que a primeira versão do UMTS na *Release 99* (R99), houve o aumento na taxa de penetração na rede e surgiu a necessidade mais serviços de dados. Assim, o *High Speed Packet Access* (HSPA) foi baseado na tecnologia UMTS, sendo basicamente uma melhoria no acesso, mantendo a rede de transmissão e *core*. Foi necessário apenas uma melhora na infra-estrutura para suportar o aumento no fluxo de dados. O HSPA consiste na melhoria no débito de *downlink* e *uplink*, composto pelo *High Speed Downlink Packet Access* (HSDPA) e *High Speed Uplink Packet Access* (HSUPA) respetivamente.

O HSDPA permite taxas de até 14.4 Mbps e foi especificado na *Release 5* do 3GPP, enquanto o HSUPA na *Release 6*. No entanto, esta não é a última versão do UMTS. O HSPA+ é um aprimoramento significativo que surge na *Release 7* com a introdução de técnicas *Multiple-Input Multiple-Output* (MIMO) no *downlink*, modulação de ordem superior para *uplink* (16-QAM) e *downlink* (64-QAM), melhorias nos protocolos e transmissão contínua de pacotes [14].

A introdução do UMTS exigiu uma nova atribuição de radiofrequências. No caso europeu, o 3GPP na *Release 99* definiu a faixa de frequência dos 2100 MHz. Apesar de ser a banda UMTS mais utilizada, alguns países usam as bandas de 850 MHz (900 MHz na Europa) e 1900 MHz. Em 2005, o 3GPP especifica a coexistência entre os dois sistemas na banda do 900 MHz, o GSM 900 MHz e o UMTS 900 MHz [15].

2.3.1 Arquitetura

A arquitetura do UMTS que foi definida pelo 3GPP na *Release 99*, está dividida em três subsistemas principais: *User Equipment* (UE), *UMTS Terrestrial RAN* (UTRAN) e *Core Network* (CN). Na Figura 2.3 são mostrados os principais elementos desta arquitetura.

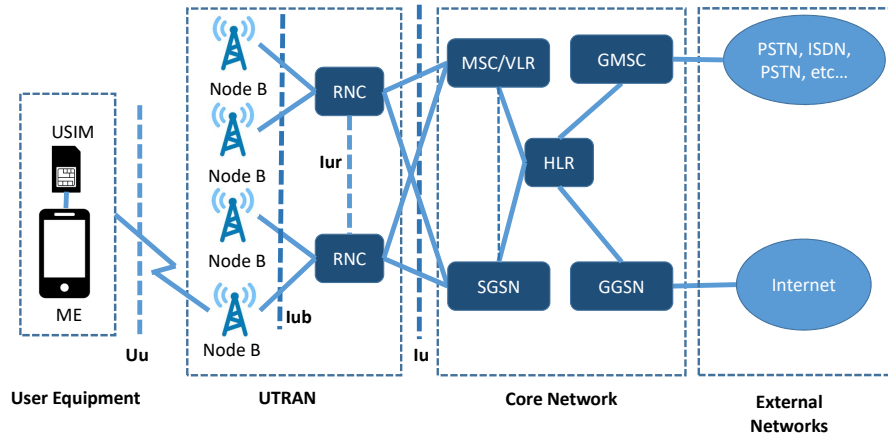


Figura 2.3: Arquitetura do UMTS (adaptado de [13]).

O UE contém o equipamento do utilizador (*Mobile Equipment* (ME)) e o *UMTS Subscriber Identity Module* (USIM), um cartão com as informações do cliente e uma chave de autenticação necessária para aceder à rede. O UTRAN é composto pelo *NodeB* (estação base), responsável pela gestão de recursos de rádio, o *Radio Network Controller* (RNC), que é o ponto de acesso de serviço para todos os serviços que a UTRAN fornece ao *Core* da rede, por exemplo, para a gestão das ligações com o UE.

Ao contrário do UE e da UTRAN que usam protocolos exclusivos para operar com a tecnologia de acesso rádio por WCDMA, o *Core Network* (CN) utiliza a estrutura do GSM, entretanto com alguns aprimoramentos, para suportar ritmos de transmissão mais elevados. No CN incluem elementos de CS e PS. Os elementos de CS são os *MSC/VLR* e *Gateway MSC* (GMSC). O MSC é responsável pela comutação de voz e inclui a base de dados VLR. O GMSC é responsável por fazer a interface entre o MSC/VLR e as redes externas. Os elementos de PS incluem o *Serving GPRS Support Node* (SGSN), um nó de comutação similar ao MSC, o *Gateway GPRS Support Node* (GGSN) funcionalidade similar do GMSC, mas em relação aos serviços de comutação de pacotes. Por fim, o *Home Location Register* (HLR) é uma base de dados com todos os perfis dos utilizadores da rede, contendo informações sobre os serviços permitidos ou proibidos ou áreas de *roaming*. Essas informações são armazenadas e mantidas enquanto o utilizador estiver ativo na rede.

2.3.2 Capacidade

Um *Channel Element* (CE) é a unidade de recursos requerida no *NodeB* para fornecer uma capacidade de canal de voz, incluindo o plano de controlo e sinalização. O número total de CEs necessários, é baseado no tipo de tráfego e depende das portadoras de rádio a serem usadas e do número de conexões simultâneas de utilizadores. Cada vez que um canal dedicado é alocado, os CEs também são alocados tanto em *uplink* como em *downlink*. Os recursos são geridos pelo *NodeB* e pelo RNC.

2.4 4G - Long Term Evolution

Ao contrário das tecnologias móveis anteriores, o LTE foi planeado para suportar apenas serviços de comutação de pacotes pelo protocolo IP. Esta nova arquitetura foca na otimização do desempenho da rede, redução de custos, utilização eficiente do espectro, menor latência e maior suporte para diferentes aplicações e plataformas, como serviços de *streaming*. Na *Release 10* apresenta um débito máximo de 3 Gbps em *downlink* e 1.5 Gbps em *uplink* [16]. Este desempenho aprimorado de serviços em comparação com as gerações anteriores das redes celulares ajudou os sistemas LTE a ganhar popularidade rápida comercial.

O LTE utiliza a técnica de *Orthogonal Frequency Division Multiplexing* (OFDM), e consiste na subdivisão da largura de banda disponível em várias subportadoras de banda estreita ortogonais, tornando-o espectralmente eficiente [17]. No *Orthogonal Frequency Division Multiplexing Access* (OFDMA) as subportadoras podem ser partilhadas entre vários utilizadores na largura de banda disponível, atribuindo dinamicamente cada utilizador a um recurso de frequência de tempo específico. A solução OFDM contempla uma taxa elevada de *Peak to Average Power Ratio* (PAPR), exigindo amplificadores de potência robustos, com altos requisitos de linearidade, aumentando o consumo de energia dos equipamentos da *Base Station* (BS). Por isso, uma solução diferente foi adotada para o *uplink*, o *Single Carrier - Frequency Division Multiple Access* (SC-FDMA), que gera um sinal com características de portadora única, e portanto, requer menor potência do UE.

Na Europa, e particularmente em Portugal, o LTE está regulamentado para operar na bandas de frequências de 800 MHz, 900 MHz, 1800 MHz, 2.1 GHz e 2.6 GHz.

2.4.1 Arquitetura

A figura 2.4 ilustra a arquitetura geral das redes LTE, em que o *Evolved Packet Core* (EPC) e o *Evolved UTRAN* (E-UTRAN) formam o *Evolved Packet Switched System* (EPS).

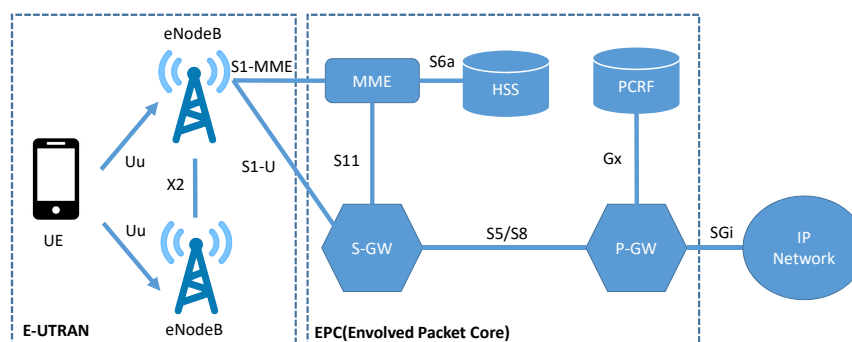


Figura 2.4: Arquitetura do LTE (adaptado de [18]).

A parte *Radio Access Network* (RAN) é denominada de E-UTRAN, e consiste na interface que conecta os UE aos *evolved NodeBs* (eNodeBs). A infra-estrutura de rádio é formada pelos seguintes elementos:

- **User Equipment:** terminais móveis e outros dispositivos que suportam o padrão LTE;
- **eNodeB:** estação base situada em toda a rede do operador. Conecta o UE via *interface Uu* à rede principal. Entre *eNodeBs* utiliza a *interface X2* e para interagir com o EPC a *interface S1*;

O EPC unifica serviço voz e dados em uma arquitetura de serviços IP. O serviço de voz LTE é denominado por *Voice over LTE* (VoLTE). Os elementos do EPC são:

- **Serving Gateway (S-GW):** responsável por transportar o tráfego de utilizadores entre os terminais móveis e externos da rede. Também conecta o E-UTRAN à rede EPC. Está ligado ao *P-GW* por meio da *interface S5/S8* para efetuar a gestão e plano do utilizador;
- **Packet Data Network Gateway (P-GW):** permite o fluxo de tráfego de dados do UE e através da *interface SGi* possibilita acesso com redes externas. Atribui endereços IP aos UE;
- **Policy Control and Charging Rules Function (PCRF):** responsável pela política de controlo, tomada de decisão de pedidos multimédia e controlo de tarifação de serviços de dados, interage com o *P-GW* usando a *interface Gx* fornecendo a política de transferência de (QoS) e as regras de cobrança;
- **Home Subscriber Server (HSS):** responsável pelo registo e pelo controlo das informações de cada dispositivo móvel;
- **Mobility Management Entity (MME):** gere conexões e atribuições de *bearers* para um UE e outras funções relacionadas ao UE, através da *interface S6a* solicita dados de subscrição e de autenticação para o HSS. Na *interface S11* troca sinalização com o MME.

2.4.2 Capacidade

A transmissão de dados em *downlink* considera a atribuição no espaço, tempo e frequência. A dimensão espacial, é dada pelas múltiplas antenas do sistema MIMO. Os recursos de tempo-frequência são compostos da seguinte forma: a maior unidade de tempo é o *trama* com 10 ms, que por sua vez é subdividida em dez *subframes* de 1 ms, cada *subframe* é subdividida em dois *slots* de 0.5 ms. Cada *slot* é composto por sete símbolos OFDM, no caso do comprimento normal do prefixo cíclico, ou seis, para o prefixo cíclico estendido.

A Figura 2.5 ilustra um *Resource Block*. Os recursos são agrupados em unidades de 12 subportadoras (ocupando no total 180 kHz com espaçamento da subportadora de 15 kHz), sendo cada unidade denominada por *Resource Block* (RB) ou *Physical Resource Block* (PRB). A menor unidade de recurso é o *Resource Element* (RE), com a duração de um símbolo OFDM, podendo ter diferentes tipos de modulações (QPSK, 16-QAM e 64-QAM).

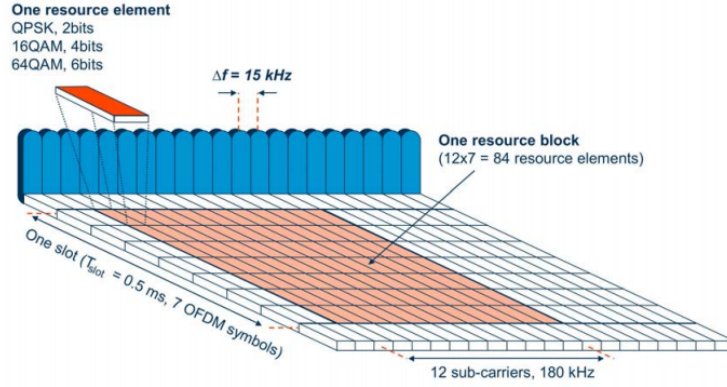


Figura 2.5: Alocação de recursos no LTE [19].

2.5 Indicadores de desempenho

Nas redes móveis, os *Key Performance Indicators* (KPI) são registrados no OSS e contêm informações sobre o desempenho da rede com granularidade detalhada. Estes indicadores consideram todos os utilizadores da rede, e medem a qualidade de serviço em termos de condições de rádio, distribuição de tráfego, utilização de recursos, entre outros parâmetros.

Os KPIs da rede de rádio podem ser usados para detetar anomalias na rede, solucionar problemas em *clusters* de células de interesse, monitorar e otimizar o desempenho da rede para uma melhor qualidade de serviço/experiência entregue ao utilizador ou para permitir um uso mais eficiente dos recursos. As unidades dos KPIs são dados em percentagem, intervalo de tempo ou unidades de tráfego. Estas medidas são agrupadas e podem ser apresentadas na sua forma absoluta, como médias, rácios ou pela sua distribuição de frequência.

2.6 Soluções de Monitorização e Funcionalidades de Poupança Energética

Um passo fundamental para avaliar o efeito de novas funcionalidades de poupança energética e de arquiteturas, é a modelação do consumo energético dos equipamentos da estação base. Esse trabalho é a continuação do estudo realizado em [20], que investiga o efeito da variação do tráfego de voz e dados no consumo energético dos equipamentos de rádio 2G, 3G e 4G. A partir dos dados de consumo energético de 6 estações base, são apresentados os modelos de consumo energético utilizando a regressão linear de efeitos fixos, obtendo erros médios entre a potência medida e a estimada que variam de 2% e 4%.

Outros modelos já foram desenvolvidos utilizando regressão linear, como os estudos apresentados em [21] e [22]. O projeto *Energy Aware Radio and Network Technologies* (EARTH) [23] investiga a relação entre a carga de utilização da estação base e o seu consumo de energia. Assim, o consumo energético de uma estação base é uma função linear dada por:

$$P_{in} = \begin{cases} N_{TRX} P_0 + \Delta_p P_{out} & 0 < P_{out} \leq P_{max} \\ N_{TRX} P_{sleep} & P_{out} = 0 \end{cases} \quad (2.1)$$

onde P_{in} é a potência consumida pela estação base, P_{out} a potência da portadora, Δ_p a taxa de crescimento dependente da potência transmitida, N_{TRX} o número de TRX e P_{max} a potência de saída da portadora para a carga máxima. Além disso, P_0 e P_{sleep} indicam o consumo de energia com carga mínima diferente de zero, e no modo *sleep*, respectivamente.

No estudo de investigação apresentado em [24] são apresentados modelos de consumo energético para os equipamentos rádio 3G e 4G, obtidos a partir de dados reais da rede Vodafone de estações base em três países. Dessa forma, conclui-se que o modelo linear utilizando o tráfego como variável explicativa, apresenta vantagem por ser robusto a alterações de horário e local, portanto, pode ser aplicado pelas operadoras móveis a toda a rede.

De facto, o coeficiente estimado correspondente à variável independente e a ordenada na origem (consumo *baseline*) obtidos no modelo linear depende do fabricante do equipamento e o tipo de tecnologia. A Figura 2.6 mostra diferenças nas curvas estimadas de consumo energético em *sites* 3G para dois fabricantes. Além disso, o modelo linear para um site 4G apresenta um perfil muito mais plano, com volumes de tráfego muito mais elevados e com menor variabilidade do que um site 3G.

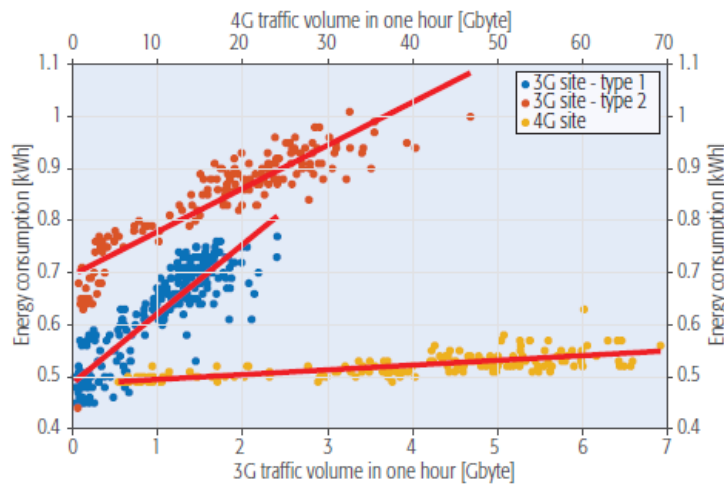


Figura 2.6: Consumo energético em *sites* 3G e 4G [24].

Para caracterizar o consumo energético, num cenário ideal seria importante monitorizar todos os equipamentos da rede ao mesmo tempo, o que não é considerado uma solução viável. Assim, uma abordagem simplificada consiste em obter uma amostra representativa, caracterizada por um perfil de tráfego e diferentes tipos de rede (por exemplo, áreas urbanas, suburbanas e rurais).

Relativamente aos métodos de poupança de energia desenvolvidos atualmente, estes concentram-se em duas áreas principais: eficiência do *hardware* e mecanismos de otimização de energia.

A eficiência do *hardware* é uma característica essencial para reduzir o consumo de energia das estações base. O desempenho da maioria dos componentes usados na atual arquitetura de rede móvel é insuficiente do ponto de vista da eficiência energética. Considerando, por exemplo, o amplificador de potência, o componente que consome a maior quantidade de energia numa estação base, mais de 90% da energia é dissipada como calor [25]. Os fabricantes de equipamentos têm evoluído neste sentido, principalmente aumentando a eficiência energética dos amplificadores de potência.

Estudos mostram que a eficiência dos amplificadores de potência pode chegar a 70%. Assim, uma quantidade substancial de economia de energia pode ser alcançada se forem adotados mais componentes eficientes na rede. No entanto, o custo de implementação para essas abordagens é elevado. Portanto, uma consideração cuidadosa sobre os aspectos operacionais e económicos pelos operadores de rede é necessária antes que as decisões sobre substituição de *hardware* sejam feitas [26].

Além disso, o consumo energético varia de acordo com a configuração da estação base. Para melhorar a cobertura, as antenas são frequentemente colocadas 20 a 60 metros acima do solo, existindo uma maior perda de potência devido à atenuação sofrida nos cabos que fazem a ligação entre as antenas e os rádios dentro do *cabinet*. Tipicamente a atenuação é de 20-50% da potência do sinal, e para uma maior atenuação de potência significa que é necessário mais energia elétrica para obter a mesma potência de saída. A solução encontrada utiliza RRUs mais próximo da antena, reduzindo a atenuação, em soluções modernas o equipamento rádio é integrado na antena - *Antenna Integrated Radio* (AIR) [27].

A ideia básica das *Green Networks* é aliviar as ineficiências resultantes do facto de que as estações base são tipicamente operadas continuamente com base nas estimativas de tráfego. Através das redes SONS é possível economizar energia monitorizando a carga de tráfego na rede e, de seguida, decidir sobre a aplicação de funcionalidades de poupança energética. Consumos de energia desnecessários, por exemplo, estações base com ar condicionado sub-carregado, podem ser evitados adotando desligamento automático e ativando técnicas de ventilação natural.

São apresentadas os principais mecanismos de poupança energética relacionadas com a redução de potência transmitida pela portadora, para cada uma das tecnologias:

- **4G Power Savings**

Nos períodos de baixo tráfego, não há necessidade de transmissão contínua de todos os canais. Nesse sentido, é possível desligar a portadora em escalas de tempo de símbolo ($71.3 \mu s$) ou mesmo desligar alguns canais [27]. Esses métodos são geridos de forma a não provocarem a redução da capacidade disponível, pois é ligado num tempo rápido o suficiente para que esteja sempre disponível para transmissões. A maior economia ocorre quando não há tráfego e apenas os canais de controlo são transmitidos.

O *Cell Sleep* consiste em desligar células LTE sobrepostas, deixando apenas as células de cobertura ativas. Em períodos de baixo tráfego, quando não é necessário grande capacidade, essa *feature* possibilita a redução do consumo de energia. O MIMO é ativado quando é necessário maior capacidade e/ou elevados valores de débito de dados. No entanto, quando não são necessárias grandes capacidades, é possível agir sobre o mecanismo MIMO desligando algum dos amplificadores. Por exemplo, pode ser desligado um amplificador numa configuração MIMO 2x2 ou até três amplificadores numa configuração MIMO 4x4. Assim, é possível alcançar uma economia substancial de energia ao mesmo tempo em que mantém o serviço LTE, embora com um *throughput* mais baixo [27].

- **3G Power Savings**

A banda de frequência no 3G pode ter uma ou mais células portadoras para aumentar o débito binário e a capacidade. O algoritmo *Traffic Aware Power Save* (TAPS) permite desligar dinamicamente as células, reduzindo o consumo energético do sector em períodos de baixo tráfego, na situação em que são desligadas todas as células, o TAPS é equivalente a *feature Sleep Mode*.

- **2G Power Savings**

A *feature* do BCCH, disponível na maioria dos fornecedores de equipamentos, economiza energia ao reduzir a potência de transmissão excessiva nos *time slots* do TRX que transporta o BCCH. A redução máxima do nível de potência permitida para um *time slot* é limitada a um nível dado pelo parâmetro de configuração *Power Reduction Offset* (PRO). Como ilustrado na Figura 2.7, a redução de potência pode ocorrer nos canais em modo *idle* e *active*, assim a potência de transmissão é regulada de acordo com os algoritmos de controlo de potência.

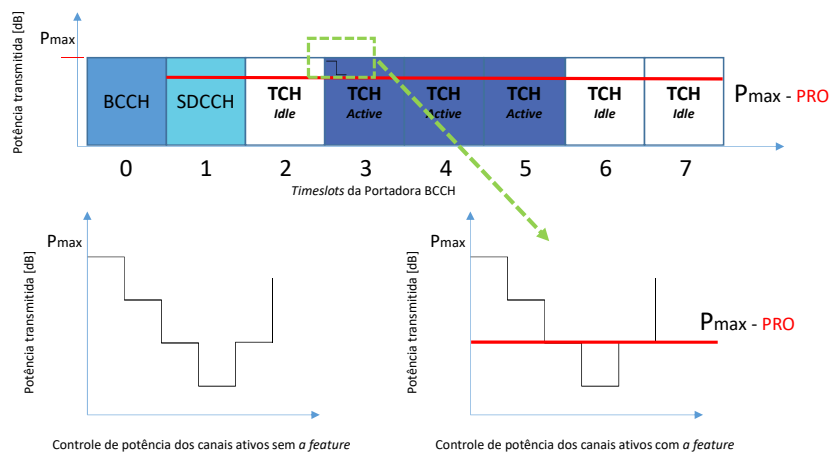


Figura 2.7: Exemplo de redução de potência para canais em *idle* e *active* (adaptado de [28]).

3

Fundamentos Estatísticos

3.1 Estatística Descritiva e Inferencial	20
3.1.1 Análise de Componentes Principais	21
3.1.2 Regressão Linear de Efeitos Fixos	22
3.1.3 Regressão Linear de Efeitos Mistos	23
3.2 Metodologia	30
3.2.1 Software utilizado	32

3.1 Estatística Descritiva e Inferencial

A análise de dados tem múltiplas abordagens e não há uma única maneira de conduzir um estudo. Neste âmbito, é inevitável destacar a importância dos métodos de inferência estatística. Para este tipo de investigação, ao invés de estudar uma população inteira, é selecionada uma amostra, e as conclusões do estudo são generalizadas para toda a população. Entretanto, para que as inferências sejam mais próximas da realidade, é necessário garantir que a amostra seja representativa, isto é, deve possuir as mesmas características da população no que diz respeito ao fenómeno em análise. O objetivo da inferência estatística consiste no desenvolvimento de procedimentos, métodos e teoremas que permitem extrair informações de dados que foram gerados por processos estocásticos (aleatórios) [29].

Os métodos estatísticos de inferência fazem estimações sobre os parâmetros da população de interesse. Entre as distribuições de probabilidade usadas para inferência, a distribuição Normal tem uma importância particular. Esta, também chamada distribuição Gaussiana é uma distribuição simétrica em torno do seu valor médio. A variável aleatória X contínua caracterizada pelos dois parâmetros: valor médio μ e variância σ^2 diz-se normal quando a sua função de densidade é dada por:

$$f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \sigma > 0. \quad (3.1)$$

Como ilustrado na Figura 3.1, para uma amostra de dimensão n , independentes com média amostral \bar{X} , proveniente de uma população com distribuição normal de variáveis aleatórias com média μ e variância σ^2 , é possível inferir sobre os parâmetros populacionais através da obtenção de estimativas pontuais e de estimativas intervalares às quais está associado um nível de confiança de 95% [30].

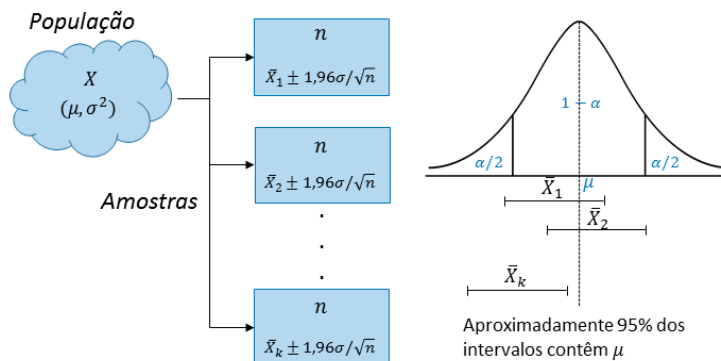


Figura 3.1: Processo de inferência estatística [30].

Utiliza-se a estatística descritiva de forma a organizar e resumir as informações obtidas de uma amostra, para assim escolher um método inferencial apropriado para o estudo [31]. A estatística descritiva permite aplicar técnicas analíticas, incluindo a construção de gráficos, tabelas e o cálculo de várias medidas descritivas, tais como medidas de posição central, de posição não central e medidas de dispersão.

Neste trabalho foram utilizadas medidas descritivas para caracterizar a localização e dispersão dos dados. Para as variáveis qualitativas nominais utilizaram-se as frequências absolutas e relativas, e

para as variáveis quantitativas, valores máximos, mínimos, mediana e percentis 25 e 75. A distribuição dos dados de potência consumida foi caracterizada pelas representações gráficas através de *boxplots*. Para a análise das possíveis associações utilizou-se o coeficiente de correlação de *Spearman* dado a não normalidade das distribuições das variáveis de interesse. Por forma a identificar quais as variáveis que mais contribuíam para explicar a variabilidade do consumo energético, e dado que algumas variáveis estão correlacionadas entre si, utilizou-se a análise de componentes principais. O desenvolvimento dos modelos de consumo energético, baseia-se na aplicação dos métodos de inferência através de modelos de regressão linear com efeitos fixos e com efeitos mistos.

3.1.1 Análise de Componentes Principais

A fim de compreender a estrutura de um conjunto de dados e como as diversas variáveis numéricas se relacionam, é utilizado o método de *Análise de Componentes Principais* (ACP). Este consiste basicamente em reescrever as coordenadas dos dados noutra sistema de eixos mais simplificado. Assim, projeta o máximo da informação contida num espaço multidimensional no mínimo de dimensões independentes, tal como representado na Figura 3.2. Para a visualização dos resultados produzidos na análise, utiliza-se o gráfico do tipo *biplot* que é semelhante ao gráfico de dispersão de duas variáveis, em que as observações são apresentadas como pontos. As variáveis categóricas não são consideradas no cálculo da ACP, mas podem ser ilustradas para representar o grupo ao qual pertencem as observações, permitindo assim detetar semelhanças e/ou diferenças entre grupos.

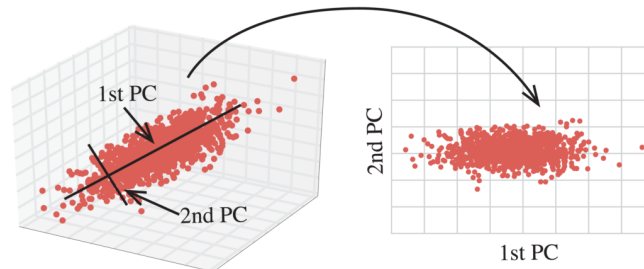


Figura 3.2: Exemplo de redução de dimensões em componentes principais [32].

As variáveis originais geram, através das suas combinações lineares, componentes principais cuja principal característica, além da ortogonalidade, é que são derivadas por ordem decrescente de importância, de modo que, por exemplo, a primeira componente principal contabiliza, tanto quanto possível, a máxima variabilidade nos dados originais. Os dados originais estão numa matriz \mathbf{X} com I linhas ($i = 1, \dots, I$) e J colunas ($j = 1, \dots, J$), sendo (x_{ij}) o valor da variável j na observação i .

As variáveis de \mathbf{X} são denotadas pelos vectores $\mathbf{x}_j (j = 1, \dots, J)$. Uma combinação linear das mesmas pode ser escrita como $\mathbf{t} = w_1 \times \mathbf{x}_1 + \dots + w_j \times \mathbf{x}_j$, onde \mathbf{t} é um novo vector no mesmo espaço das variáveis \mathbf{x} . Em notação matricial, podemos escrever $\mathbf{t} = \mathbf{X}\mathbf{w}$, sendo \mathbf{w} o vector de pesos que condensa as variáveis com os elementos $w_j (j = 1, \dots, J)$. A projecção das colunas de \mathbf{X} em \mathbf{t} é calculada usando a seguinte equação de regressão ordinária:

$$\mathbf{X} = \mathbf{t}\mathbf{p}^T + \mathbf{E} \quad (3.2)$$

onde \mathbf{p} é o vector de coeficientes da regressão e \mathbf{E} a matriz de resíduos. Assim, as variáveis de \mathbf{X} podem então ser substituídas por \mathbf{t} que retém a maioria das informações relevantes [33].

3.1.2 Regressão Linear de Efeitos Fixos

A regressão linear permite modelar relações entre variáveis e prever o valor de uma ou mais variáveis dependentes a partir de um conjunto de variáveis independentes. Os dados de entrada \mathbf{X} consistem em observações com p variáveis preditoras e através da relação linear na Equação 3.3 é possível estimar o valor de \mathbf{Y} .

$$\mathbf{Y} = \beta_0 + \sum_{j=1}^p \mathbf{X}_j \beta_j + \epsilon \quad (3.3)$$

onde termo β_0 é o termo independente (*intercept*); β_j 's são os coeficientes a serem estimados; ϵ é o vector de erros aleatórios que contém a variabilidade de \mathbf{Y} que não é explicada pelos preditores [34].

Reescrevendo a Equação 3.3 na forma matricial obtém-se a Equação 3.4, o vector de coeficientes $\boldsymbol{\beta}$ de dimensão $(p + 1)$, associados a cada variável independente na matriz $N \times (p + 1)$, e por fim ϵ o vector de resíduos com dimensão N .

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \epsilon \quad (3.4)$$

O *Residual Sum of Squares* (RSS) é uma medida da variabilidade dos valores estimados em torno dos valores reais, e pode ser definido pela Equação 3.5. O método dos mínimos quadrados permite estimar os coeficientes $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$ do modelo de forma a minimizar a função do RSS.

$$RSS(\boldsymbol{\beta}) = \sum_{i=1}^N (y_i - f(x_i))^2 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (3.5)$$

Assim, por ser uma função quadrática, existe sempre um mínimo, podendo não ser único. Diferenciando a Equação 3.5, é possível obter uma única solução para cada coeficiente, dada na Equação 3.6. Para isso, é necessário assumir que a matriz $(\mathbf{X}^T \mathbf{X})$ é positiva definida, ou seja, é simétrica, invertível e os determinantes das sub-matrizes principais são positivos [35].

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (3.6)$$

A estimativa da matriz de variância-covariância de $\hat{\boldsymbol{\beta}}$ é derivada de 3.6,

$$var(\boldsymbol{\beta}) = (\mathbf{X}^T \mathbf{X})^{-1} \sigma_e^2 \quad (3.7)$$

sendo a estimativa do parâmetro de variância residual σ_e^2 :

$$\hat{\sigma}^2 = \frac{1}{N - p - 1} \sum_{i=1}^N (y_i - \hat{y}_i) \quad (3.8)$$

Para fazer inferências sobre os parâmetros do modelo, são necessárias suposições adicionais. Assuma-se que os erros devem ter distribuição $N(\mathbf{0}, \sigma_e^2)$ (Normal com vector vector de média zero e variância constante). Por fim, a matriz de variância-covariância dos resíduos é dada por:

$$\text{var}(\mathbf{Y}) = \mathbf{V} = \sigma_e^2 \mathbf{I} = \text{var}(\boldsymbol{\epsilon}) \quad (3.9)$$

De forma geral, o modelo de regressão linear de efeitos fixos tem quatro pressupostos principais acerca da distribuição dos resíduos que devem ser utilizados na validação do modelo [36]. Por ordem decrescente de importância tem-se:

1. **Linearidade:** a relação entre as variáveis independentes e a variável dependente deve ser linear. Pode ser verificada por meio do gráfico de dispersão do valor estimado *versus* valor observado. Num diagnóstico satisfatório este gráfico deve mostrar uma tendência linear;
2. **Independência:** os resíduos devem ser independentes entre si. Para analisar a existência de autocorrelação utiliza-se a *Autocorrelation Function* (ACF);
3. **Homocedasticidade:** os resíduos devem ter variância constante. Para analisar este pressuposto utiliza-se o gráfico dos resíduos *versus* valores ajustados, verificando se os pontos estão distribuídos aleatoriamente em torno do zero e sem apresentar um padrão definido;
4. **Normalidade:** os resíduos devem seguir uma distribuição normal. Utiliza-se o gráfico Q-Q (*Quantil-Quantil*) para comparar a distribuição dos resíduos do modelo com a distribuição normal.

3.1.3 Regressão Linear de Efeitos Mistos

Para fazer inferências sobre os parâmetros no modelo de regressão linear de efeitos fixos, é necessário cumprir os pressupostos. Entretanto, nem sempre é possível que todos os pressupostos sejam satisfeitos, por exemplo, o pressuposto da independência, pode não ser verificado nos casos em que a população está estruturada por grupos, o que pode levar a conclusões incorretas ou estimativas erradas. Nesses casos, observações dentro de um mesmo grupo podem apresentar características semelhantes, indicando que não há independência entre eles [37].

O modelo de efeitos mistos fornece uma abordagem alternativa para resolver as limitações do modelo comum de regressão linear. É composto por efeitos fixos, associados ao efeito médio dos preditores na variável de resposta, e por efeitos aleatórios associados aos indivíduos selecionados aleatoriamente da população [38]. Dessa forma, consegue-se tratar a não-independência dos erros, pois as observações dentro de cada grupo são correlacionadas e devem partilhar o mesmo efeito aleatório. No caso em que os dados estão estruturados hierarquicamente em diversos níveis de efeitos aleatórios, são designados modelos de efeitos mistos multinível.

A decisão sobre considerar um efeito como fixo ou aleatório não é trivial e depende muito do estudo a ser realizado e do contexto dos dados. A inferência sobre a variável de resposta é obtida para a população utilizando amostras independentes e representativas [39]. Em certos problemas, o facto de se ignorar a estrutura de agrupamento dos dados pode levar a uma inferência incorreta, ou seja, é

possível obter uma conclusão oposta se o agrupamento não for levado em conta, como relatado no paradoxo de Simpson [40].

O modelo de regressão linear misto, de acordo com [41], pode-se descrever pela Equação 3.10, onde \mathbf{Y} é o vector das observações da variável resposta $n \times 1$, n é o número de observações, $\mathbf{X}\boldsymbol{\beta}$ é definida como a matriz de efeitos fixos e $\mathbf{Z}\mathbf{b}$ a matriz de efeitos aleatórios.

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\epsilon} \quad (3.10)$$

Para cada grupo $i = 1, \dots, M$, a Equação 3.10 pode ser reescrita:

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \quad \text{com } i = 1, \dots, M \quad (3.11)$$

\mathbf{Y}_i é o vector ($n_i \times 1$) da variável de resposta para o i -ésimo grupo, com n_i observações; $\boldsymbol{\beta}$ o vector ($p \times 1$) com p coeficientes de efeitos fixos; \mathbf{X}_i a matriz ($n_i \times p$) de covariáveis dos efeitos fixos. \mathbf{Z}_i é a matriz ($n_i \times q$) de covariáveis dos efeitos aleatórios e $\boldsymbol{\epsilon}_i$ o vector ($n_i \times 1$) de resíduos intra-grupo. Assume-se que os resíduos tem distribuição $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \mathbf{R}_i)$ (Normal com vector médio nulo e matriz de variância-covariância \mathbf{R}_i) e a distribuição do efeito aleatório $\mathbf{b}_i \sim N(0, \mathbf{G})$ (Normal com vector médio nulo e matriz de variância-covariância positiva-definida \mathbf{G}), as componentes de \mathbf{b}_i são independentes entre si e entre $\boldsymbol{\epsilon}_i$.

A escolha do efeito aleatório irá depender da interação existente entre as variáveis nominais candidatas, e além disso está dependente da forma como os dados foram obtidos. Nos modelos mistos com vários níveis de efeitos aleatórios, é possível contabilizar a variabilidade em simultâneo nos diversos níveis. Para o caso do modelo com dois níveis de efeitos aleatórios, temos:

$$\mathbf{Y}_{ij} = \mathbf{X}_{ij}\boldsymbol{\beta} + \mathbf{Z}_{i,j}\mathbf{b}_i + \mathbf{Z}_{ij}\mathbf{b}_{ij} + \boldsymbol{\epsilon}_{ij}, \text{ com } i = 1, \dots, M \text{ e } j = 1, \dots, M_i \quad (3.12)$$

onde \mathbf{Y}_{ij} é o vector ($n_{ij} \times 1$) para o j -ésimo grupo do segundo nível ($j = 1, \dots, M_i$) que está aninhado no i -ésimo grupo do primeiro nível; \mathbf{X}_{ij} é a matriz de covariáveis dos efeitos fixos; $\mathbf{Z}_{i,j}$ é a matriz ($n_{ij} \times q_1$) de covariáveis dos efeitos aleatórios de primeiro nível; \mathbf{Z}_{ij} é a matriz ($n_{ij} \times q_2$) de covariáveis dos efeitos aleatórios do segundo nível. As condições subjacentes ao modelo são: $\boldsymbol{\epsilon}_{ij} \sim N(0, \mathbf{R}_{ij})$ (Normal com vector médio nulo e matriz de variância-covariância \mathbf{R}_{ij}); $\mathbf{b}_i \sim N(0, \mathbf{G}_1)$ (Normal com vector de média zero e matriz de variância-covariância positiva-definida \mathbf{G}_1); $\mathbf{b}_{ij} \sim N(0, \mathbf{G}_2)$ (Normal com vector de média zero e matriz de variância-covariância positiva-definida \mathbf{G}_2); \mathbf{b}_i e \mathbf{b}_{ij} são independentes para diferentes i 's ou j 's e entre $\boldsymbol{\epsilon}_{ij}$.

O modelo definido em 3.11 tem duas fontes de variabilidade: a variabilidade *intra-grupos* (i) ou *entre-grupos* (ii). Para derivar a estrutura de variância neste modelo é necessário considerar o efeito fixo, o efeito aleatório e o residual. Assim a variância de \mathbf{Y}_i é dada por:

$$\text{var}(\mathbf{Y}_i) = \text{var}(\mathbf{X}_i\boldsymbol{\beta}) + \text{var}(\mathbf{Z}_i\mathbf{b}_i) + \text{var}(\boldsymbol{\epsilon}_i), \text{ com } i = 1, \dots, M \quad (3.13)$$

sendo a variância dos efeitos fixos $\text{var}(\mathbf{X}_i\boldsymbol{\beta}) = 0$, a variação *intra-grupo* dada por $\text{var}(\boldsymbol{\epsilon}_i) = \mathbf{R}_i$, e sendo a variância dos efeitos aleatórios $\text{var}(\mathbf{Z}_i\mathbf{b}_i) = \mathbf{Z}_i\text{var}(\mathbf{b}_i)\mathbf{Z}_i^T = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^T$. Tendo em consideração que \mathbf{b}_i e $\boldsymbol{\epsilon}_i$ são independentes para diferentes grupos, a matriz final de variância-covariância é dada:

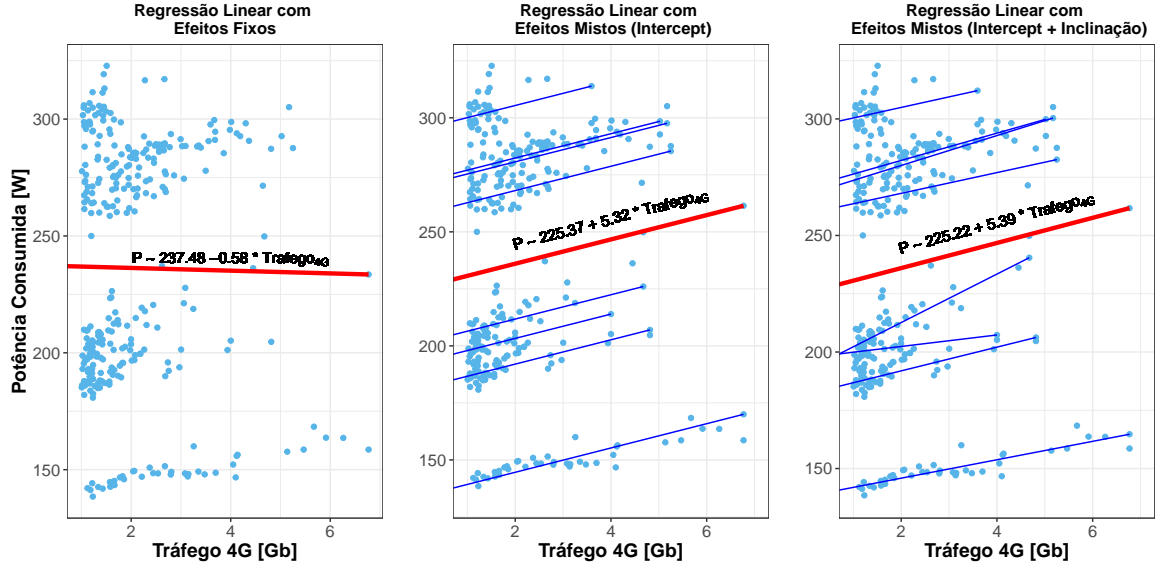


Figura 3.3: Comparação de estimativas entre o modelo de regressão de efeitos fixos e mistos

$$\text{var}(\mathbf{Y}_i) = \mathbf{V} = \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^T + \mathbf{R}_i, \text{ com } i = 1, \dots, M \quad (3.14)$$

A estrutura das matrizes \mathbf{R}_i e \mathbf{G} depende do conhecimento sobre o método de obtenção dos dados. Estas matrizes são adequadas para melhorar a representação dos dados, podendo traduzir diferentes estruturas de dependência entre as observações. A escolha correta da estrutura de covariância assegura uma melhor estimação dos parâmetros do modelo [42].

Os modelos mistos permitem que os coeficientes β 's de regressão sejam diferentes para cada grupo para melhor estimar a reta de efeito fixo. A Figura 3.3 apresenta um exemplo de estimativas da potência média consumida dos equipamentos de rádio em função do volume de tráfego 4G em períodos de 15 minutos. As amostras têm 40 observações para cada um dos 8 equipamentos de rádio. É possível verificar que na estimativa obtida utilizando o modelo de efeitos fixos, o tráfego teve pouca relação com a potência consumida. Já quando a estimativa é obtida com os modelos de efeitos mistos, com *intercepts* e inclinações aleatórias, em ambos os casos as variações no tráfego explicam muito melhor o consumo energético.

Considerando a Equação 3.11, o modelo de *intercepts* aleatórios é dado por:

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{X} + b_{0i} + \epsilon_i, \text{ com } i = 1, \dots, M \quad (3.15)$$

onde o termo b_{0i} é o *intercept* aleatório, que captura o desvio do grupo i em relação a β_0 . O modelo estima os *intercepts* para cada grupo. Os resíduos seguem uma distribuição $\epsilon_i \sim N(0, \mathbf{R}_i)$, com $\mathbf{R}_i = \sigma_e^2 \mathbf{I}$, em que \mathbf{I} a matriz de identidade de dimensão $n_i \times n_i$. Os efeitos aleatórios são representados por $\mathbf{b}_i \sim N(0, \mathbf{G})$ (Normal com vector de média zero e matriz de variância-covariância positiva-definida \mathbf{G}), em que $\mathbf{G} = \sigma_0^2 \mathbf{I}$.

O modelo 3.16 em relação ao definido em 3.15 acrescenta o termo aleatório $b_{1i} \mathbf{X}$ que pode melhorar as estimativas ao considerar que existem diferenças no coeficiente de inclinação β_1 para cada grupo.

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{X} + b_{0i} + b_{1i} \mathbf{X} + \boldsymbol{\epsilon}_i, \text{ com } i = 1, \dots, M \quad (3.16)$$

O modelo assume uma distribuição para o erro $\boldsymbol{\epsilon}_i \sim N(0, \sigma_\epsilon^2 \mathbf{I})$ e para os efeitos aleatórios $\mathbf{b}_i \sim N(0, \mathbf{G})$. Neste modelo, a diagonal da matriz \mathbf{G} tem dois parâmetros a estimar: a variabilidade que corresponde ao termo independente entre os grupos σ_0^2 e a variabilidade de desvios do coeficiente de inclinação σ_1^2 .

Para o caso em que se tem dois efeitos aleatórios, a matriz \mathbf{G} é, portanto, diagonal em blocos tendo cada bloco dimensão correspondente ao número efeitos aleatórios para cada grupo. Assim, a matriz de variância-covariâncias \mathbf{G} é positiva, sendo dada por:

$$\mathbf{G} = \begin{pmatrix} \text{var}(b_0) & \text{cov}(b_{01}) \\ \text{cov}(b_{10}) & \text{var}(b_1) \end{pmatrix} = \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix} \quad (3.17)$$

A estrutura da matriz definida em 3.17 considera a existência de correlação nos efeitos que correspondem ao termo independente e também nos que correspondem ao declive, no entanto, é possível assumir outras estruturas para a matriz de covariância dos efeitos aleatórios. Por exemplo, a matriz diagonal que considera a correlação entre os efeitos nula, definida por:

$$\mathbf{G} = \begin{pmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix} \quad (3.18)$$

Outras matrizes definidas positivas de variâncias-covariâncias podem ser utilizadas para melhorar as estimativas do modelo, como por exemplo a matriz identidade, a matriz simétrica e a matriz simétrica composta.

Métodos de Estimação

Para estimar os parâmetros do modelo, normalmente utiliza-se o método de máxima verosimilhança, *Maximum Likelihood Estimation* (MLE). Os valores dos parâmetros são encontrados de tal forma que maximizam a probabilidade que as estimativas do modelo reproduzam os dados que foram realmente observados. A função de verosimilhança (L) é uma função dos parâmetros de um modelo estatístico que permite inferir sobre o seu valor a partir de um conjunto de observações [43].

Para derivar as estimativas, a função de verosimilhança é maximizada em relação ao parâmetro estimado [44]. Assim, as estimativas de $\boldsymbol{\beta}$ e a matriz de variância-covariância da Equação 3.14 são obtidas através do cálculo dos zeros da derivada da função do logaritmo da função de verosimilhança, definida por:

$$\log(L) = -\frac{1}{2} \log|\mathbf{V}| - \frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (3.19)$$

$$\mathbf{X}' \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = 0 \quad (3.20)$$

Assim, a estimativa do parâmetro $\boldsymbol{\beta}$ é dada por:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}' \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{V}^{-1} \mathbf{y} \quad (3.21)$$

A solução da Equação 3.21 não pode ser especificada numa única equação, processos iterativos devem ser usados para obter estimativas dos parâmetros. Um desses processos é o conhecido algoritmo de *Newton-Raphson*, que consiste em resolver repetidamente uma aproximação da função do logaritmo verosimilhança.

As estimativas para a matriz de variância-covariância \mathbf{V} do modelo misto utilizando o método MLE apresentam enviesamento. Assim, é recomendado o método *Residual Maximum Likelihood Estimation* (RMLE), que se baseia em separar a probabilidade utilizada para estimar \mathbf{V} daquela que é usada para estimar β . A implicação dessa separação é que modelos com diferentes efeitos fixos não são comparáveis. No entanto, para modelos com os mesmos efeitos fixos, podem ser usadas as probabilidades do logaritmo de verosimilhança obtidas através do método RMLE.

Seleção de Modelos

O objetivo da seleção de modelos é escolher um modelo entre uma quantidade finita de modelos candidatos. Existem vários critérios para selecionar o modelo mais parcimonioso¹, ou seja, um modelo mais simples que envolva um menor número de parâmetros deve ser escolhido em relação ao mais complexo, desde que a qualidade do ajustamento seja similar.

O *Akaike Information Criterion* (AIC) é uma métrica que permite estimar a qualidade relativa de um certo modelo estatístico, dado um conjunto de dados. É utilizada com muita frequência para encontrar o melhor modelo para o caso de estudo [46]. Este utiliza o logaritmo da função de verosimilhança para medir a qualidade do ajuste, além disso penaliza modelos com excessivo número de parâmetros. Dado n variáveis preditoras, k o número de parâmetros estimados e \hat{L} , a estimativa da função de verosimilhança, define-se o valor de AIC:

$$AIC = 2k - 2\log(\hat{L}) \quad (3.22)$$

Dado um conjunto de modelos candidatos, o modelo escolhido é aquele com o valor mínimo de AIC. Assim, o AIC retribui a qualidade de ajuste, mas também inclui uma penalização que é uma função crescente do número de parâmetros estimados e dessa forma equilibra a qualidade e a complexidade do ajustamento. A penalização reduz o *overfitting*, pois aumentar o número de parâmetros no modelo pode reduzir o erro. O AIC é baseada na teoria da informação² e não usa o paradigma estatístico tradicional definido na teoria da decisão com testes de hipóteses. Portanto, não gera um valor-p, ou seja, não chega a conclusões sobre a significância estatística e não rejeita nenhum modelo.

A hipótese a ser testada nos modelos mistos pode ser que um ou mais efeitos sejam zero. O *Teste de Razão de Verosimilhanças* (TRV) é comumente usado para avaliar a diferença entre modelos aninhados. Um modelo é considerado aninhado com outro se o primeiro puder ser gerado pela imposição de restrições aos parâmetros do segundo. Modelos aninhados com diferentes efeitos fixos e com os mesmos efeitos aleatórios e, portanto, com a mesma estrutura de variância, devem ser estimados por MLE e

¹O princípio da parcimónia consiste em privilegiar os modelos mais simples, desde que a qualidade do ajuste seja similar [45].

²A teoria da informação é um ramo da teoria da probabilidade e da matemática estatística que lida com sistemas de comunicação. Foi criada por *Shannon* em 1948 para aplicar no processo de transferência de informação entre dois ou mais pontos e reproduzi-lo exatamente ou aproximadamente. É utilizada para quantificar a informação perdida ao ajustar um modelo. A teoria tem várias aplicações nas mais diversas áreas, incluindo inferência estatística, criptografia, computação quântica, dentre outras [47].

comparados com a utilização do TRV para avaliar a significância do efeito adicionado. A estatística de teste é dada por:

$$2(\log(\widehat{L}_1) - \log(\widehat{L}_0)) \quad (3.23)$$

O teste calcula a diferença do valor da probabilidade do logaritmo da verosimilhança entre os dois modelos, sendo \widehat{L}_1 a estimativa de verosimilhança do modelo mais geral, isto é, com k_1 graus de liberdade³, e \widehat{L}_0 a do modelo aninhado com $(k_1 - 1)$ graus de liberdade. Se a diferença encontrada pelo teste for muito pequena a partir do nível de significância, os dois modelos são estatisticamente iguais, podemos concluir, para um dado nível de significância que os dois modelos não diferem com significado estatístico.

Validação do modelo

Os métodos de seleção apenas escolhem o melhor modelo em comparação com outros, mas não é possível afirmar se o modelo escolhido verifica os pressupostos distribucionais exigidos pelo modelo de regressão linear. Esta etapa garante que as suposições necessárias para a realização da inferência sejam validadas.

Esse processo inclui a verificação de possíveis violações de pressupostos do modelo e também uma avaliação de quão bem o modelo captura as características dos dados [48]. Os pressupostos no modelo de efeitos mistos assume-se que os efeitos aleatórios são independentes, seguem uma distribuição normal com valor médio nulo, e com variância constante [49].

A verificação dos principais pressupostos é normalmente feita através da análise gráfica dos resíduos e dos valores ajustados. Para o caso em que o interesse recaia somente na inferência dos efeitos fixos, as estimativas são válidas mesmo que não se verifique a hipótese de distribuição normal [50].

Modelação da Heterocedasticidade e da Dependência

Por definição, o modelo linear de efeitos mistos impõe restrições para os resíduos dentro de cada grupo, assim a estrutura da matriz residual é dada por $\mathbf{R}_i = \sigma_e^2 \mathbf{I}$. Essa condição não se verifica em muitos conjuntos de dados, principalmente em séries temporais, onde várias observações são realizadas sobre o mesmo grupo e podem estar correlacionadas ou mesmo não ter variância constante. Contudo, é possível alterar essas condições modelando a estrutura de variâncias-covariâncias dos erros aleatórios. Para o caso do modelo com um nível de efeitos aleatórios, os erros para cada grupo são dados por:

$$\epsilon_i \sim N(0, \sigma_e^2 \mathbf{\Lambda}_i), \text{ com } i = 1, \dots, M \quad (3.24)$$

onde $\mathbf{\Lambda}_i$ é uma matriz positiva-definida e pode ser decomposta no produto da matriz \mathbf{W}_i de variância e a matriz \mathbf{C}_i de correlação para erros dentro do mesmo grupo, sendo definida por:

$$\mathbf{\Lambda}_i = \mathbf{W}_i \mathbf{C}_i \mathbf{W}_i, \text{ com } i = 1, \dots, M \quad (3.25)$$

A modelação da estrutura de variância dos erros dentro dos grupos implica a utilização de uma função de variância para ter um modelo heterocedástico, no qual se atribui menor peso às observações com maior variância. Por exemplo, a variância dos resíduos pode ser descrita por:

³Graus de liberdade são a quantidade de informação que os dados fornecem para estimar os valores de parâmetros populacionais desconhecidos, e calcular a variabilidade dessas estimativas.

$$Var(\epsilon_i) = \sigma_e^2 |v_i|^{2\delta} \quad (3.26)$$

correspondendo a função de variância:

$$g(v_i, \delta) = |v_i|^\delta \quad (3.27)$$

A Equação 3.27 define uma função de variância que é potência do valor absoluto da covariável v_i . O parâmetro δ pode assumir qualquer valor real, de modo que é utilizado para os casos em que a variância aumenta ou diminui com o valor absoluto da covariável. Além disso, pode assumir valores diferentes numa variável de agrupamento, ou mesmo possuir outros tipos de função como de variância fixa, variância constante e exponencial [49].

Para a modelação da dependência dos erros aleatórios para séries temporais, é possível assumir estruturas de correlações seriais [44]. Na matriz 3.28 considera-se o grupo i com quatro observações, assim a estrutura de simetria composta assume uma correlação θ constante para os erros aleatórios.

$$C_i = \begin{pmatrix} \sigma^2 & \theta & \theta & \theta \\ \theta & \sigma^2 & \theta & \theta \\ \theta & \theta & \sigma^2 & \theta \\ \theta & \theta & \theta & \sigma^2 \end{pmatrix} \quad (3.28)$$

Na estrutura da matriz 3.29 auto-regressiva de primeira ordem AR(1), assume-se que a correlação entre observações consecutivas diminui à medida que as distâncias aumentam, dado pelo parâmetro ρ . As medidas mais próximas têm correlações mais altas do que as medidas com maior tempo entre si. Esta estrutura não se aplica a dados desbalanceados, ou seja, é restrita às medições que sejam obtidas em intervalos igualmente espaçados.

$$C_i = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix} \quad (3.29)$$

Métricas de erro

As métricas de erro são utilizadas para informar acerca do quanto a previsão pode desviar-se do valor real [51]. Este conhecimento adicional fornece uma melhor percepção sobre a precisão da previsão. Assim, o *Root Mean Square Error* (RMSE) é a raiz quadrada da média das diferenças quadradas dos resíduos, Equação 3.30, e o *Mean Absolute Percentage Error* (MAPE) mede a dimensão do erro em termos percentuais, Equação 3.31.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i)^2} \quad (3.30)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\left| \frac{e_i}{y_i} \right| \right) \times 100 \quad (3.31)$$

A Equação 3.32 expressa o coeficiente de determinação, também chamado de R^2 , usado para quantificar a capacidade explicativa do modelo, ou seja, é uma medida da proporção da variância da variável resposta que é explicada pela equação de regressão linear quando estão envolvidas as variáveis independentes [51].

$$R^2 = \frac{SQR}{SQT} = 1 - \frac{SQR}{SQT}, -1 \leq R^2 \leq 1 \quad (3.32)$$

No entanto, o facto de $R^2 \simeq 1$, não significa ter um bom ajustamento aos dados, dado que a adição de uma variável independentemente da sua capacidade explicativa irá sempre provocar um aumento do valor do coeficiente. Assim sendo, é preferível utilizar o R_a^2 ajustado em que se considera a adição variáveis. Na Equação 3.33, p é o número de graus de liberdade e n o número de observações.

$$R_a^2 = 1 - \left(\frac{N - 1}{N - (p + 1)} \right) (1 - R^2) \quad (3.33)$$

Um desafio ao usar modelos de efeitos mistos é determinar como relatar a quantidade de variância explicada. O valor do R^2 em modelos mistos costuma ser bastante alto e aumenta com o número de termos de efeito aleatório, simplesmente porque muita variabilidade nos dados se resume à variabilidade devido ao agrupamento. Entretanto, é rara a apresentação do valor de variância explicada utilizando o R^2 como uma estatística de resumo relevante de modelos de efeitos mistos. Alguns métodos de cálculo do R^2 foram comparados em [52] e [53], contudo não existe um consentimento estabelecido e vários estudos discutem sobre a necessidade dessa métrica para modelos mistos.

3.2 Metodologia

Tendo em conta os fundamentos estatísticos apresentados na secção anterior, utiliza-se a seguinte metodologia para o desenvolvimento dos modelos para os equipamentos de rádio:

1. **Análise descritiva dos dados:** utilizam-se medidas estatísticas das variáveis, representações gráficas dos dados e análise de componentes principais para identificar a estrutura de agrupamento dos dados, bem como as variáveis mais importantes para explicar a variabilidade total entre as variáveis;
2. **Escolha dos efeitos aleatórios:** é escolhida uma variável de agrupamento mantendo a estrutura do efeito fixo, e caso haja a possibilidade de dois níveis é averiguado a pertinência do efeito de nível com o teste TRV;
3. **Escolha dos efeitos fixos:** estimam-se modelos pelo método MLE e, depois entre os candidatos utiliza-se a métrica de AIC para modelos não aninhados e o TRV para modelos aninhados, a fim de averiguar se as variáveis explicativas no modelo são significativas;
4. **Reajuste dos efeitos aleatórios:** a hipótese que existem diferenças nos coeficientes de inclinação entre grupos é testada. O procedimento de seleção do efeito aleatório depende do objetivo da análise. Existem duas perspectivas:

4.1 **Perspectiva 1:** consiste no modelo que utiliza todas as variáveis preditoras como efeitos aleatórios [54]. Apesar de reduzir o erro residual, ocorre com frequência a sub-parametrização do modelo, o que pode levar a obter modelo não interpretáveis, mais difíceis de ajustar devido à sua grande complexidade.

4.2 **Perspectiva 2:** apenas incluem os efeitos aleatórios que contribuem significativamente para modelar a probabilidade, usando o teste de razão de verosimilhança [55].

Neste trabalho optou-se pela Perspectiva 2, no qual será selecionado apenas um efeito de inclinação aleatório. Depois de escolhida a variável para o efeito de inclinação, é verificado o intervalo de confiança dos termos da matriz \mathbf{G} de efeitos aleatórios, e caso contenha o zero serão testadas outras estruturas.

5. **Análise de pressupostos:** utilizam-se métodos gráficos para a validação do modelo, em ordem decrescente de importância:

5.1 **Linearidade:** a suposição de linearidade significa que a relação entre as variáveis explicativas e a variável de resposta é linear. Através do gráfico de dispersão dos valores reais *versus* valores estimados é possível verificar se segue uma relação linear;

5.2 **Independência:** o modelo de regressão assume que os erros são independentes. Para isso, calcula-se a função de autocorrelação (ACF) dos resíduos considerando a estrutura de agrupamento dos dados;

5.3 **Homocedasticidade:** se a variância dos erros de regressão não for constante, pode ser adicionado a ponderação de variância, definida por uma função. A diferente variação não afeta o mais importante pressuposto de um modelo de regressão, a linearidade. Com o gráfico dos resíduos *versus* potência estimada é verificado se os erros estão aleatoriamente distribuídos em torno de zero;

5.4 **Normalidade:** os resíduos devem seguir uma distribuição normal. De facto, com a finalidade de estimar a reta de regressão (em comparação com a previsão de pontos de dados individuais), a suposição de normalidade é pouco importante [36].

6. **Modelação da heterocedasticidade e dependência:** caso haja violação dos pressupostos de linearidade, independência e homocedasticidade, será aplicado algum método de correção. Se a linearidade não é cumprida, podem ser utilizadas transformações nos preditores. No caso da falta de independência, podem ser aplicadas estruturas de correlação ou outras funções. Entretanto é possível que ao corrigir um pressuposto, o modelo se torne complexo e ocorra falha de convergência;

7. **Apresentação das estimativas:** as estimativas intervalares dos parâmetros de efeitos fixos. A validação do modelo não recai sobre o período de treino, pelo que será apresentado um exemplo da aplicação do modelo no período de teste, bem como métricas de erros para verificar a sua precisão: RMSE, MAPE e R_a^2 .

3.2.1 Software utilizado

Para o desenvolvimento deste trabalho foi utilizado a linguagem de programação **R** para a criação e análise dos modelos. Especificamente, para o desenvolvimento dos modelos mistos, o *package* **nlme** - *Linear Mixed-Effects Models* [56].

4

Caracterização e Análise dos Dados de Consumo Energético

4.1	Equipamento de Medida	34
4.2	Caracterização das Estações Base	35
4.3	Vendor 1 - Ericsson	36
4.3.1	Equipamentos de Rádio 4G	36
4.3.2	Equipamentos de Rádio 3G	39
4.3.3	Equipamentos de Rádio 2G	41
4.3.4	Equipamentos de Rádio Partilhado 2G/3G	42
4.3.5	Equipamentos de Banda Base	43
4.4	Vendor 2 - Huawei	44
4.4.1	Equipamentos de Rádio 4G	44
4.4.2	Equipamentos de Rádio 3G	46
4.4.3	Equipamentos de Rádio Partilhado 2G/3G	47
4.4.4	Equipamentos de Banda Base	48

Para o desenvolvimento dos modelos de consumo energético foram utilizados três tipos de dados distintos: *Energy Management* com medidas de consumo energético dos equipamentos que constituem a estação base, seguido de dados de *Performance Management* (PM) que reporta os KPIs que são dados estatísticos de desempenho da rede, por fim dados de *Configuration Management* (CM) com as configurações dos equipamentos, por exemplo, o modelo de *hardware*, a banda de frequência e a potência máxima transmitida.

4.1 Equipamento de Medida

Relativamente aos dados de consumo energético, foram obtidos a partir de medidas registradas por *Dataloggers* (produto desenvolvido pela empresa CELFINET). O *Datalogger* possui um sensor de efeito *Hall*, e tem como função registrar as medidas e enviá-las para uma base de dados externa [57].

O sensor de efeito *Hall*, ilustrado na Figura 4.1, é utilizado para medir a corrente eléctrica contínua, I_f , que ao percorrer um condutor gera no espaço ao redor, um campo magnético. A intensidade desse campo é proporcional à corrente eléctrica, e pode ser sentido por um material ferromagnético no qual está acoplado o sensor de efeito *Hall*. Por sua vez, surge neste sensor uma tensão induzida, V_h , proporcional à intensidade do campo magnético, e à corrente I_f a ser medida. O circuito electrónico amplifica essa voltagem e digitaliza-a.

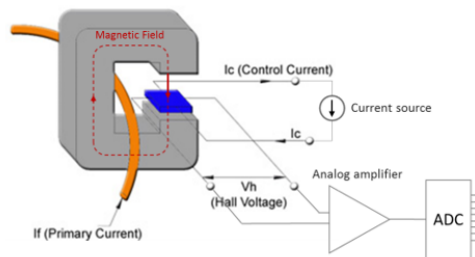


Figura 4.1: Sensor de Efeito *Hall* [57].

Dessa forma, os sensores atuam de forma não intrusiva para medir a corrente que flui através do cabo de alimentação dos equipamentos monitorizados nas estações base. A potência consumida é a taxa de trabalho realizada quando uma corrente eléctrica passa por uma diferença de potencial, dada por:

$$P = U \cdot I \quad (4.1)$$

onde P é a medida de potência instantânea em *Watt*, U a diferença de potencial em *Volts* e I a corrente medida em *Amperes*. O *Datalogger* mede valores instantâneos de potência em *Watt* (**W**). A energia, é uma medida do trabalho realizado no tempo. Neste caso, para calcular o consumo energético em **kWh**, calcula-se a potência média consumida em uma hora.

4.2 Caracterização das Estações Base

A Figura 4.2 apresenta os componentes típicos de uma BS. A energia elétrica é fornecida em corrente alternada e controlada por uma *Power Supply*, de seguida distribuída para o retificador e para a unidade de ar-condicionado (AVAC). O retificador é o módulo de distribuição de corrente contínua, este converte 220 V AC em 48 V DC. A partir do retificador, a energia é distribuída para as RRUs, BBUs, equipamentos de transmissão, entre outros, como a alimentação das baterias, sensores auxiliares, iluminação e alarmística. A principal fonte de fornecimento de energia dos equipamentos da BS é através da rede elétrica e, em caso de falha eléctrica, as baterias podem fornecer energia em corrente contínua até 12 horas.

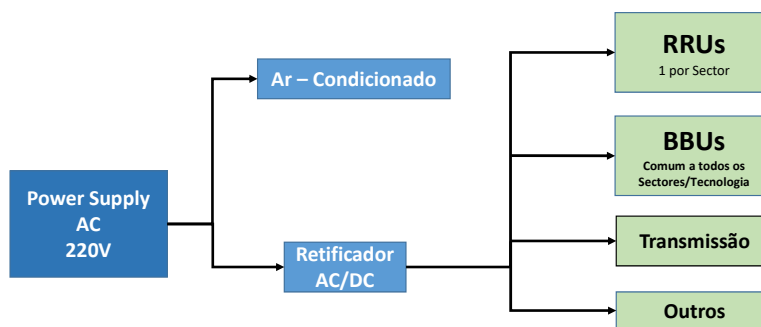


Figura 4.2: Exemplo dos equipamentos que compõem uma estação base.

As BSs são construídas de diferentes formas, de acordo com o planeamento da rede, podendo operar em diferentes tipos de tecnologias e bandas de frequência. Dessa forma, neste trabalho serão analisadas RRUs e BBUs independentes da BS na qual pertencem. Foram amostrados equipamentos de uma rede real, em áreas rurais, suburbanas e urbanas, de dois *vendors* distintos: *Ericsson*, com 171 RRUs e 47 BBUs, distribuídas por 20 *sites* localizados em Portugal, e *Huawei*, com 52 RRUs e 13 BBUs em 7 *sites* em Espanha. A quantidade de equipamentos por tecnologia para cada *vendedor* é mostrado na Tabela 4.1. Além dos equipamentos que operam em modo dedicado, em ambos os *vendors* existem equipamentos que são partilhados na banda dos 900 MHz em 2G/3G. As amostras foram divididas em 11 dias para treino e 9 para teste.

Tabela 4.1: Quantidade de equipamentos por tecnologia e *vendedor*.

	Ericsson		Huawei	
	RRU	BBU	RRU	BBU
4G	90	17	27	5
3G	45	18	6	-
2G	15	12	-	-
2G/3G	21	-	19	8
Total	171	47	52	13

4.3 Vendor 1 - Ericsson

4.3.1 Equipamentos de Rádio 4G

Foram monitorizados 90 equipamentos de rádio 4G, a operar nas bandas dos 800 MHz, 1800 MHz e 2600 MHz. Sendo a banda de frequência de 800 MHz com largura de banda de 10 MHz, e as demais com 20 MHz. A Tabela 4.2 apresenta as principais características dos equipamentos amostrados, tais como o modelo do *hardware*, a banda de frequências e a potência máxima da portadora.

Tabela 4.2: Configurações dos equipamentos de rádio 4G - *Ericsson*.

	Modelo de <i>Hardware</i>	Banda de Frequência	Potência Máx. da Portadora	Nº Equipamentos
Config. 1	Radio 2217	1800 MHz	120 W (4x30W)	3
Config. 2		800 MHz	60 W (2x30W)	3
Config. 3	Radio 4415	1800 MHz	80 W (4x20W)	3
Config. 4		2600 MHz		3
Config. 5	RRUS 32A	2600 MHz	120 W (4x30W)	3
Config. 6	RRUS 12	1800 MHz	60 W (2x30W)	18
Config. 7	RRUS 11	2600 MHz	60 W (2x30W)	3
Config. 8		800 MHz		54

Verifica-se no *boxplot* da Figura 4.3, que os equipamentos com a mesma configuração têm valores medianos próximos. No entanto, nas configurações 1 e 2, apesar de terem o mesmo modelo de *hardware*, os equipamentos da configuração 1 consomem duas vezes mais do que os equipamentos da configuração 2. Uma possível razão para essa diferença é que os equipamentos da configuração 1 operam com MIMO 4x4, utilizando 2 RRUs distintas. Entretanto, não é possível relacionar um maior consumo de potência com o facto do equipamento possuir MIMO 4x4, como verificado nas configurações 3 e 4 (nº 07 a 12) onde apresentam um valor semelhante.

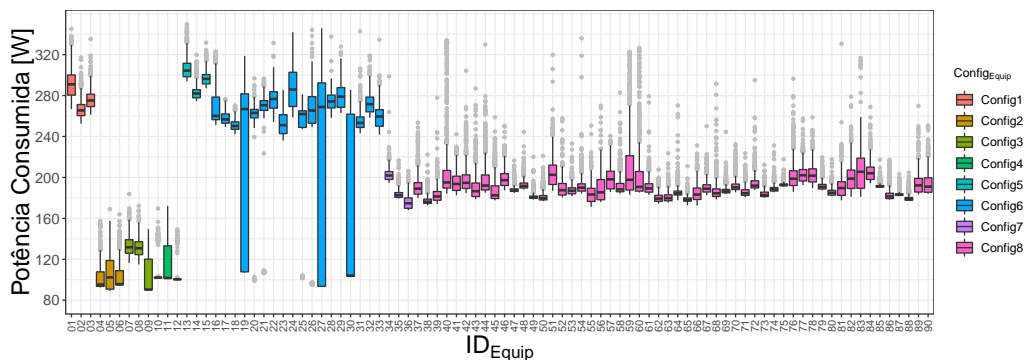


Figura 4.3: *Boxplot* da potência consumida pelos rádios 4G - *Ericsson*.

Abaixo são apresentadas as variáveis explicativas consideradas no modelo para explicar o consumo energético:

- T_{4G} : somatório do volume total de bits de tráfego e sinalização. Unidade: [Gb];
- $F_{MicroSleep}$: percentagem do tempo de ativação da *feature Micro Sleep*, que desliga a portadora no tempo de símbolo ($71.3 \mu s$). Unidade: [%];

- F_{Sleep} : variável binária que indica se determinada portadora está em modo *stand by* ("OFF"), corresponde à ativação da *feature* de poupança energética *Sleep Mode*.

A Tabela 4.3 apresenta um resumo das medidas descritivas das variáveis independentes, caracterizando a distribuição dos dados através da mediana, máximo, mínimo e percentis 25 e 75.

Tabela 4.3: Medidas estatísticas descritivas referentes às variáveis independentes 4G - *Ericsson*.

	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
T_{4G} [Gb]	0	0.04	0.33	1.08	34.61
$F_{MicroSleep}$ [%]	0	0	0	0	63.0
F_{Sleep}	0 = 'ON'	-	-	-	1 = 'OFF'

No conjunto de dados amostrados, 11 equipamentos (n° 09-12, 19-21, 25-27 e 30) apresentam períodos com a *feature* de *Sleep Mode* activa. O tempo de ativação da *feature* é diferente em cada equipamento. Observa-se na Figura 4.4 um equipamento que é desligado cerca de 10 horas por dia, onde a potência reduz de 250 W para 100 W. Nos períodos que o equipamento está ligado a potência consumida está positivamente correlacionada com o volume de tráfego, tendo coeficiente de correlação de *Spearman* de 0.80.

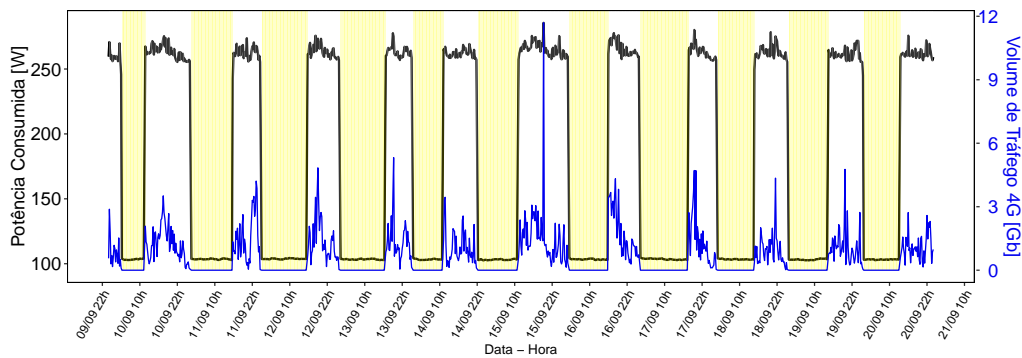


Figura 4.4: Potência consumida pelo equipamento n° 30 - *Ericsson*.

Para analisar a redução no consumo energético mensal, consideram-se os equipamentos amostrados com 16 horas ativas e 8 horas em modo *stand by*. A Figura 4.5 apresenta o resultado dessa análise, atingindo uma poupança média mensal de cerca de 20% nos equipamentos da configuração 6, e 12% nas configurações 3 e 4.

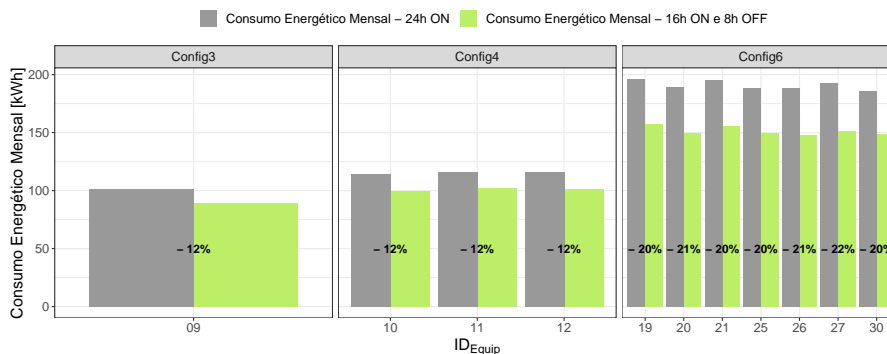


Figura 4.5: Poupança energética mensal devido a ativação da *feature* *Sleep* - *Ericsson*.

Ao desligar os equipamentos nos períodos de baixo tráfego é feito de forma automática na rede, os utilizadores são transferidos para outra célula 4G de forma a não afetar a qualidade do serviço. Outra solução, com menor impacto na rede para poupança energética consiste na *feature Micro Sleep* que desliga a portadora em intervalos de tempo de símbolo ao detetar baixo tráfego.

Dos equipamentos amostrados, 9 equipamentos (nº 04-12) contém períodos com a *feature Micro Sleep* ativa. A Figura 4.6 mostra que a partir do dia 12/09 a *feature* está ativada entre 40% a 60% do tempo no intervalo de 15 minutos, com isso a potência média reduziu de 145 W para 129 W, ou seja, cerca de 11%.

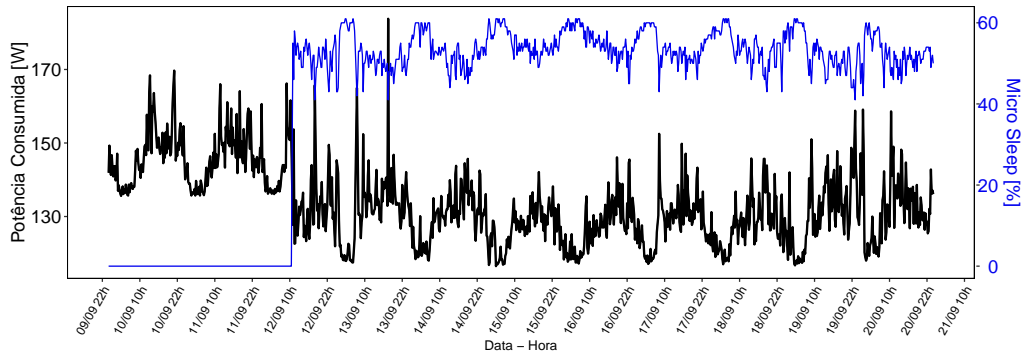


Figura 4.6: Potência consumida pelo equipamento nº 07 - *Ericsson*.

A Figura 4.7 mostra o *Biplot* do ACP, onde se verifica que as duas primeiras dimensões representam 72.5% dos dados, sendo 45.7% da variabilidade explicada na primeira dimensão e 26.8% na segunda dimensão. Os vetores indicam a direção da contribuição das variáveis independentes e a potência consumida. Verifica-se que as *features* de poupança são negativamente correlacionados com a potência, esta por sua vez está positivamente correlacionada com o tráfego.

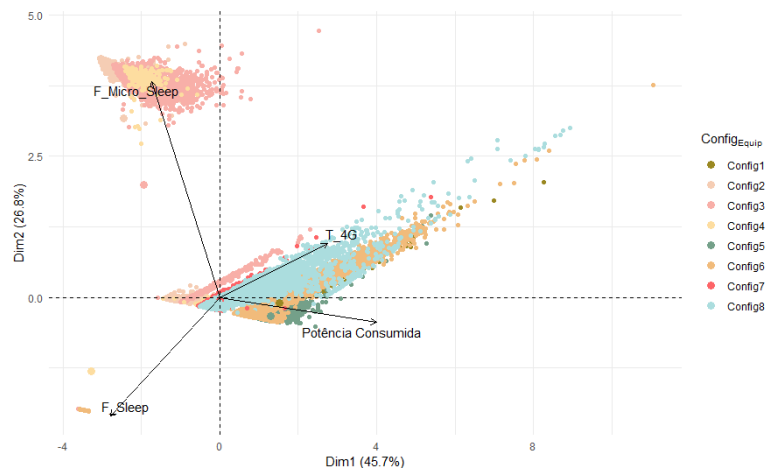


Figura 4.7: ACP das variáveis utilizadas no modelo para equipamentos de rádio 4G - *Ericsson*.

4.3.2 Equipamentos de Rádio 3G

A Tabela 4.4 apresenta as principais características dos equipamentos 3G, com os modelos de *hardware* distintos, a banda de frequência e o número de células. Os 45 equipamentos de rádio possuem 4 configurações de *hardware*, todas são do tipo RRU com a exceção da configuração 1 que é somente *Radio Unit* (RU), ou seja, com *hardware* dentro do *cabinet* e não junto à antena. O número de células está relacionado com a banda de frequência, sendo nos 900 MHz com 1 célula portadora, e em 2100 MHz com 2 células portadoras.

Tabela 4.4: Configurações dos equipamentos de rádio 3G - *Ericsson*.

	Modelo de <i>Hardware</i>	Banda de Frequência	Nº Células	Nº Equipamentos
Config. 1	RUS 01	900 MHz	1	18
Config. 2	RRUS 12	2100 MHz	2	12
Config. 3	RRUS 11	2100 MHz	2	3
Config. 4	RRUW	2100 MHz	2	12

A Figura 4.8 consiste no gráfico conjunto dos *boxplot* das medidas de consumo energético por equipamento, sendo que a mediana da potência consumida entre os equipamentos varia entre 125 W e 240 W. No geral, os equipamentos da configuração 1 (nº 01 a 18), possuem menor consumo mediano. Este pode estar associado ao facto da configuração operar com apenas uma célula ativa nos 900 MHz e os demais com 2 células portadoras nos 2100 MHz.

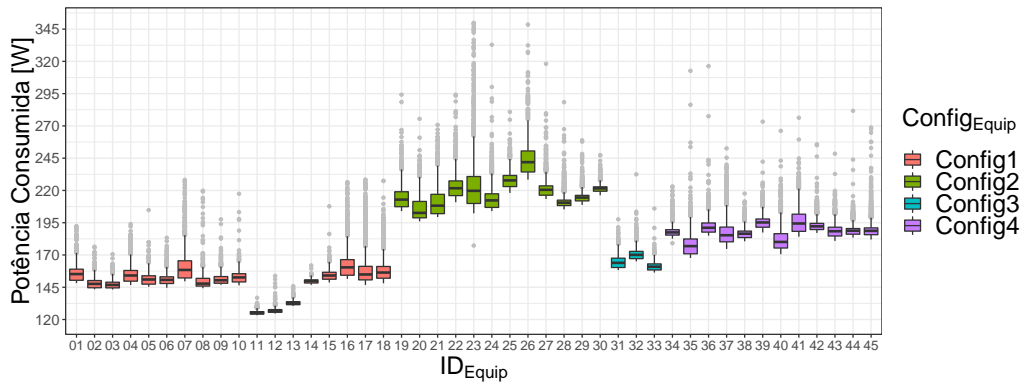


Figura 4.8: *Boxplot* da potência consumida pelos rádios 3G - *Ericsson*

Abaixo são apresentadas as variáveis explicativas consideradas no modelo para estimar o consumo energético, todas referentes ao tráfego gerado em *downlink* por período de 15 minutos:

- T_{CS_3G} : volume acumulado de tráfego de voz utilizando o somatório dos diferentes tipos de serviços 12.2 kbps, 64 kbps e *Adaptive Multi-Rate Wideband* (AMR-WB) a 16 kbps. Unidade: [Mb].
- T_{PS_R99} : volume acumulado de tráfego de dados com uma um débito binário máximo, nos diferentes serviços de 16 a 384 kbps, correspondendo à *Release 99*. Unidade: [Mb].
- T_{HSDPA} : volume acumulado de tráfego do canal de transporte em HSDPA, com *throughput* máximo de 14 Mbps. Unidade: [Mb].

A Tabela 4.5 apresenta as medidas estatísticas das variáveis independentes. Observa-se que há um maior volume de tráfego gerado HSDPA em relação aos demais tipos de tráfego.

Tabela 4.5: Medidas estatísticas descritivas referentes às variáveis independentes 3G - *Ericsson*.

	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
T_{CS_3G} [Mb]	0	0.44	5.46	16.92	240.40
T_{PS_R99} [Mb]	0	0.12	0.29	0.59	4.97
T_{HSDPA} [Mb]	0	4.48	45.39	232.41	4987.36

O resultado da análise ACP é ilustrado na Figura 4.9, onde as duas primeiras dimensões explicam 73.4% da variabilidade dos dados. Pode-se verificar que o tráfego de voz CS é predominante para as observações pertencentes à configuração 1 (900 MHz).

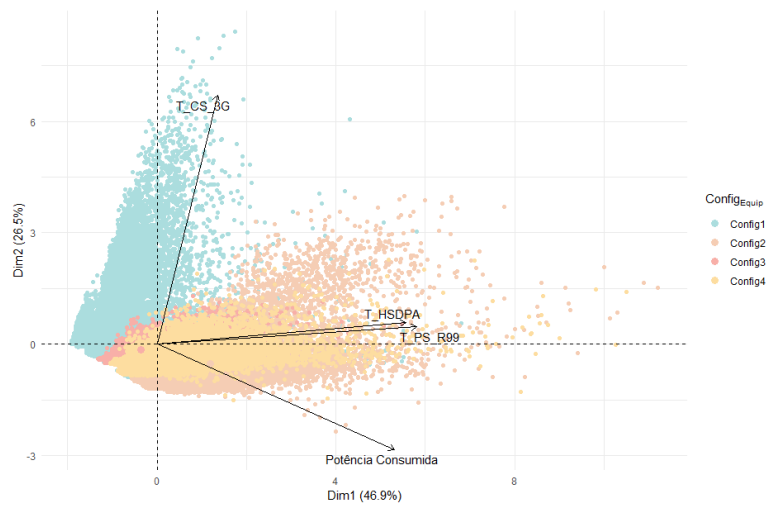
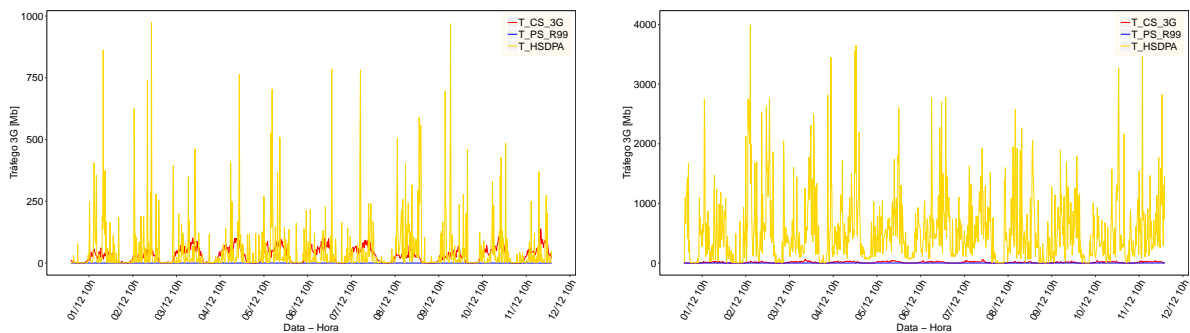


Figura 4.9: ACP das variáveis utilizadas no modelo para equipamentos de rádio 3G - *Ericsson*.

O volume de tráfego varia entre as diferentes células e ao longo do próprio dia. A Figura 4.10a mostra as variações de tráfego de uma célula na banda de 900 MHz, onde predomina o tráfego de voz, com o valor mediano de 29.33 Mb, o tráfego HSDPA com 6.17 Mb e o tráfego PS R99 com 0.08 Mb. Por outro lado, a Figura 4.10b apresenta o tráfego de uma célula na banda de 2100 MHz, com valor mediano HSDPA de 361.76 Mb, PS R99 de 0.67 Mb e o tráfego de voz com 9.38 Mb.



(a) Célula na banda dos 900 MHz

(b) Célula na banda dos 2100 MHz

Figura 4.10: Evolução do tráfego 3G nas células de 900 MHz e 2100 MHz.

4.3.3 Equipamentos de Rádio 2G

Os 15 equipamentos de rádio monitorizados em 2G encontram-se instalados dentro do *cabinet* e tem modelo de *hardware* "RUS 01". Operam na banda de frequência dos 900 MHz, com largura de banda de 5 MHz, e são compostos por 2 TRX por sector, sendo o primeiro para o BCCH e restantes canais de sinalização, e o segundo para canais de tráfego TCH. A *feature* de *Power Saving* do BCCH altera o limite de regulação da potência máxima de transmissão. Durante os períodos monitorizados, as células têm o parâmetro PRO constante em 6 dB. A potência máxima de transmissão configurada é de 43 dBm (20 W), reduzindo a *feature* esse valor para 37 dBm (5 W). A potência mediana dos equipamentos varia entre 147 W e 157 W, como verificado no *boxplot* da Figura 4.11.

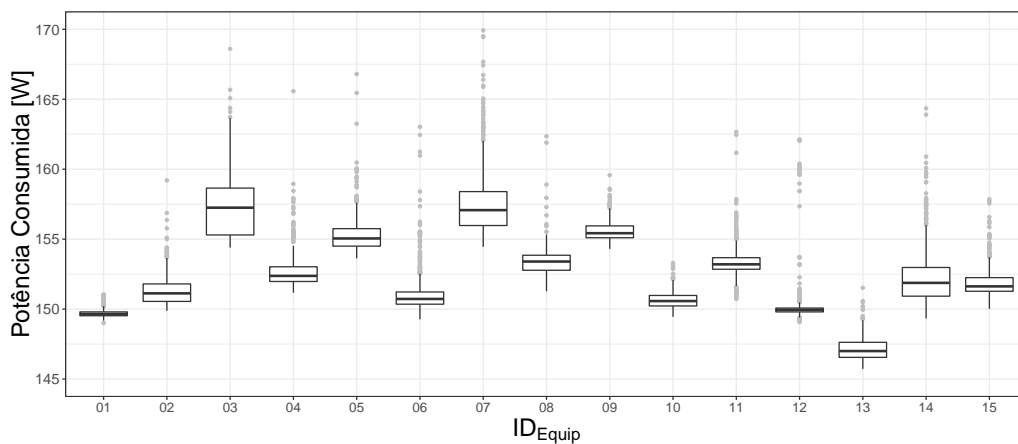


Figura 4.11: *Boxplot* da potência consumida pelos rádios 2G - *Ericsson*.

Os KPIs relacionados com o tráfego ao nível de voz e dados na tecnologia 2G, no período de 1 hora, são definidos por:

- T_{CS_2G} : volume de tráfego de TCHs alocado aos canais *full-rate* e *half-rate* ao nível da célula. Unidade: [Erl];
- T_{PS_2G} : volume total de dados EDGE e GPRS transportados no *downlink*. Unidade: [Mb].

As medidas que caracterizam os diferentes tipos de tráfego em 2G são apresentadas na Tabela 4.6. Pela análise ACP, na Figura 4.12, pode-se concluir que os volumes tráfego em CS e PS estão correlacionados positivamente com a potência.

Tabela 4.6: Medidas estatísticas descritivas referentes às variáveis independentes em 2G

	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
T_{CS_2G} [Erl]	0	0.005	0.08	0.40	4.63
T_{PS_2G} [Mb]	0	0.03	0.22	1.05	247.57

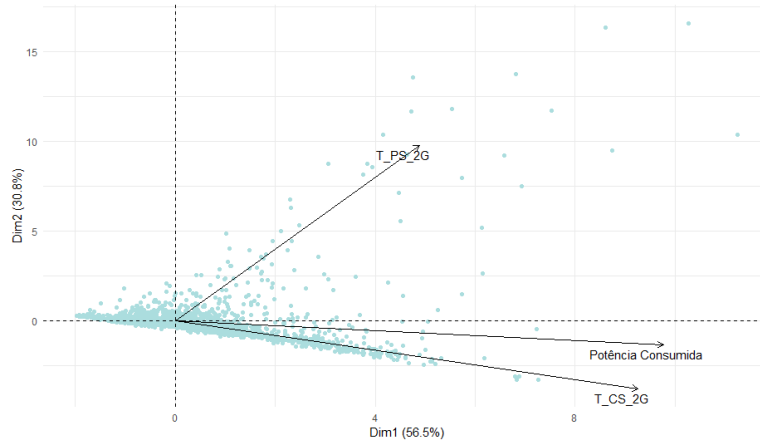


Figura 4.12: ACP das variáveis utilizadas no modelo para equipamentos de rádio 2G - *Ericsson*.

4.3.4 Equipamentos de Rádio Partilhado 2G/3G

Os rádios partilhados entre as tecnologias 2G e 3G, operam na banda de frequência de 900 MHz, e têm o modelo de *hardware* "RRUS 12". Na Figura 4.13, é possível verificar que a potência consumida nos 21 equipamentos amostrados varia entre 250 W e 280 W.

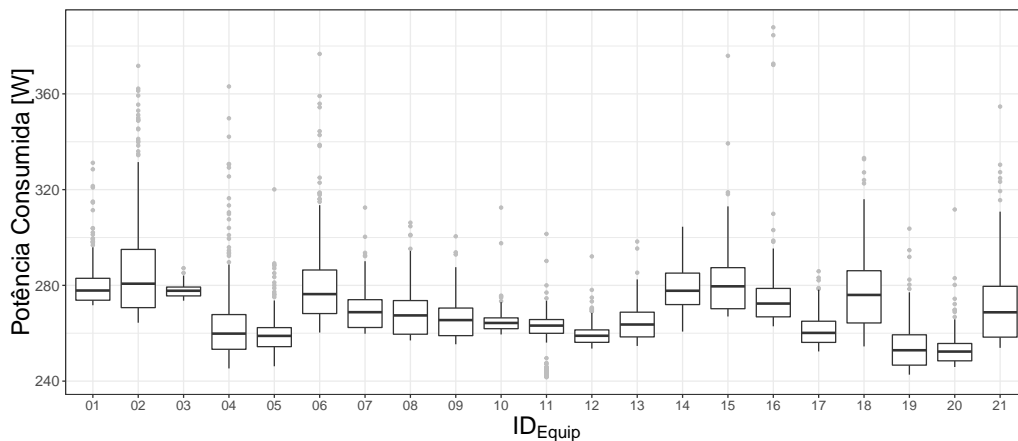


Figura 4.13: *Boxplot* da potência consumida pelos rádios partilhado 2G/3G - *Ericsson*.

Utilizam-se ambos os tráfegos em 2G e 3G, para explicar o consumo energético. A Tabela 4.7 apresenta as medidas descritivas variáveis independentes, verificando-se que o tráfego em HSDPA possui maior utilização mediana. Todas as variáveis independentes têm correlações positivas com a potência consumida, conforme verificado na análise ACP da Figura 4.14.

Tabela 4.7: Medidas estatísticas descritivas referentes às variáveis independentes 2G/3G - *Ericsson*.

	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
T_{CS_2G} [Erl]	0	0.03	0.31	0.86	6.48
T_{PS_2G} [Mb]	0	0.20	1.10	4.82	615.92
T_{CS_3G} [Mb]	0	1.60	13.24	37.44	332.15
T_{PS_R99} [Mb]	0	0.07	0.29	0.29	102.79
T_{HSDPA} [Mb]	0	1.64	20.43	119.3	8270.02

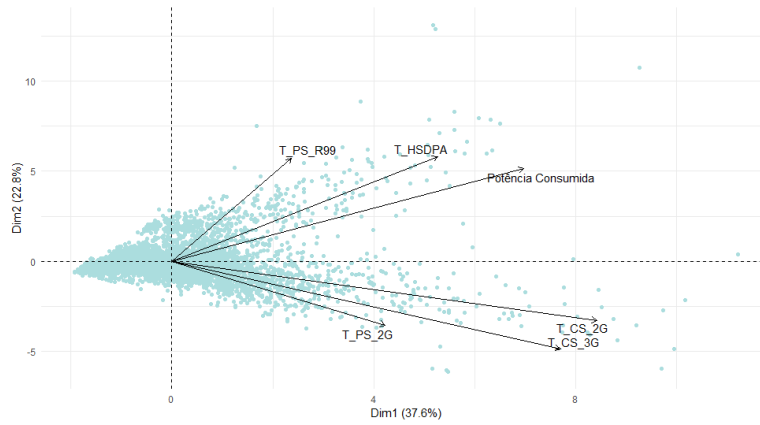


Figura 4.14: ACP das variáveis utilizadas no modelo para equipamentos de rádio partilhado 2G/3G - *Ericsson*.

4.3.5 Equipamentos de Banda Base

Foram monitorizados 17 equipamentos de banda base 4G. A Figura 4.15 apresenta as medidas de potência consumida em cada equipamento. Verifica-se que o valor mediano varia entre 106 W e 165 W, sendo a variabilidade das medidas é quase constante. Apesar de não seguir um padrão definido, optou-se por expressar as diferenças do consumo utilizando a capacidade máxima de processamento dos dados, em termos de número de células, por exemplo, uma BBU pode processar dados de 9 células, requerendo assim um equipamento de maior potência.

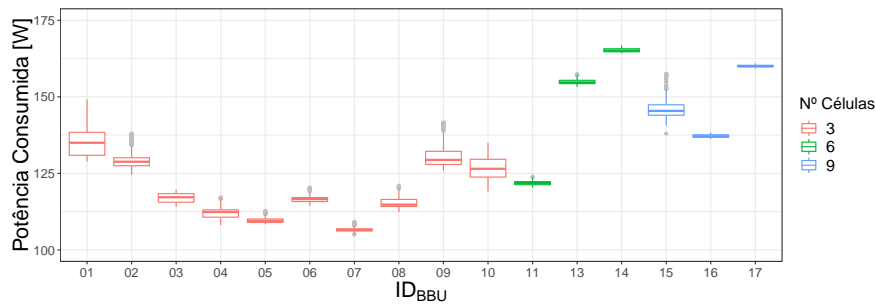


Figura 4.15: *Boxplot* da potência consumida pelas BBUs 4G - *Ericsson*.

Os equipamentos de banda base 3G, foram caracterizados de acordo com a capacidade máxima de *Channel Element* (CE), que podem ser: 384, 768 ou 1152. Na Figura 4.16 observa-se que as BBUs com maiores capacidade de CE tem maior consumo.

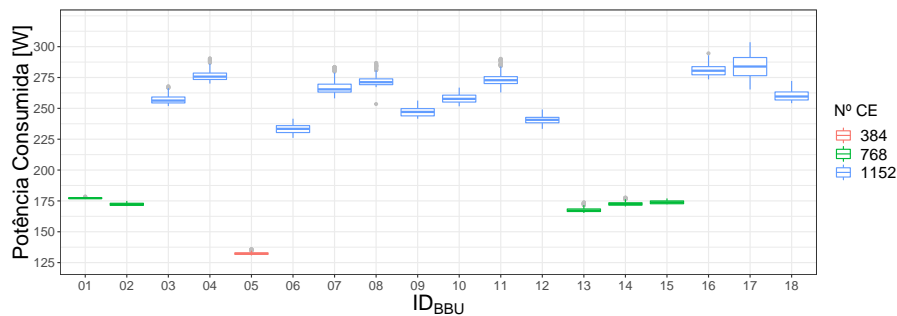


Figura 4.16: *Boxplot* da potência consumida pelas BBUs 3G - *Ericsson*.

As BBU's 2G tem potência média consumida muito semelhante, variando apenas entre 34 W e 45 W, conforme verificado na Figura 4.3.

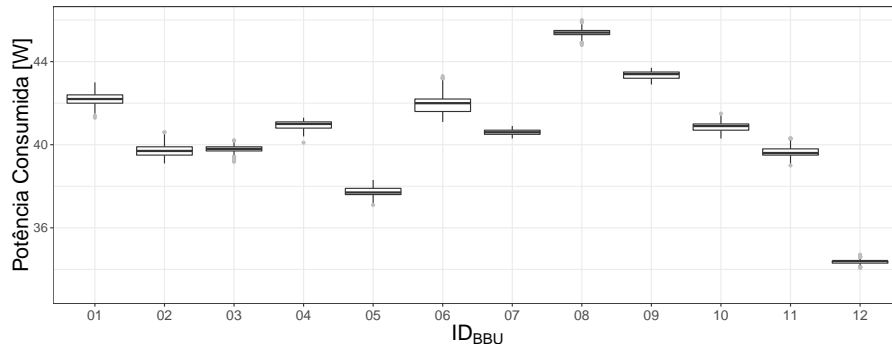


Figura 4.17: *Boxplot* da potência consumida pelas BBU's 2G - *Ericsson*.

4.4 *Vendor 2 - Huawei*

4.4.1 Equipamentos de Rádio 4G

Os 27 equipamentos monitorizados têm três configurações de *hardware* diferentes, todas com MIMO 2x2 e operam nas bandas de frequência dos 800 MHz e 1800 MHz, com largura de banda de 10 MHz e 20 MHz respectivamente, conforme apresentado na Tabela 4.8. A distribuição de potência consumida pelos equipamentos é demonstrada no gráfico *Boxplot* da Figura 4.18. Observa-se que os equipamentos da configuração 1 têm valores medianos mais próximos, em contraste com a maior variabilidade dos demais equipamentos.

Tabela 4.8: Configurações dos equipamentos de rádio utilizados no modelo 4G - *Huawei*.

	Modelo de <i>Hardware</i>	Banda de Frequência	Potência Máx. da Portadora	Nº Equipamentos
Config. 1	RRU 3268	800 MHz	2x40W	12
Config. 2	RRU 3928	1800 MHz	2x40W	3
Config. 3	RRU 3929	1800 MHz	2x40W	12

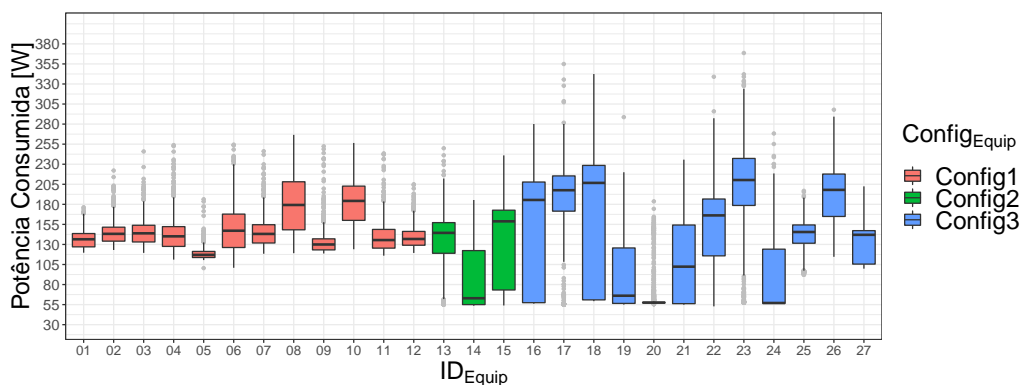


Figura 4.18: *Boxplot* da potência consumida pelos rádios 4G - *Huawei*

As medidas estatísticas das variáveis explicativas são apresentas na Tabela 4.9. Utiliza-se o volume de tráfego T_{4G} em Gb, no período de 15 minutos, a *feature Sleep Mode* como variável numérica, e a *feature Radio Frequency Channel Shutdown* (F_{RF_OFF}) que representa a percentagem de tempo em que alguns canais de rádio frequência da célula se encontram desligados, para poupança de energia.

Tabela 4.9: Medidas estatísticas descritivas referentes às variáveis independentes em 4G - *Huawei*.

	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
T_{4G} [Gb]	0	0.14	0.94	3.06	31.46
F_{RF_OFF} [%]	0	0	2.2	0	100
F_{Sleep} [%]	0	0	0	0	100

A Figura 4.19 mostra uma comparação da energia consumida nos períodos em que a *feature de Sleep* está ativa e inativa. Considerando o período de 16 horas do equipamento ligado e 8 horas em *stand by*, é possível alcançar uma poupança média mensal de cerca de 20%.

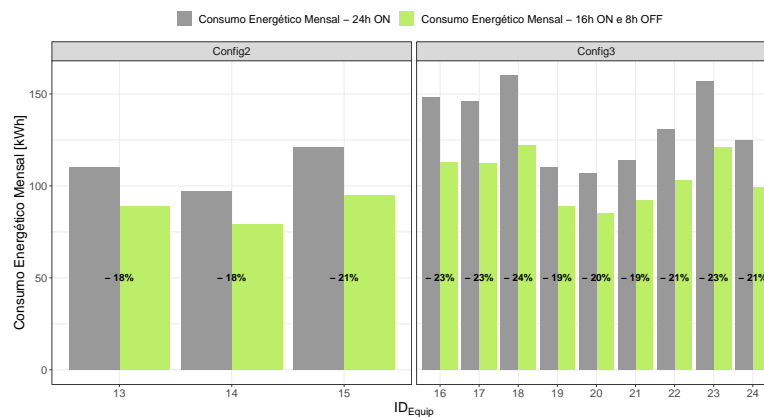


Figura 4.19: Poupança energética mensal com a ativação da *feature Sleep* - *Huawei*.

A *feature Radio Frequency Channel Shutdown* está ativada em todos os equipamentos que operam na banda dos 1800 MHz. A Figura 4.20 apresenta um exemplo do seu funcionamento, em que nos períodos da noite está 100% ativa, desligando assim automaticamente os amplificadores de potência ao emitir menos canais.

É possível concluir pela análise de ACP, ilustrada na Figura 4.21, que o tráfego se relaciona positivamente com a potência consumida, enquanto as *features* de poupança energética têm direções opostas, como era expectável.

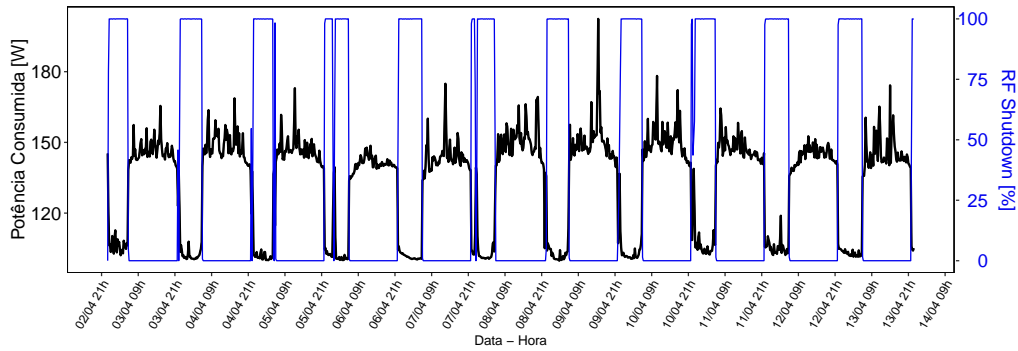


Figura 4.20: Potência consumida pelo equipamento n° 27 com a *feature Radio Frequency Channel Shutdown* ativa - *Huawei*.

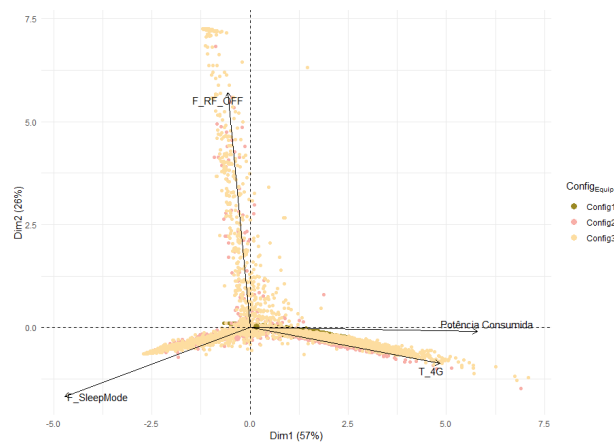


Figura 4.21: ACP das variáveis utilizadas no modelo de rádio 4G - *Huawei*.

4.4.2 Equipamentos de Rádio 3G

Foram amostrados 6 equipamentos de rádio 3G na banda de frequência 2100 MHz, e de modelo de *hardware* "RRU3928". Observa-se na Figura 4.22 que os equipamentos n° 02, 03 e 06 têm uma maior variabilidade do consumo de potência, devido à *feature Sleep* se encontrar ativa.

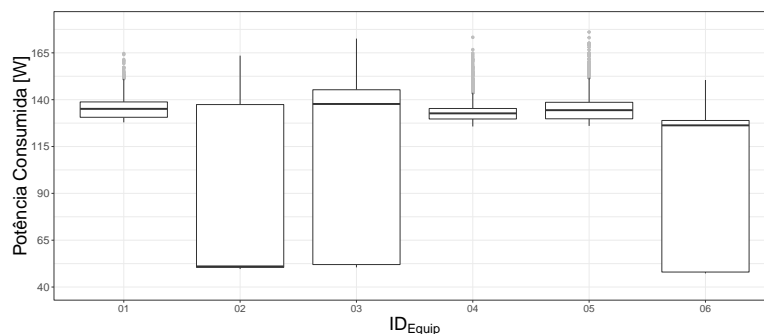


Figura 4.22: *Boxplot* da potência consumida dos equipamentos de rádio 3G - *Huawei*.

A Tabela 4.10 apresenta o resumo das medidas estatísticas das seguintes variáveis independentes: volume de tráfego em PS, T_{PS_R99} e T_{HSDPA} , ambos em Mb, volume de tráfego CS, T_{CS_3G} em Erl, e a *feature Sleep* (ON/OFF).

A análise de ACP, na Figura 4.23, confirma a correlação do consumo de potência dos equipamentos com as variáveis referentes ao volume de tráfego.

Tabela 4.10: Medidas estatísticas descritivas referentes às variáveis independentes em 3G - *Huawei*.

	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
T_{CS_3G} [Erl]	0	0	0.40	0	15.86
T_{PS_R99} [Mb]	0	0	0.06	0.18	14.70
T_{HSDPA} [Mb]	0	2.61	79.32	324.63	5636.15
F_{Sleep}	0="ON"	-	-	-	1="OFF"

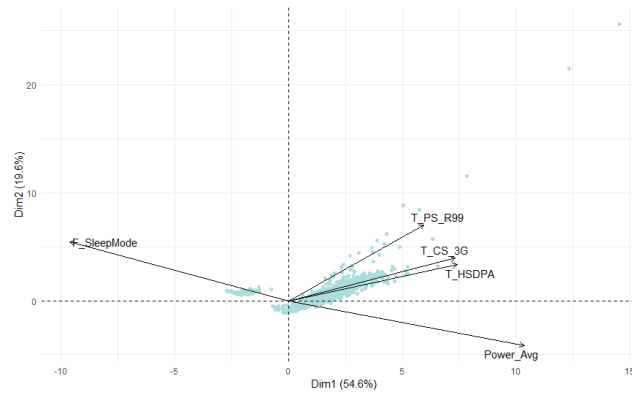


Figura 4.23: ACP das variáveis utilizadas no modelo de rádio 3G - *Huawei*.

4.4.3 Equipamentos de Rádio Partilhado 2G/3G

Os modelo de *hardware* dos 19 equipamentos amostrados de rádio partilhado são apresentados na Tabela 4.11. Pela análise do *boxplot* na Figura 4.24 é possível verificar que o consumo de potência mediano varia entre 160 W e 205 W.

Tabela 4.11: Configurações dos equipamentos de rádio utilizados no modelo partilhado 2G/3G - *Huawei*.

	Modelo de <i>Hardware</i>	Banda de Frequência	Nº Equipamentos
Config. 1	DBS 3900	900 MHz	2
Config. 2	RRU 3928	900 MHz	12
Config. 3	RRU 3929	900 MHz	5

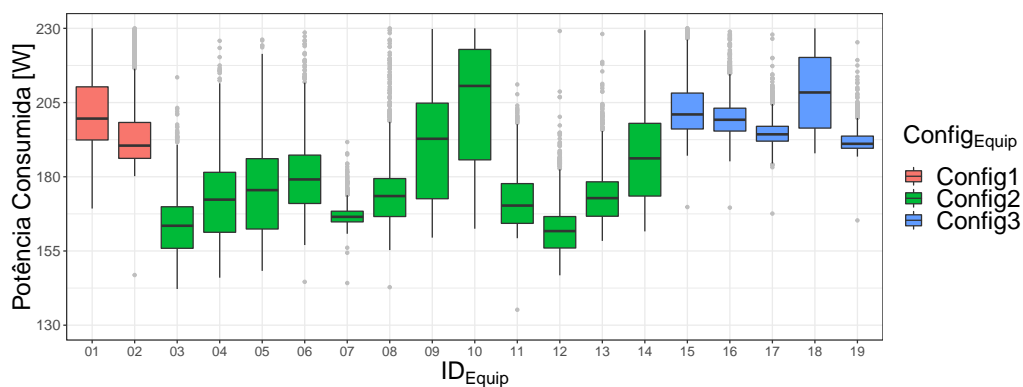


Figura 4.24: *Boxplot* da potência consumida nos equipamentos de rádio partilhado 2G/3G - *Huawei*.

As medidas estatísticas das variáveis independentes são apresentadas na Tabela 4.12. Utilizou-se

novamente o tráfego de voz 2G, dado por T_{CS_2G} (Erl) e o tráfego de dados, T_{PS_2G} (Mb). Além do tráfego em 3G, dado pelo volume de tráfego em PS, T_{PS_R99} e T_{HSDPA} , ambos em Mb, volume de tráfego CS, T_{CS_3G} em Erl. Pela análise de ACP na Figura 4.25, verifica-se que todas as variáveis que representam tráfego consideradas tem correlação positiva com a potência consumida.

Tabela 4.12: Medidas estatísticas descritivas referentes às variáveis independentes em 2G/3G - *Huawei*.

	Mínimo	Percentil 25	Mediana	Percentil 75	Máximo
T_{CS_2G} [Erl]	0	0.00	0.04	0.60	10.59
T_{PS_2G} [Mb]	0	2.55	4.96	9.81	194.07
T_{CS_3G} [Erl]	0	0.16	1.20	4.50	34.26
T_{PS_R99} [Mb]	0	0.07	0.13	0.27	52.22
T_{HSDPA} [Mb]	0	16.13	98.80	319.24	11118.75

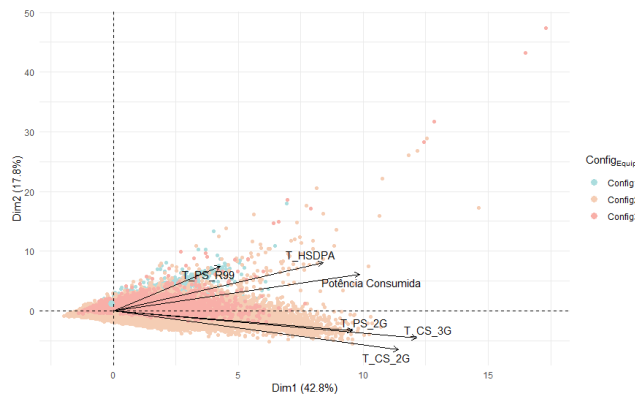


Figura 4.25: ACP das variáveis utilizadas no modelo de rádio partilhado 2G/3G - *Huawei*.

4.4.4 Equipamentos de Banda Base

Foram monitorizados 5 equipamentos de banda base 4G e 8 equipamentos 3G. A Figura 4.26 apresenta a potência consumida dos equipamentos por tecnologia. Devido à amostra ser reduzida, não foram encontradas características que permitissem explicar a variabilidade do consumo.

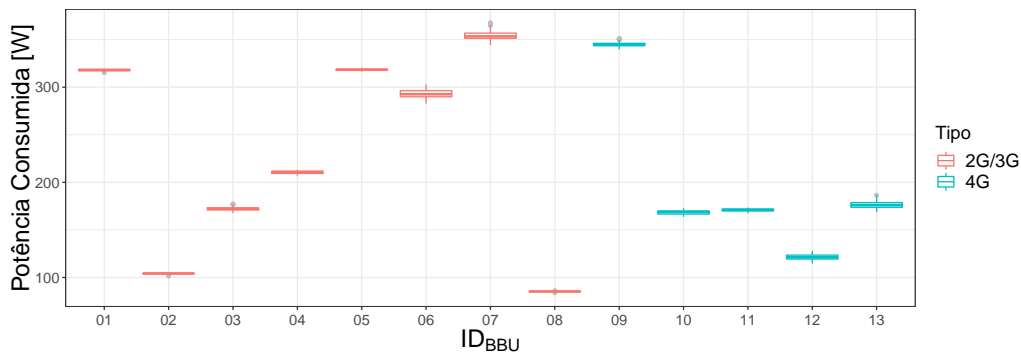


Figura 4.26: *Boxplot* da potência consumida pelas BBU's 3G e 4G - *Huawei*

5

Desenvolvimento dos Modelos de Consumo Energético

5.1	Vendor 1 - Ericsson	50
5.1.1	Modelo RRU 4G	50
5.1.2	Modelo RRU 3G	57
5.1.3	Modelo RRU 2G	59
5.1.4	Modelo RRU Partilhado 2G/3G	60
5.1.5	Modelo BBU	63
5.2	Vendor 2 - Huawei	64
5.2.1	Modelo RRU 4G	64
5.2.2	Modelo RRU 3G	66
5.2.3	Modelo RRU Partilhado 2G/3G	68
5.3	Estimativas e Métricas dos Modelos	70
5.4	Extrapolação do Consumo Energético	71

Neste capítulo é apresentada a construção e a evolução dos modelos de consumo energético para os equipamentos dos *vendors Ericsson e Huawei*. A partir do resultado obtido no modelo, é estimado o consumo energético anual em uma rede real. Por fim, a simulação de um cenário de *switch-off* das tecnologias 2G e/ou 3G.

5.1 Vendor 1 - Ericsson

Foram desenvolvidos quatro modelos de consumo energético para o rádio e três modelos de banda base. Nomeadamente, equipamentos de rádio 4G que operam na banda dos 800, 1800 e 2600 MHz, equipamentos 3G na banda dos 900 e 2100 MHz, equipamentos 2G na banda dos 900 MHz, e por fim os equipamentos de rádio partilhados entre 2G e 3G na banda dos 900 MHz. Para os equipamentos de banda base, foram desenvolvidos modelos para cada tecnologia.

5.1.1 Modelo RRU 4G

Duas abordagens foram consideradas para o desenvolvimento do modelo 4G. A primeira abordagem baseia-se na regressão linear de efeitos fixos, enquanto que a segunda utilizou-se o modelo de efeitos aleatórios.

Abordagem I: Modelo de regressão linear de efeitos fixos

O modelo da Equação 5.1 estima a potência consumida, (P_{4G}), considerando o volume de tráfego 4G (T_{4G}) em Gb. O coeficiente β_0 reflete o consumo *baseline* do equipamento, β_1 representa a alteração média da potência consumida por uma variação unitária do tráfego, e o termo do erro ϵ . A Figura 5.1 mostra a reta estimada pelo modelo, onde é possível verificar que a maioria das observações não se ajustam à reta.

$$P_{4G} = \beta_0 + \beta_1 T_{4G} + \epsilon \quad (5.1)$$

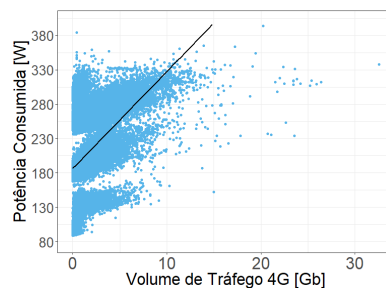


Figura 5.1: Reta de regressão linear ajustada do modelo 4G - *Ericsson*.

A Figura 5.2 apresenta a análise dos resíduos, onde se verifica que não se cumpre o pressuposto da homocedasticidade, pois os pontos não estão aleatoriamente distribuídos em torno de zero, apresentando alguma tendência. Verifica-se também a falta de independência dos resíduos, com forte autocorrelação entre os resíduos.

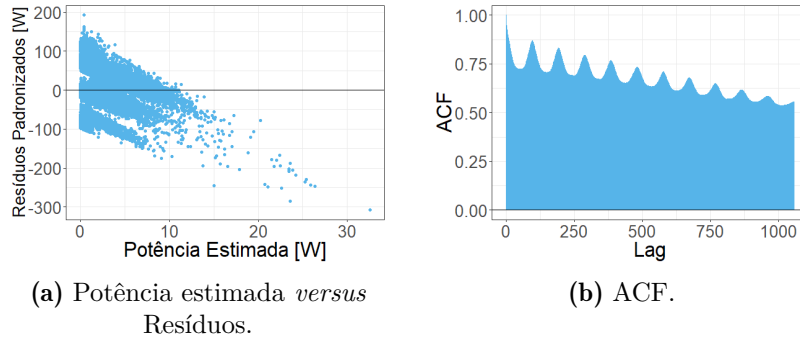


Figura 5.2: Análise residual do modelo 4G com regressão linear de efeitos fixos - *Ericsson*.

O modelo da Equação 5.1 apresenta o valor de 0.20 para o coeficiente de determinação ajustado, e desvio padrão do erro $\hat{\sigma}_e = 44.62$ W. Tendo em conta a análise apresentada, é possível que faltem variáveis preditoras importantes, ou mesmo que os equipamentos amostrados possam diferir no valor do consumo *baseline*. A fim de verificar se o consumo *baseline* difere entre os equipamentos, um segundo modelo de regressão linear é analisado, dado na Equação 5.2, onde utiliza, além do volume de tráfego, a variável nominal (ID_{Equip}) com a identificação única dos equipamentos amostrados.

$$P_{4G} = \beta_0 + \beta_1 T_{4G} + \beta_2 ID_{Equip} + \epsilon \quad (5.2)$$

Na regressão linear as variáveis nominais são consideradas como variáveis fictícias, e para estimar o coeficiente β_2 utilizam-se diferenças relativas a partir de um grupo de referência. O ajustamento do modelo 5.2 teve um aumento da variabilidade explicada, tendo-se obtido $R_a^2 = 0.84$, e uma redução do desvio padrão do erro para $\hat{\sigma}_e = 20.26$ W, quando comparado com o modelo 5.1.

Na tentativa de melhorar a qualidade do modelo, são usadas outras variáveis nominais com importância na explicação do consumo energético, e nesse sentido foram ajustados diversos modelos, cujos resultados apresentados na Tabela 5.1. As variáveis utilizadas nos modelos foram: configuração do equipamento (*Config*: Configuração 1 a 8), largura de banda (*BW*: 10 MHz e 20 MHz) e banda de frequências (*Banda_{Freq}*: 800 MHz, 1800 MHz e 2600 MHz).

Tabela 5.1: Comparação entre modelos de regressão linear de efeitos fixos 4G - *Ericsson*.

Modelo	Efeito Fixo	R_a^2	$\hat{\sigma}_e$ [W]
M1	$\beta_0 + \beta_1 T_{4G}$	0.20	44.62
M2	$\beta_0 + \beta_1 T_{4G} + \beta_2 ID_{Equip}$	0.84	20.26
M3	$\beta_0 + \beta_1 T_{4G} + \beta_2 Config$	0.79	23.55
M4	$\beta_0 + \beta_1 T_{4G} + \beta_2 BW$	0.31	42.58
M5	$\beta_0 + \beta_1 T_{4G} + \beta_2 Banda_{Freq}$	0.34	41.64

O modelo 5.2 (M2) continua a ser o melhor, com o menor o erro residual. De seguida, o modelo M3 com a variável nominal que representa as configurações dos equipamentos. No entanto, não é de interesse que o modelo compare equipamentos ou configurações específicas. Deve-se considerar um modelo que seja capaz de generalizar para novos equipamentos rádio e em simultâneo controlar a variabilidade proveniente da amostra.

Numa outra abordagem o modelo da Equação 5.1 é ajustado utilizando as observações de cada equipamento em separado. A Figura 5.3a apresenta os resultados dos coeficientes β_0 e β_1 estimados em

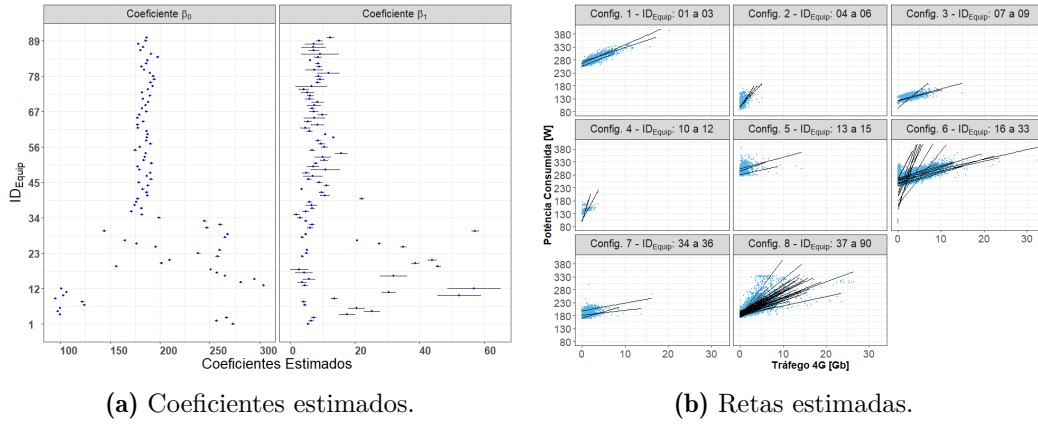


Figura 5.3: Estimativa dos coeficientes de regressão linear de efeitos fixos ajustados por equipamento no modelo 4G - *Ericsson*.

cada equipamento. Verifica-se grande variabilidade nas estimativas de β_0 nos primeiros trinta e três equipamentos, depois estabiliza-se em torno de 180 W. Os coeficientes β_1 tem valor médio estimado de 11.90 W/Gb. A Figura 5.3b mostra a reta estimada, onde se observa que os equipamentos da mesma configuração possuem *intercepts* (β_0) mais próximos. No entanto, existem muitas desvantagens nesta abordagem, pois não é facilmente interpretada quando há muitos grupos, e normalmente não permitem a generalização, ou seja, ignoram as semelhanças que possam existir entre os grupos.

Abordagem II: Modelo de regressão linear de efeitos mistos

Para obter melhores estimativas será testado o modelo de regressão linear de efeitos mistos. Os modelos mistos adaptam-se à situação em que não há controle sobre todas as variáveis e os dados são estruturados. Por exemplo, não é possível ter o controle das condições ambientais a que os equipamentos estão submetidos, nem mesmo medir as perdas de dissipação de energia dos equipamentos, por efeito *Joule*. Assim como também é possível existirem correntes induzidas pela proximidade dos cabos dentro do *cabinet*, entre outros fatores que possam aumentar a incerteza da medição.

No modelo da Equação 5.3, o efeito fixo é representado pelo termo com os coeficientes β_0 e β_1 , além disso considera o efeito aleatório devido às diferenças do consumo *baseline* dos equipamentos, existindo ainda a parcela correspondente ao erro aleatório. Sendo i o número de equipamentos, assume-se a distribuição dos resíduos $\epsilon_i \sim N(\mathbf{0}, \mathbf{R}_i)$ (Normal com vector médio nulo e matriz de variância-covariância $\mathbf{R}_i = \sigma_e^2 \mathbf{I}$) e a distribuição do efeito aleatório $b_i \sim N(\mathbf{0}, \mathbf{G})$ (Normal com vector médio nulo e matriz de variância-covariância $\mathbf{G} = \sigma_0^2 \mathbf{I}$).

$$\mathbf{P}_{(4G)ij} = \underbrace{\beta_0 + \beta_1 \mathbf{T}_{4G}}_{\text{Efeito Fixo}} + \underbrace{b_{0i}}_{\text{Efeito Aleatório}} + \underbrace{\epsilon_i}_{\text{Erro}}, \text{ com } i = 1, \dots, 90 \quad (5.3)$$

Relativamente aos resultados obtidos no modelo acima, verificou-se que a estimativa do desvio padrão do erro é de $\hat{\sigma}_e = 20.26$ W, e a estimativa do efeito aleatório $\hat{\sigma}_0 = 41.55$ W, representando 67.22% da variabilidade explicada pelo efeito aleatório. O resíduo deste modelo é inferior em relação ao erro do modelo da Equação 5.1 de efeitos fixos, e semelhante ao da Equação 5.2. A vantagem no modelo misto é que permite a generalização, pois estima os coeficientes contabilizando as influências dos efeitos aleatórios. Pode-se concluir que apenas os efeitos fixos não são suficientes para explicar

a variabilidade do consumo de potência, sendo o efeito aleatório devido às diferenças de consumo *baseline* entre equipamentos destinado a descrever uma parcela do erro.

É natural que equipamentos com a mesma configuração tenham valores de consumo *baseline* mais próximos. O modelo 5.4 considera as variações aleatórias dos *intercepts* em dois níveis: nível 1, das configurações, e nível 2, dos equipamentos aninhados nas configurações. Assume-se a distribuição dos resíduos $\epsilon_{ij} \sim N(\mathbf{0}, \mathbf{R}_{ij})$ (Normal com vector médio nulo e matriz de variância-covariância $\mathbf{R}_{ij} = \sigma_e^2 \mathbf{I}$), sendo $i = 8$ configurações e $j = 90$ equipamentos. A distribuição do efeito aleatório correspondente ao nível 1 é dado por $\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G}_1)$ (Normal com vector médio nulo e matriz de variância-covariância $\mathbf{G}_1 = \sigma_{0(Nível1)}^2 \mathbf{I}$), e ao nível 2 dado por $\mathbf{b}_{ij} \sim N(\mathbf{0}, \mathbf{G}_2 = \sigma_{0(Nível2)}^2 \mathbf{I})$.

$$\mathbf{P}_{(4G)ij} = \underbrace{\beta_0 + \beta_1 \mathbf{T}_{4G}}_{\text{Efeito Fixo}} + \underbrace{b_{0i}}_{\substack{\text{Efeito Aleatório} \\ \text{Nível 1}}} + \underbrace{b_{0ij}}_{\substack{\text{Efeito Aleatório} \\ \text{Nível 2}}} + \underbrace{\epsilon_{ij}}_{\text{Erro}}, \text{ com } i = 1, \dots, 8; j = 1, \dots, 90 \quad (5.4)$$

O resultado do modelo definido na Equação 5.4 apresenta desvio padrão para o nível 2 de $\hat{\sigma}_{0(Nível2)} = 12.58$ W, e para o nível 1 $\hat{\sigma}_{0(Nível1)} = 73.15$ W e desvio padrão do erro $\hat{\sigma}_e = 20.26$ W. Assim, as diferenças do consumo *baseline* entre configurações explicam 69% da variabilidade do consumo de potência, e apenas 11% está associado às diferenças entre equipamentos.

Este modelo é comparado com o anterior através do teste TRV, a fim de verificar se a estrutura com efeitos aleatórios em dois níveis é a mais indicada. O resultado do teste é apresentado na Tabela 5.2, onde se verifica com base no valor-p, a rejeição da hipótese nula, portanto, a diferença entre os modelos é estatisticamente significativa.

Tabela 5.2: Resumo do critério de seleção do efeito aleatório de agrupamento para o modelo 4G - *Ericsson*.

Modelo	Efeito Aleatório	g.l.	AIC	TRV	Teste	Hipótese Nula	Valor-p
M1	b_{0i}	4	842386				
M2	$b_{0i} + b_{0ij}$	5	842210	177.66	1 vs 2	$b_{0ij} = 0$	<0.001

Uma vez decidido a estrutura do efeito aleatório, é possível melhorar o modelo adicionando variáveis explicativas de efeito fixo. Para tal, estimam-se diferentes modelos, onde o método de estimação MLE é utilizado para que os modelos possam ser comparados. A Tabela 5.3 apresenta o resultado do TRV, onde se inicia com o modelo apenas composto pelo *intercept* (β_0), e depois se adiciona cada uma das variáveis para testar a sua significância. Além do tráfego, as variáveis testadas foram as *features Sleep Mode* (\mathbf{F}_{Sleep}) e *Micro Sleep* ($\mathbf{F}_{MicroSleep}$). Verifica-se que ambas as variáveis contribuem de forma significativa para a capacidade preditiva do modelo.

É possível que os equipamentos, além de apresentarem diferenças no consumo *baseline*, possam também diferir em algum dos coeficientes de inclinação. Na Tabela 5.4 é apresentado o resultado do teste ao adicionar o efeito de inclinação aleatória para cada preditor, em ambos os níveis. Verifica-se que o efeito é significativo em todos os modelos, sendo escolhido o modelo M10, com menor valor de AIC.

Tabela 5.3: Resumo do critério de seleção do efeito fixo para o modelo 4G - *Ericsson*.

Modelo	Efeito Fixo	g.l.	AIC	TRV	Teste	Hipótese Nula	Valor-p
M3	β_0	4	872316				
M4	$\beta_0 + \beta_1 \mathbf{T}_{4G}$	5	842214	30103.79	3 vs 4	$\beta_1 = 0$	< 0.001
M5	$\beta_0 + \beta_1 \mathbf{T}_{4G} + \beta_2 \mathbf{F}_{MicroSleep}$	6	722368	27803.8	4 vs 5	$\beta_2 = 0$	< 0.001
M6	$\beta_0 + \beta_1 \mathbf{T}_{4G} + \beta_2 \mathbf{F}_{MicroSleep} + \beta_3 \mathbf{F}_{Sleep}$	7	694566	119848.42	5 vs 6	$\beta_3 = 0$	< 0.001

Tabela 5.4: Resumo do critério de seleção do efeito aleatório para o modelo 4G - *Ericsson*.

Modelo	Efeito Aleatório	g.l.	AIC	TRV	Teste	Hipótese Nula	Valor-p
M7	$b_{0i} + b_{0ij}$	7	694565				
M8	$b_{0i} + b_{0ij} + b_{1i} \mathbf{T}_{4G} + b_{1ij} \mathbf{T}_{4G}$	11	647061	47512.03	7 vs 8	$b_{1i} = b_{1ij} = 0$	<0.001
M9	$b_{0i} + b_{0ij} + b_{1i} \mathbf{F}_{MicroSleep} + b_{1ij} \mathbf{F}_{MicroSleep}$	11	646705	47867.75	7 vs 9	$b_{1i} = b_{1ij} = 0$	<0.001
M10	$b_{0i} + b_{0ij} + b_{1i} \mathbf{F}_{Sleep} + b_{1ij} \mathbf{F}_{Sleep}$	11	640401	54171.9	7 vs 10	$b_{1i} = b_{1ij} = 0$	<0.001

A Figura 5.4 apresenta a variabilidade dos desvios dos *intercepts* estimados no nível 1. Observa-se que as configurações com o mesmo modelo de *hardware* têm desvios próximos, exceto a Configuração 1 e 2, pois a primeira utiliza dois equipamentos individuais para operar em MIMO 4x4. Portanto, a diferença de potência *baseline* está relacionada com as configurações dos equipamentos amostrados, e a variância dos efeitos aleatórios pode ser considerada inexplicada, no sentido em que as causas detalhadas da variação são desconhecidas. Assim sendo, dentre as variáveis disponíveis não é possível encontrar outros fatores que possam ser incluídos no modelo.

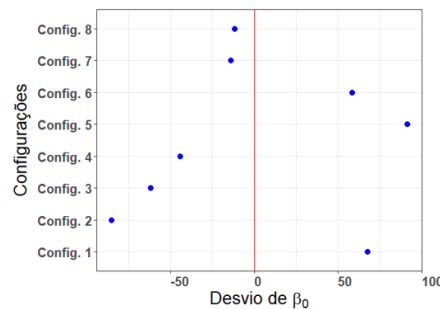


Figura 5.4: Desvios dos coeficientes estimados do modelo 4G - *Ericsson*.

Validação do Modelo

De acordo com a análise da Figura 5.5, constata-se que os resíduos estão distribuídos aleatoriamente com média zero. O facto dos resíduos não estarem normalmente distribuídos não é considerado grave.

Na Figura 5.6a pode ser verificado o pressuposto de linearidade, onde se constata que a potência estimada tem uma tendência linear. Observa-se a falta de independência dos resíduos, conforme demonstra a Figura 5.6b, com autocorrelação significativa entre os mesmos e decaimento lento, além de correlações positivas e negativas com periodicidade.

A fim de corrigir a falta de independência dos resíduos, a abordagem seguida consiste em utilizar componentes auto-regressivas ou funções periódicas no tempo. Embora o interesse do modelo não seja a evolução temporal do consumo de potência, o que é facto é que se observa a dependência dos resíduos com um padrão periódico. Uma possível razão pode ser a influência das oscilações da temperatura ambiente à qual os equipamentos estão sujeitos.

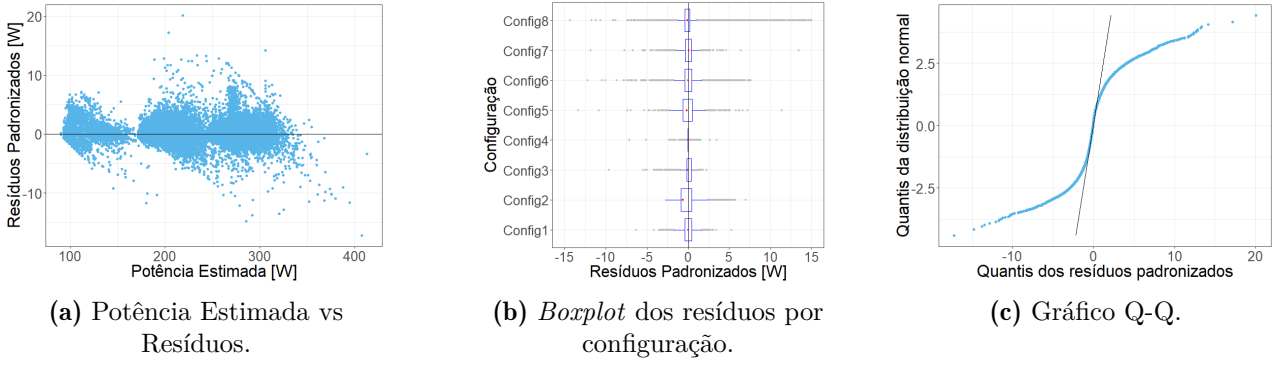


Figura 5.5: Análise residual para verificação da homocedasticidade e normalidade dos erros no modelo 4G - *Ericsson*.

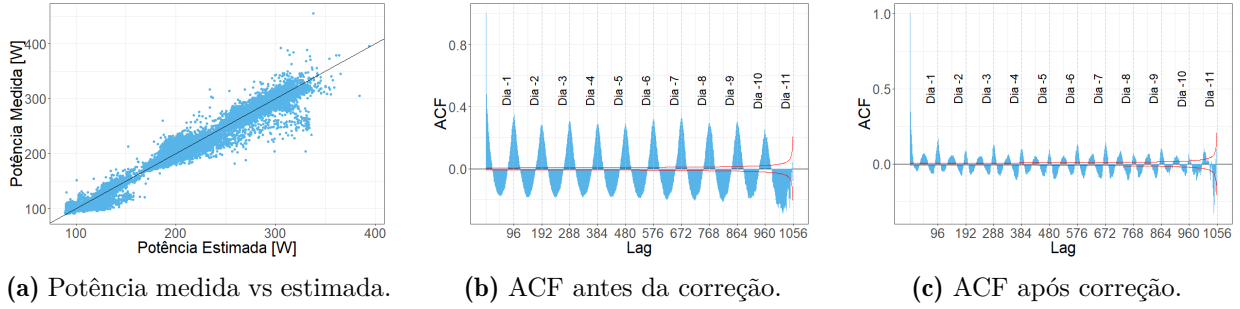


Figura 5.6: Análise residual para verificação da linearidade e independência no modelo 4G - *Ericsson*.

É acrescentado ao modelo o efeito aleatório de uma função de *seno* e *coosseno* em cada intervalo de tempo amostrado, dado por \mathbf{H}_{rad} . Como as observações foram obtidas a cada 15 minutos, 1 dia possui 96 observações; portanto, ao converter em radianos, tem-se um intervalo entre 0 e 6.28 rad. O resultado da ACF após esta correção é mostrado na Figura 5.6c. Observa-se que esta transformação reduziu a amplitude da auto-correlação, contudo ainda há decaimento lento nos primeiros atrasos. Uma solução possível, seria utilizar o processo auto-regressivo AR(1), mas devido à dimensão da amostra, o modelo teve um alto custo computacional e não apresentou ganhos significativos.

Modelo Final

As estimativas do efeito fixo para o modelo final descrito na Equação 5.5 são apresentadas na Tabela 5.5. O aumento da potência consumida pelos equipamentos está diretamente relacionado com o tráfego, com uma taxa de 6.10 W/Gb. Em relação às *features* de poupança, tem-se uma redução de 97 W no modo de *Sleep*, e uma redução de, no máximo, 25.63 W quando utiliza o *Micro Sleep*.

$$\begin{aligned}
 P_{(4G)ij} &= \beta_0 + \beta_1 \mathbf{T}_{4G} + \beta_2 \mathbf{F}_{MicroSleep} + \beta_3 \mathbf{F}_{Sleep} + b_{0i} + \\
 &b_{0ij} + b_{1i} \mathbf{F}_{Sleep} + b_{1ij} \mathbf{F}_{Sleep} + b_{2ij} \sin \mathbf{H}_{rad} + b_{3ij} \cos \mathbf{H}_{rad} + \epsilon_{ij} \\
 &\text{com } i = 1, \dots, 8 \text{ configurações ; } j = 1, \dots, 90 \text{ equipamentos} \\
 &b_i \sim N(\mathbf{0}, \mathbf{G}_1); b_{ij} \sim N(\mathbf{0}, \mathbf{G}_2); \epsilon_{ij} \sim N(\mathbf{0}, \sigma_e^2 \mathbf{I})
 \end{aligned} \tag{5.5}$$

Em 5.6 apresentam-se as estimativas de erro, dos elementos de variância e covariância da matriz estimada \mathbf{G}_1 de efeito aleatório. Em 5.7 apresenta-se a matriz estimada \mathbf{G}_2 de efeito aleatório dos equipamentos (nível 2).

Tabela 5.5: Estimativa dos intervalos de confiança a 95% para os parâmetros de efeito fixo do modelo 4G - *Ericsson*.

		Estimativa	IC (95%)
		Pontual	
(Intercept) [W]	$\hat{\beta}_0$	194.816	[149.313,240.319]
T_{4G} [Gb]	$\hat{\beta}_1$	6.110	[6.073,6.147]
$F_{MicroSleep}$ [%]	$\hat{\beta}_2$	-25.634	[-26.310,-24.958]
F_{Sleep} [ON/OFF]	$\hat{\beta}_3$	-96.152	[-139.402,-52.902]

$$\hat{\sigma}_e^2 = 6.14W \quad ; \quad \hat{\mathbf{G}}_1 = \begin{bmatrix} \hat{\sigma}_0^2 & \hat{\rho}\hat{\sigma}_0\hat{\sigma}_1 \\ \hat{\rho}\hat{\sigma}_0\hat{\sigma}_1 & \hat{\sigma}_1^2 \end{bmatrix} = \begin{bmatrix} (65.71)^2 & -0.99 \\ -0.99 & (63.62)^2 \end{bmatrix} \quad (5.6)$$

$$\hat{\mathbf{G}}_2 = \begin{bmatrix} \hat{\sigma}_0^2 & \hat{\rho}\hat{\sigma}_0\hat{\sigma}_1 & \hat{\rho}\hat{\sigma}_0\hat{\sigma}_2 & \hat{\rho}\hat{\sigma}_0\hat{\sigma}_3 \\ \hat{\rho}\hat{\sigma}_1\hat{\sigma}_0 & \hat{\sigma}_1^2 & \hat{\rho}\hat{\sigma}_1\hat{\sigma}_2 & \hat{\rho}\hat{\sigma}_1\hat{\sigma}_3 \\ \hat{\rho}\hat{\sigma}_2\hat{\sigma}_0 & \hat{\rho}\hat{\sigma}_2\hat{\sigma}_1 & \hat{\sigma}_2^2 & \hat{\rho}\hat{\sigma}_2\hat{\sigma}_3 \\ \hat{\rho}\hat{\sigma}_3\hat{\sigma}_0 & \hat{\rho}\hat{\sigma}_3\hat{\sigma}_1 & \hat{\rho}\hat{\sigma}_3\hat{\sigma}_2 & \hat{\sigma}_3^2 \end{bmatrix} = \begin{bmatrix} (7.30)^2 & -0.53 & 0.04 & -0.52 \\ -0.53 & (1.83)^2 & 0.25 & 0.55 \\ 0.04 & 0.25 & (2.83)^2 & -0.18 \\ -0.52 & 0.55 & -0.18 & (3.02)^2 \end{bmatrix} \quad (5.7)$$

Na Figura 5.7 podemos observar a evolução da estimativa da potência consumida, resultante do modelo final, ao longo de 9 dias de teste, para um equipamento que opera na banda de 1800 MHz. Observa-se que a partir do dia 24/09 a *feature micro sleep* é desativada, e a potência média aumenta de 129 W para 145 W, ou seja, cerca de 11%. A Figura 5.8, mostra a aplicação do modelo a um equipamento na banda de 1800 MHz, com períodos com a *feature sleep mode* ativa, sendo que nesses períodos a potência reduz cerca de 150 W.

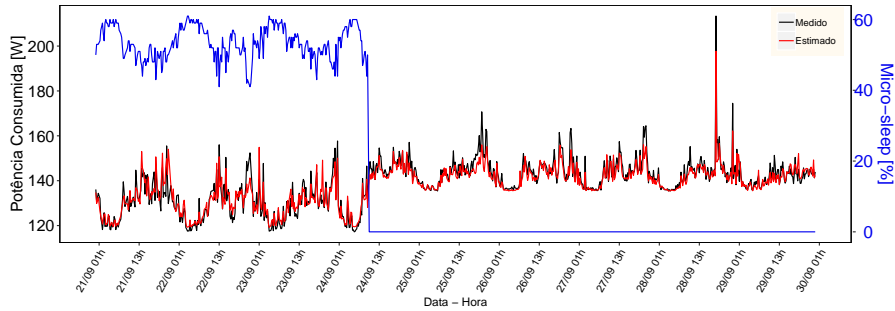


Figura 5.7: Aplicação do modelo 4G ao equipamento nº 07 da configuração 3 (1800 MHz) - *Ericsson*

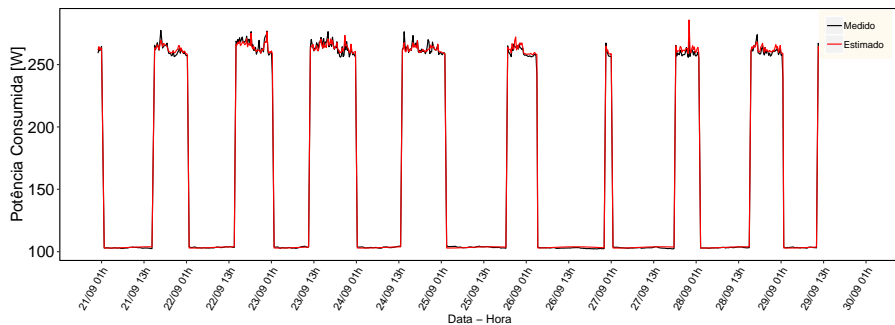


Figura 5.8: Aplicação do modelo 4G ao equipamento nº 30 da configuração 6 (1800 MHz) - *Ericsson*

Por fim, para verificar se a adição dos efeitos aleatórios se traduz em melhores resultados, realizou-se uma comparação dos modelos de efeitos fixos e de efeitos mistos, cujo resultado se encontra apresentado na Tabela 5.6. Dado que é rejeitada a hipótese de nulidade dos coeficientes correspondentes aos efeitos aleatórios, concluímos pela sua significância estatística, e portanto pela vantagem na adição dos efeitos aleatórios.

Tabela 5.6: Comparação entre efeitos fixos e efeitos misto no modelo 4G - *Ericsson*.

	RMSE [W]	MAPE [%]	R_a^2	g.l.	AIC	TRV	Hipótese Nula	Valor-p
Modelo de Efeitos Fixos	34.03	12.62	0.55	7	940143			
Modelo de Efeitos Mistos	6.14	1.65	0.99	20	616462	323707.9	$b_{ij} = b_i = 0$	< 0.001

5.1.2 Modelo RRU 3G

Para o modelo 3G utilizam-se como variáveis independentes ambos os tráfegos PS, T_{PS_R99} e T_{HSDPA} (Mb), bem como o tráfego de voz CS, representado por T_{CS_3G} (Erl). Foram estimados modelos considerando o efeito aleatório para a variável nominal associada à banda de frequências (900 MHz e 2100 MHz) e à configuração do *hardware* (Configurações de 1 a 4). Aplicando o TRV, foi selecionado o modelo com efeito ao nível 1 das configurações, e para o nível 2 dos equipamentos. Relativamente à escolha do efeito aleatório de inclinação, é considerada a variável T_{HSDPA} apenas no nível 2. À semelhança do modelo 4G, utiliza-se a função de *seno* e *coseno* na tentativa de remover a periodicidade dos resíduos.

$$P_{(3G)ij} = \beta_0 + \beta_1 T_{CS_3G} + \beta_2 T_{PS_R99} + \beta_3 T_{HSDPA} + b_{0i} + b_{0ij} + b_{1i} T_{HSDPA} + b_{2ij} \sin(\mathbf{H}_{rad}) + b_{3ij} \cos(\mathbf{H}_{rad}) + \epsilon_{ij} \quad (5.8)$$

com $i = 1, \dots, 4$ configurações ; $j = 1, \dots, 45$ equipamentos

$$b_i \sim N(\mathbf{0}, \mathbf{G}_1); b_{ij} \sim N(\mathbf{0}, \mathbf{G}_2); \epsilon_{ij} \sim N(\mathbf{0}, \sigma_e^2 \mathbf{I})$$

Validação do Modelo

Os pressupostos do modelo selecionado são apresentados na Figura 5.9, onde se verifica variância constante e média nula dos resíduos, não verificando o critério da normalidade.

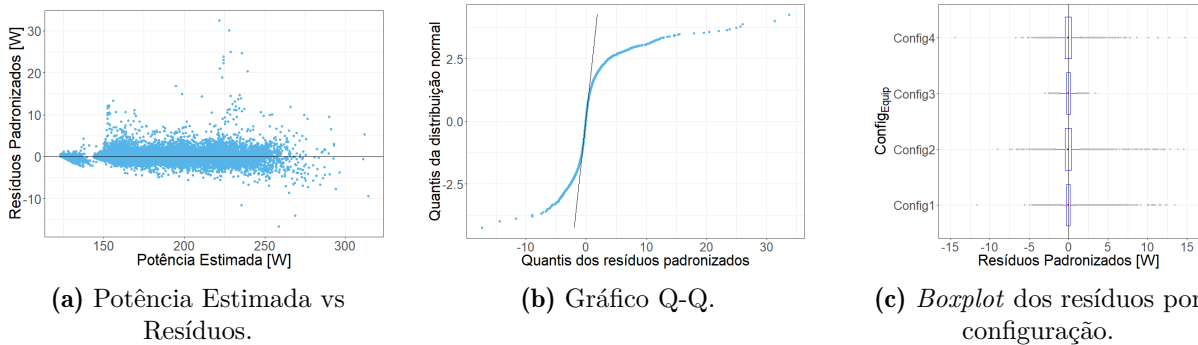


Figura 5.9: Análise residual para verificação da homocedasticidade e normalidade no modelo 3G - *Ericsson*.

A linearidade é verificada na Figura 5.10a. Em relação à independência dos resíduos, na Figura 5.10b observa-se um decaimento lento a partir do primeiro atraso e periodicidade de correlação positiva

e negativa. O resultado da correção é verificado na Figura 5.10c, com redução da amplitude de autocorrelação.

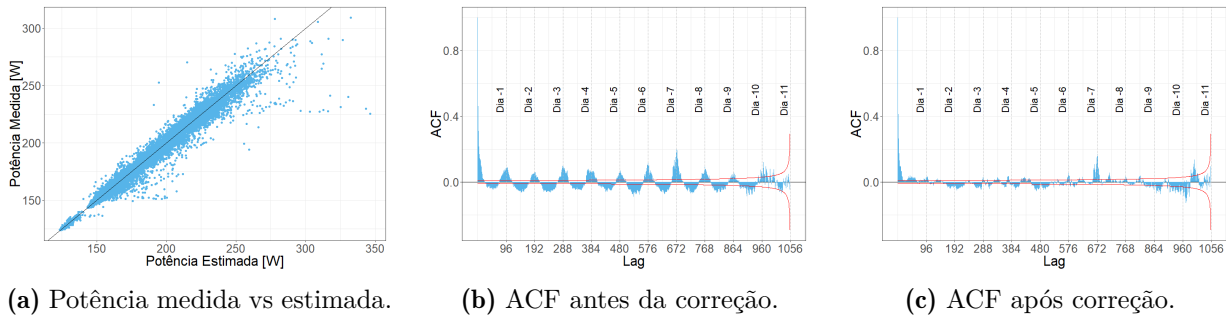


Figura 5.10: Análise residual para verificação da linearidade e independência no modelo 3G - *Ericsson*.

A Tabela 5.7a apresenta as estimativas intervalares para os coeficientes relativamente ao modelo da Equação 5.8. Paralelamente, foi estimado o mesmo modelo com as variáveis padronizadas para identificar qual o tipo de tráfego que tem maior importância na explicação da variabilidade da potência consumida. Na Tabela 5.7b é possível verificar que quando o tráfego em HSDPA varia unitariamente, a potência consumida aumenta em média 4.134 W. O tráfego R99 surge como a variável de menor importância (0.434 W), depois da voz (1.990 W).

Tabela 5.7: Estimativas intervalares a 95% para os coeficientes do modelo 3G - *Ericsson*.

	Estimativa Pontual	IC (95%)		Estimativa Pontual	IC (95%)		
(Intercept) [W]	$\hat{\beta}_0$	168.027	[141.451,194.602]	(Intercept) [W]	$\hat{\beta}_0$	181.475	[153.595,209.355]
T_{CS_3G} [Mb]	$\hat{\beta}_1$	0.084	[0.081,0.087]	T_{CS_3G} [Mb]	$\hat{\beta}_1$	1.990	[1.947, 2.034]
T_{PS_R99} [Mb]	$\hat{\beta}_2$	1.176	[1.071,1.282]	T_{PS_R99} [Mb]	$\hat{\beta}_2$	0.434	[0.390,0.478]
T_{HSDPA} [Mb]	$\hat{\beta}_3$	0.011	[0.009,0.012]	T_{HSDPA} [Mb]	$\hat{\beta}_3$	4.134	[3.287,4.981]

(a) Estimativas com variáveis não-padronizadas. (b) Estimativas com variáveis padronizadas.

A Figura 5.11 e 5.12 apresentam exemplos da aplicação do modelo para o período de teste nos equipamentos em ambas as bandas, 900 MHz e 2100 MHz.

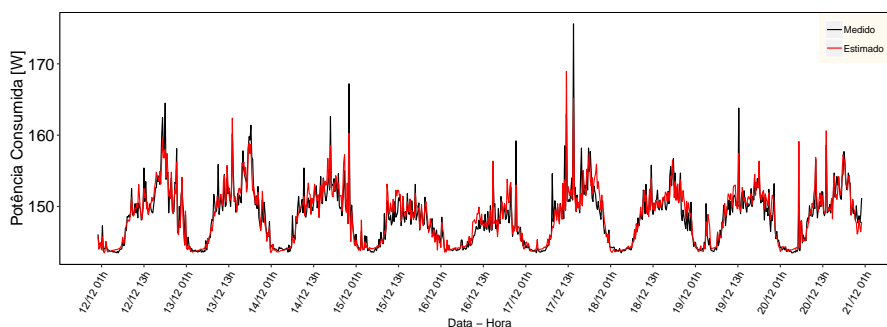


Figura 5.11: Aplicação do modelo 3G ao equipamento n° 02 da configuração 1 (900 MHz) - *Ericsson*.

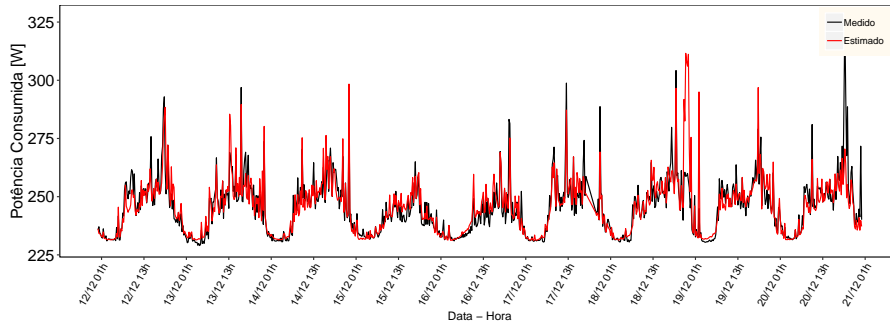


Figura 5.12: Aplicação do modelo 3G ao equipamento nº 26 da configuração 2 (2100 MHz) - *Ericsson*.

5.1.3 Modelo RRU 2G

Para o modelo 2G, utilizam-se como variáveis independentes o tráfego de voz, T_{CS_2G} (Erl), bem como o tráfego de dados PS, T_{PS_2G} (Mb). Todos os equipamentos possuem as mesmas configurações, assim sendo o modelo permanece com apenas 1 nível.

Seleciona-se então o efeito aleatório do *intercept* e inclinação na variável T_{CS_2G} . Durante a seleção dos modelos, foi verificado que a estimativa de correlação entre os elementos da matriz de efeitos aleatórios \mathbf{G} , continha o zero, gerando a hipótese da independência dos efeitos aleatórios. De forma a restringir parâmetros da matriz de efeitos aleatórios \mathbf{G} , outros modelos foram estimados com a matriz composta simétrica, identidade e diagonal. Conclui-se que a estrutura com a matriz diagonal foi a mais indicada, sendo o modelo dado por:

$$P_{(2G)_i} = \beta_0 + \beta_1 T_{CS_2G} + \beta_2 T_{PS_2G} + b_{0i} + b_{1i} T_{CS_2G} + \epsilon_i$$

com $i = 1, \dots, 15$ equipamentos (5.9)

$$\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G}); \epsilon_i \sim N(\mathbf{0}, \sigma_e^2 \mathbf{I})$$

Validação do Modelo

A Figura 5.13 apresenta a análise dos resíduos, na qual se constata que estão dispersos em torno de um valor médio nulo, apesar da existência de alguns *outliers*. De uma forma geral, o modelo cumpre o pressuposto de variância constante, no entanto o pressuposto da normalidade não é verificado.

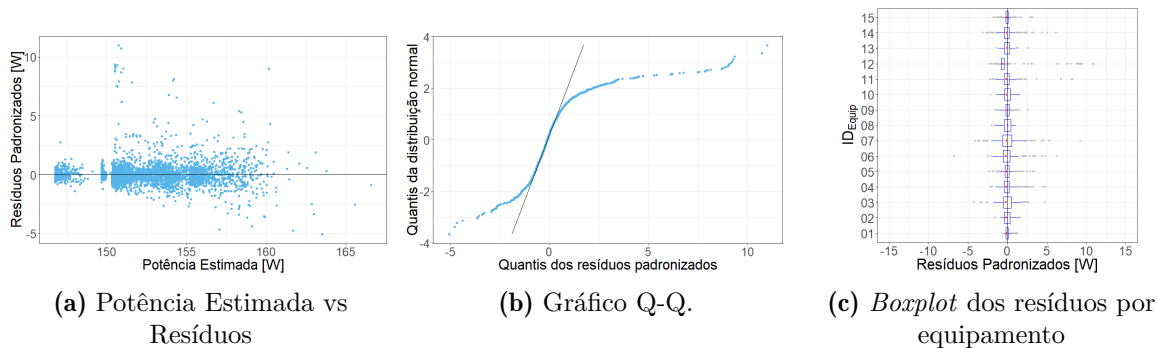


Figura 5.13: Análise residual para verificação da homocedasticidade no modelo 2G - *Ericsson*.

Na análise da Figura 5.14, observa-se linearidade entre a potência medida e estimada, a autocorrelação apresenta-se significativa, sendo o primeiro atraso com decaimento lento, foi ponderada a correção da dependência dos dados com um processo AR(1), no entanto, não se obtiveram ganhos significativos.

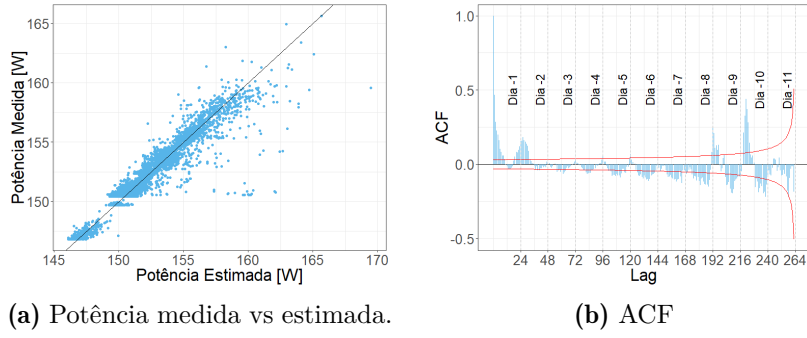


Figura 5.14: Análise residual para verificação da linearidade e independência no modelo 2G - *Ericsson*.

As estimativas intervalares dos efeitos fixos são mostrados na Tabela 5.8a e com as variáveis padronizadas na Tabela 5.8b. Destaca-se que uma maior influência do tráfego CS, no entanto ambos os coeficientes são baixos, o que evidencia baixa variabilidade do consumo energético nos equipamentos 2G.

Tabela 5.8: Estimativas intervalares a 95% para os coeficientes do modelo 2G - *Ericsson*.

	Estimativa Pontual	IC (95%)		Estimativa Pontual	IC (95%)		
(Intercept) [W]	$\hat{\beta}_0$	151.989	[150.745,153.233]	(Intercept) [W]	$\hat{\beta}_0$	152.731	[151.313,154.149]
T_{CS_2G} [Erl]	$\hat{\beta}_1$	2.065	[1.626,2.505]	T_{CS_2G} [Erl]	$\hat{\beta}_1$	0.619	[0.370, 0.867]
T_{PS_2G} [Mb]	$\hat{\beta}_2$	0.034	[0.031,0.0367]	T_{PS_2G} [Mb]	$\hat{\beta}_2$	0.329	[0.293,0.364]

(a) Estimativas com variáveis não-padronizadas.

(b) Estimativas com variáveis padronizadas.

A Figura 5.15 apresenta um exemplo da aplicação do modelo no período de teste. É possível observar a baixa variabilidade da potência consumida e a estimativa aproxima-se do valor medido.

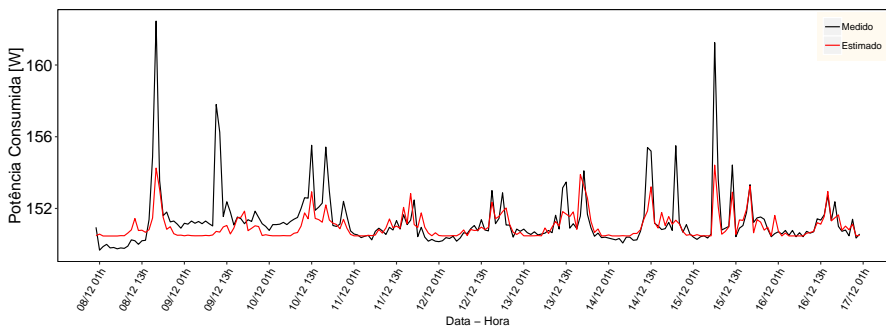


Figura 5.15: Aplicação do modelo 2G ao equipamento n° 06 - *Ericsson*.

5.1.4 Modelo RRU Partilhado 2G/3G

No modelo de consumo energético para os rádios partilhados entre as tecnologias 2G e 3G que operam na banda de frequências dos 900 MHz, foram utilizadas as mesmas variáveis de tráfego 2G e tráfego 3G apresentadas nos modelos anteriores.

O modelo tem *intercepts* e inclinação aleatória na variável T_{HSDPA} , sendo a \mathbf{G} matriz diagonal. É considerada a heterocedasticidade utilizando uma função de variância. Para modelar a estrutura dos resíduos em relação à dependência das observações, utiliza-se a função *seno* e *coseno* como efeito aleatório.

$$\begin{aligned}
 P_{(2G.3G)_i} = & \beta_0 + \beta_1 \mathbf{T}_{CS_2G} + \beta_2 \mathbf{T}_{PS_2G} + \beta_3 \mathbf{T}_{CS_3G} + \beta_4 \mathbf{T}_{PS_R99} + \\
 & \beta_5 \mathbf{T}_{HSDPA} + b_{0i} + b_{1i} \mathbf{T}_{HSDPA} + b_{2i} \sin \mathbf{H}_{rad} + b_{3i} \cos \mathbf{H}_{rad} + \epsilon_i \\
 & \text{com } i = 1, \dots, 21 \text{ equipamentos} \\
 & \mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G}); \epsilon_i \sim N(0, \sigma_\epsilon^2 |\mathbf{v}_i|^{2\delta})
 \end{aligned} \tag{5.10}$$

Validação do Modelo

A análise dos resíduos do modelo selecionado é apresentada na Figura 5.16, em que é possível verificar a linearidade, entretanto não se verifica a normalidade. O *boxplot* dos resíduos por equipamento mostra que os resíduos apresentam média nula.

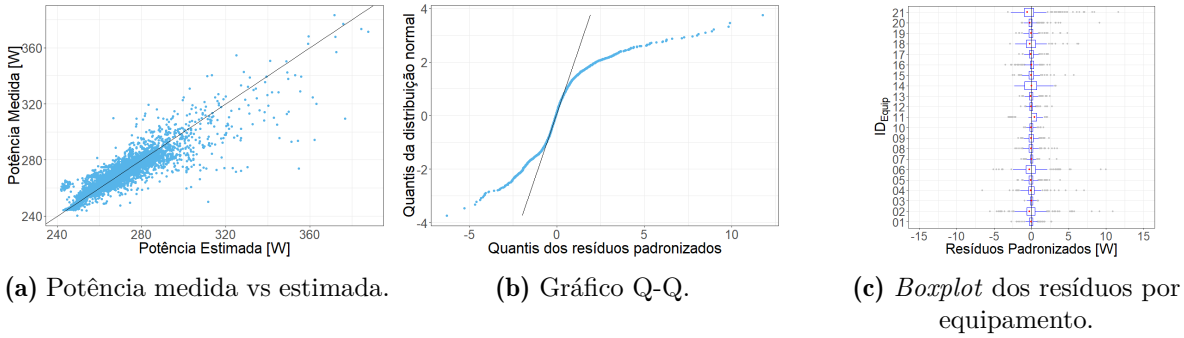


Figura 5.16: Análise residual para verificação da homocedasticidade no modelo partilhado 2G/3G - *Ericsson*.

A autocorrelação durante o desenvolvimento do modelo é ilustrada na Figura 5.17a, e após correção obtém-se os resultado que consta da Figura 5.17b, em que se observa uma redução da autocorrelação periódica que se aproxima do nível do ruído branco.

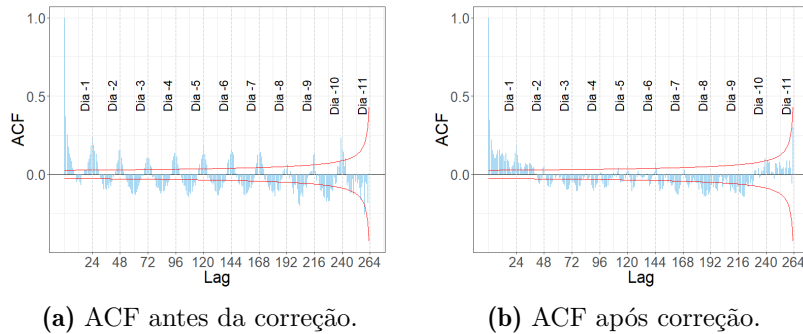


Figura 5.17: Análise residual para verificação da independência no modelo de rádio partilhado 2G/3G - *Ericsson*.

A Figura 5.18a mostra que o aumento da variabilidade dos resíduos cresce em função do aumento da potência estimada, indicando a falha de homocedasticidade. Assim, após a correção com uma função de potência para a variância, dada por $|\mathbf{v}_i|^{2\delta}$, em que \mathbf{v}_i são os valores estimados e δ o termo de potência. Na Figura 5.18b constata-se que os erros se distribuem aleatoriamente sem tendência.

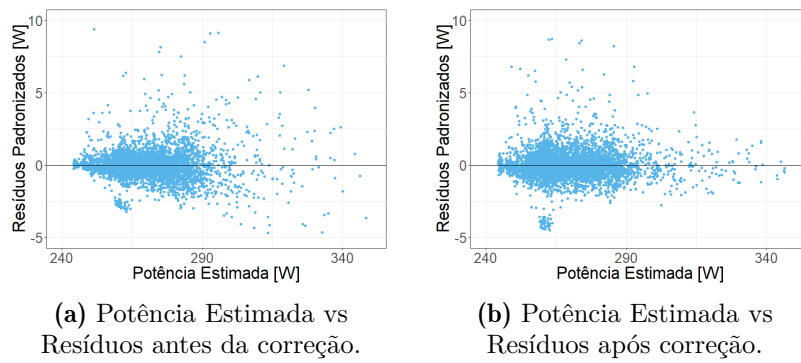


Figura 5.18: Análise residual para verificação da homocedasticidade no modelo de rádio partilhado 2G/3G - *Ericsson*.

As estimativas intervalares para os coeficientes do modelo são apresentadas na Tabela 5.9. O consumo *baseline* estimado é de 262.601 W, e cresce a uma taxa constante em função dos diferentes tipos de tráfego. Além disso, com as variáveis padronizadas verifica-se que as variações do tráfego em HSDPA têm maior impacto no consumo de potência. O tráfego R99 tem menor importância (0.228 W), depois das variáveis restantes que tem importância semelhantes.

Tabela 5.9: Estimativas intervalares a 95% para os coeficientes do modelo 2G/3G - *Ericsson*.

		Estimativa	IC (95%)			Estimativa	IC (95%)
		Pontual				Pontual	
(Intercept) [W]	$\hat{\beta}_0$	262.601	[258.871,266.331]	(Intercept) [W]	$\hat{\beta}_0$	269.429	[258.871,273.537]
T_{CS_2G} [ErI]	$\hat{\beta}_1$	3.108	[2.673,3.543]	T_{CS_2G} [ErI]	$\hat{\beta}_1$	1.624	[1.363,1.885]
T_{PS_2G} [Mb]	$\hat{\beta}_2$	0.048	[0.040,0.055]	T_{PS_2G} [Mb]	$\hat{\beta}_2$	1.149	[0.959,1.338]
T_{CS_3G} [Mb]	$\hat{\beta}_3$	0.066	[0.057,0.076]	T_{CS_3G} [Mb]	$\hat{\beta}_3$	1.090	[0.768,1.413]
T_{PS_R99} [Mb]	$\hat{\beta}_4$	0.758	[0.234,1.281]	T_{PS_R99} [Mb]	$\hat{\beta}_4$	0.228	[0.061,0.395]
T_{HSDPA} [Mb]	$\hat{\beta}_5$	0.018	[0.013,0.022]	T_{HSDPA} [Mb]	$\hat{\beta}_5$	4.048	[2.243,5.853]

(a) Estimativas com variáveis não-padronizadas.

(b) Estimativas com variáveis padronizadas.

Por fim, na Figura 5.19, pode-se verificar o resultado da aplicação do modelo num equipamento, em que a potência estimada se aproxima do valor real.

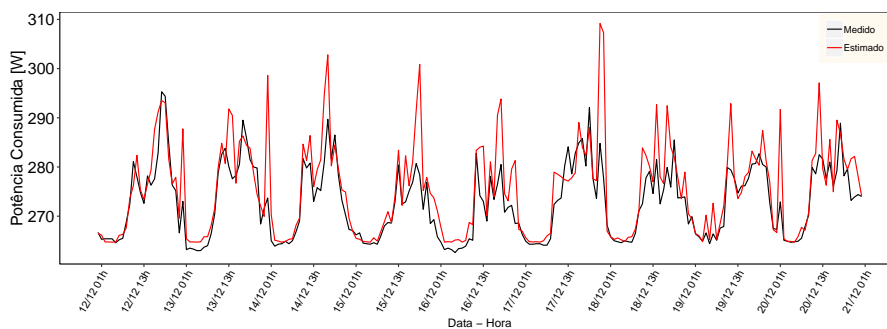


Figura 5.19: Aplicação do modelo partilhado 2G/3G ao equipamento n° 02 - *Ericsson*.

5.1.5 Modelo BBU

O consumo energético das BBUs mantém-se praticamente constante ao longo do tempo, dessa forma para o desenvolvimento dos modelos é utilizado apenas o valor médio. O modelo BBU 4G (Equação 5.11) considera o número de células ativas. O *intercept* (β_0), reflete o consumo energético de uma BBU com 3 células ativas.

$$P_{BBU.4G} = \beta_0 + \beta_1(N_{Cells} - 3) + \epsilon \quad (5.11)$$

Tabela 5.10: Estimativas intervalares a 95% para os coeficientes do modelo BBU 4G - *Ericsson*.

		Estimativa Pontual	IC (95%)
(Intercept) [W]	$\hat{\beta}_0$	121.433	[113.289, 129.577]
($N_{Cells} - 3$)	$\hat{\beta}_1$	5.233	[2.429, 8.036]

O modelo 3G, dado na Equação 5.12, utiliza como variável preditora a capacidade máxima do equipamento em número de CE (N_{CE}). Em que o *intercept*, reflete o consumo energético quando o *hardware* tem 384 CE de capacidade instalada.

$$P_{BBU.3G} = \beta_0 + \beta_1\left(\frac{N_{CE}}{384} - 1\right) + \epsilon \quad (5.12)$$

Tabela 5.11: Estimativas intervalares a 95% para os coeficientes do modelo BBU 3G - *Ericsson*.

		Estimativa Pontual	IC (95%)
(Intercept) [W]	$\hat{\beta}_0$	106.220	[82.737, 129.701]
($\frac{N_{CE}}{384} - 1$)	$\hat{\beta}_1$	77.040	[63.357, 90.727]

O modelo 2G de regressão linear de efeitos fixos, dado na equação 5.13, considera apenas o valor médio para estimar a potência consumida. As estimativas intervalares do modelo são apresentados na Tabela 5.12.

$$P_{BBU.2G} = \beta_0 + \epsilon \quad (5.13)$$

Tabela 5.12: Estimativas intervalares a 95% para os coeficientes do modelo BBU 2G - *Ericsson*.

		Estimativa Pontual	IC (95%)
(Intercept) [W]	β_0	40.54	[38.782, 42.313]

5.2 Vendor 2 - Huawei

Foram desenvolvidos três modelos, nomeadamente para os equipamentos de rádio 4G que operam na banda dos 800 MHz e 1800 MHz, equipamentos 3G na banda dos 2100 MHz, e por fim os equipamentos de rádio partilhados entre 2G e 3G na banda dos 900 MHz.

5.2.1 Modelo RRU 4G

No modelo 4G consideram-se as variáveis correspondentes ao volume de tráfego T_{4G} (Gb) e as *features* de poupança energética F_{Sleep} e F_{RF_OFF} . O modelo da Equação 5.14 tem *intercepts* e inclinações aleatórias na variável T_{4G} . A matriz G_1 e G_2 tem estrutura diagonal. Relativamente aos resíduos, utiliza-se a função de *seno* e *coseno* para corrigir a falta da independência dos resíduos. Além disso, é utilizada uma modelação da heterocedasticidade com um função de variância constante que é dependente da covariável (F_{Sleep}).

$$\begin{aligned}
 P_{(4G)ij} &\sim \beta_0 + \beta_1 T_{4G} + \beta_2 F_{RF_OFF} + \beta_3 F_{Sleep} + b_{0i} + b_{0ij} + b_{1i} T_{4G} + \\
 &\quad b_{1ij} T_{4G} + b_{2i} \sin H_{rad} + b_{3i} \cos H_{rad} + \epsilon_{ij} \\
 &\text{com } i = 1, \dots, 4 \text{ configurações ; } j = 1, \dots, 27 \text{ equipamentos} \\
 &\quad b_i \sim N(\mathbf{0}, G_1); b_{ij} \sim N(\mathbf{0}, G_2); \epsilon_{ij} \sim N(\mathbf{0}, \sigma_e^2 \delta_s^2);
 \end{aligned}
 \tag{5.14}$$

Validação do Modelo

A Figura 5.20a apresenta o resultado da ACF durante o processo de seleção do modelo, em que se verificou a falta de independência dos resíduos. Após a correção, na Figura 5.20b é confirmado que a auto-correlação reduziu.

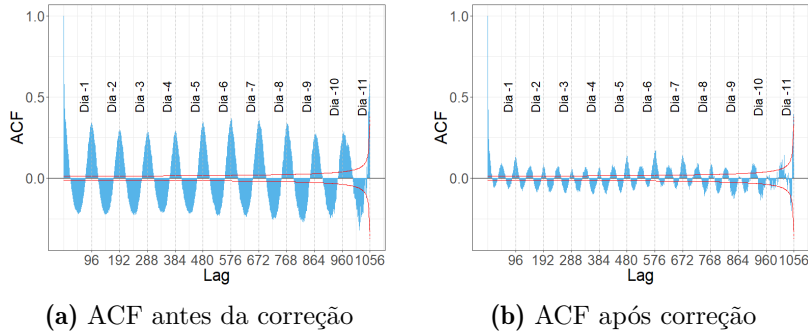


Figura 5.20: Análise residual para verificação da independência no modelo 4G - *Huawei*.

Na Figura 5.21a observa-se que os resíduos apresentam alguma tendência na variabilidade do erro com o aumento dos valores ajustados. Somente neste modelo, a variável F_{Sleep} é considerada como numérica no intervalo de $[0,1]$. Na Figura 5.21b, observa-se os períodos em que o equipamento está ligado, a variabilidade dos resíduos é maior, reduzindo ao desligar o mesmo. Devido a isso, foi aplicado uma função de variância constante por grupo, dado por δ_s^2 , em que s é a variável binária ON/OFF, onde o período "OFF" corresponde aos períodos em que $F_{Sleep} > 0$. A correção pode ser verificada na Figura 5.21c em que os resíduos se distribuem de forma aleatória em torno de zero. Na Figura 5.22 é possível verificar o pressuposto da linearidade. A normalidade novamente não é verificada.

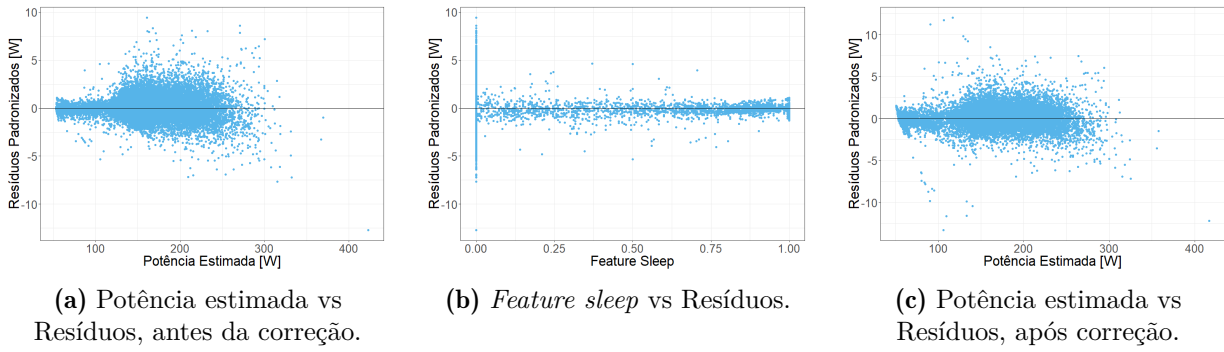


Figura 5.21: Análise residual para verificação da homocedasticidade no modelo 4G - *Huawei*.

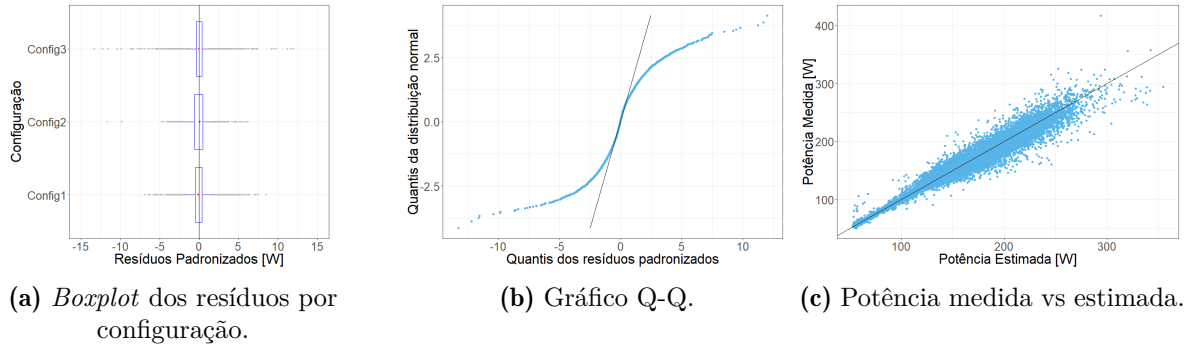


Figura 5.22: Análise residual para verificação da normalidade, linearidade no modelo 4G - *Huawei*.

A Tabela 5.13 apresenta o resultado das estimativas intervalares, onde se verifica uma potência *baseline* de 140.858 W, crescendo a uma taxa constante em função do volume de tráfego de 7.891 W/Gb. Relativamente às *features* de poupança energética, a F_{RF_OFF} permite reduzir no máximo 40.262 W, em comparação com a situação com *Sleep Mode*, que reduz 87.276 W.

Tabela 5.13: Estimativas intervalares a 95% para os coeficientes do modelo 4G - *Huawei*.

		Estimativa Pontual	IC (95%)
(Intercept) [W]	$\hat{\beta}_0$	140.858	[132.584,149.132]
T_{4G} [Gb]	$\hat{\beta}_1$	7.891	[3.453,12.328]
F_{RF_OFF} [%]	$\hat{\beta}_2$	-40.262	[-41.417,-39.107]
F_{Sleep} [%]	$\hat{\beta}_3$	-87.276	[-87.724,-86.827]

A Figura 5.23 apresenta a aplicação do modelo, observa-se que a *feature Sleep Mode* desliga totalmente a portadora reduzindo a potência para 50 W. Em comparação com a *feature Radio Frequency Channel Shutdown*, na Figura 5.24, que o consumo reduz para 100 W, e mantém a sinalização da portadora.

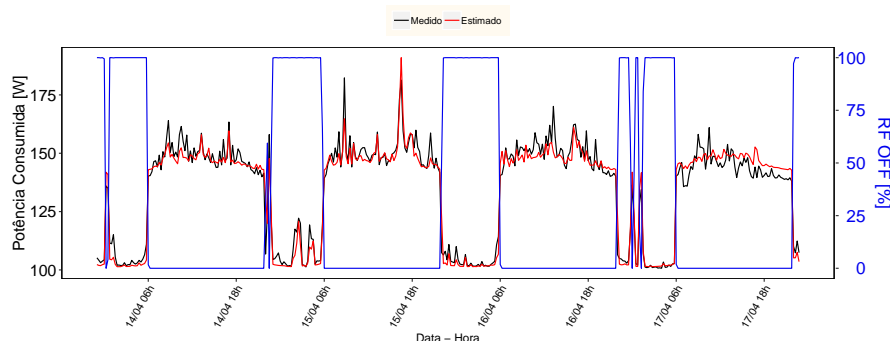


Figura 5.24: Aplicação do modelo 4G ao equipamento n° 27 da configuração 3 (1800 MHz) - *Huawei*.

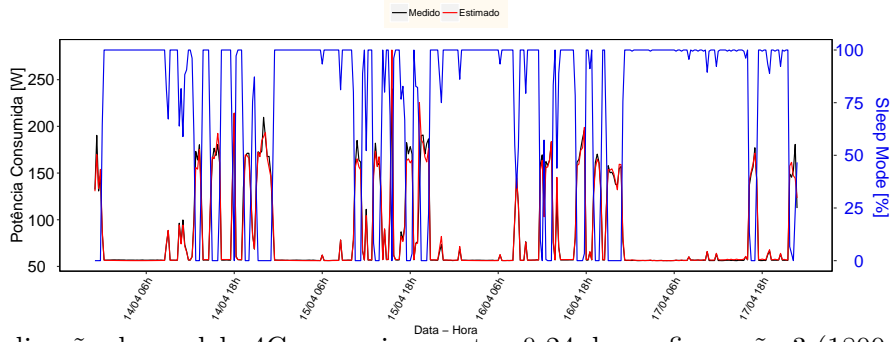


Figura 5.23: Aplicação do modelo 4G ao equipamento n^o 24 da configuração 3 (1800 MHz) - *Huawei*.

5.2.2 Modelo RRU 3G

O modelo 5.15 tem as seguintes variáveis predictoras: volume de tráfego PS (T_{PS_R99} e T_{HSDPA} em Mb), volume de tráfego CS (T_{CS_3G} em Erl) e a *feature Sleep Mode* (F_{Sleep}). O efeito aleatório consiste em *intercepts*, inclinação aleatória na variável F_{Sleep} e um termo de *seno* e *coseno* para atenuar o efeito da dependência dos resíduos. Considera-se que a matriz \mathbf{G} é diagonal. A variância é definida como uma função de potência dos valores ajustados.

$$\begin{aligned} P_{(3G)_i} &= \beta_0 + \beta_1 T_{CS_3G} + \beta_2 T_{PS_R99} + \beta_3 T_{HSDPA} + \beta_4 F_{Sleep} + \\ & b_{0i} + b_{1i} F_{Sleep} + b_{2i} \sin \mathbf{H}_{rad} + b_{2i} \cos \mathbf{H}_{rad} + \epsilon_i \\ & \text{com } i = 1, \dots, 6 \text{ equipamentos} \\ & b_i \sim N(\mathbf{0}, \mathbf{G}) \quad ; \quad \epsilon_i \sim N(\mathbf{0}, \sigma_e^2 |\mathbf{v}_i|^{2\delta}) \end{aligned} \quad (5.15)$$

Validação do Modelo

A Figura 5.25 apresenta a ACF antes e após correção, no qual se verificam a redução da autocorrelação periódica. No entanto, ainda mantém o decaimento lento no primeiro *lag* e a sua correção apresenta custos computacionais elevados.

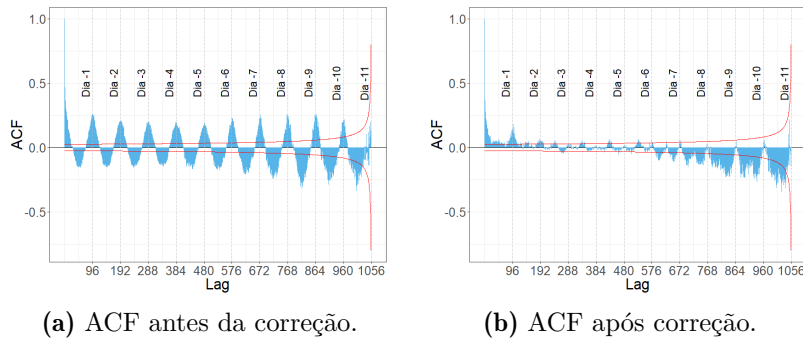


Figura 5.25: Análise residual para verificação da independência no modelo 3G - *Huawei*.

Na Figura 5.26a observa-se que a variação dos resíduos aumenta proporcionalmente com os valores ajustados. Desta forma, na Figura 5.26b verificam-se os resíduos depois aplicar a função de potência, dado por $|\mathbf{v}_i|^{2\delta}$. A Figura 5.27 apresenta a análise do pressuposto da linearidade em que é verificado, os resíduos têm média zero dentro de cada equipamento e não cumpre a normalidade.

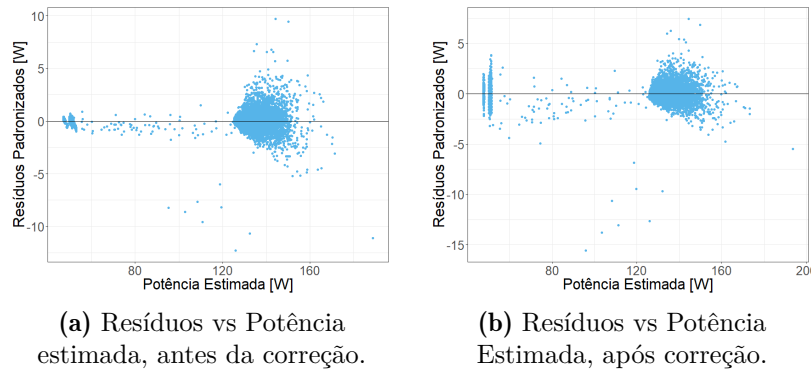


Figura 5.26: Análise residual para verificação de homoscedasticidade no modelo 3G - *Huawei*.

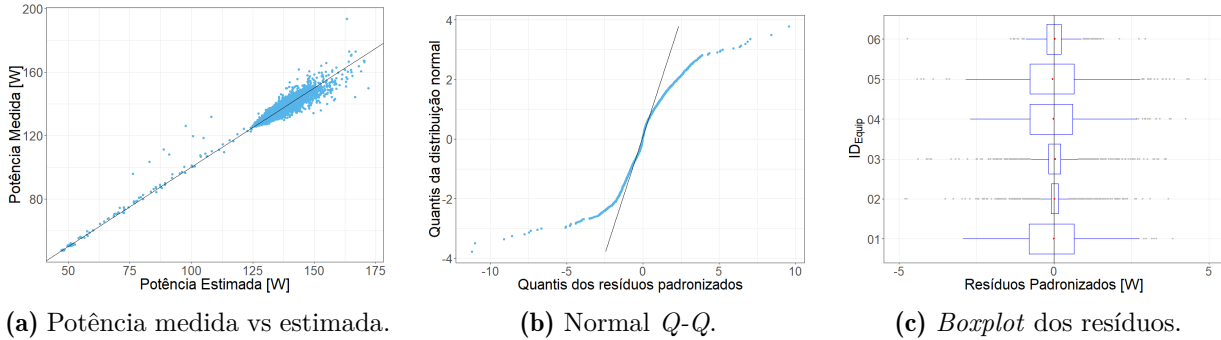


Figura 5.27: Análise residual para verificação da linearidade e normalidade no modelo 3G - *Huawei*.

As estimativas intervalares do modelo são dadas na Tabela 5.14a. Na Tabela 5.14b verifica-se que as variações do tráfego HSDPA tem maior variação unitária na potência consumida do rádio. De seguida no tráfego de voz, e por fim no tráfego PS R99. A Figura 5.28 apresenta a aplicação do modelo, em que os períodos que o rádio está desligado a potência reduz-se cerca de 100 W.

Tabela 5.14: Estimativas intervalares a 95% para os coeficientes do modelo 3G - *Huawei*.

	Estimativa			Estimativa		
	Pontual	IC (95%)		Pontual	IC (95%)	
(Intercept) [W]	$\hat{\beta}_0$	130.240	[128.279,132.540]	(Intercept) [W]	$\hat{\beta}_0$	132.337 [130.941,149.377]
T_{CS_3G} [Erl]	$\hat{\beta}_1$	1.116	[1.062,1.171]	T_{CS_3G} [Erl]	$\hat{\beta}_1$	1.484 [1.3449,1.623]
T_{PS_R99} [Mb]	$\hat{\beta}_2$	1.338	[1.081,1.594]	T_{PS_R99} [Mb]	$\hat{\beta}_2$	0.052 [0.064,0.075]
T_{HSDPA} [Mb]	$\hat{\beta}_3$	0.010	[0.010,0.011]	T_{HSDPA} [Mb]	$\hat{\beta}_3$	3.412 [3.305,3.518]
F_{Sleep} [ON/OFF]	$\hat{\beta}_4$	-80.298	[-81.542,-79.054]	F_{Sleep} [ON/OFF]	$\hat{\beta}_4$	-82.452 [-85.192,-79.712]

(a) Estimativas com variáveis não-padronizadas.

(b) Estimativas com variáveis padronizadas.

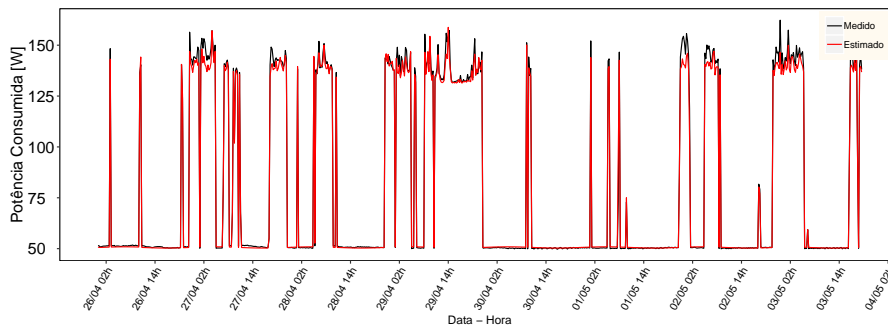


Figura 5.28: Aplicação do modelo 3G ao equipamento n° 02 - *Huawei*.

5.2.3 Modelo RRU Partilhado 2G/3G

As variáveis utilizadas no modelo de rádio partilhado 2G/3G, dado na Equação 5.16, foram o tráfego em 2G (T_{CS_2G} em Erl e T_{PS_2G} em Mb) e o tráfego 3G (T_{CS_3G} em Erl, T_{PS_R99} e T_{HSDPA} , ambos em Mb). O efeito aleatório de inclinação selecionado está na variável T_{CS_2G} . A variância dos resíduos é uma função de potência dos valores ajustados.

$$\begin{aligned} P_{(2G.3G)_{ij}} = & \beta_0 + \beta_1 T_{CS_2G} + \beta_2 T_{PS_2G} + \beta_3 T_{CS_3G} + \beta_4 T_{PS_R99} + \\ & \beta_5 T_{HSDPA} + b_{0i} + b_{0ij} + b_{1i} T_{CS_2G} + b_{2i} \sin \mathbf{H}_{rad} + b_{2i} \cos \mathbf{H}_{rad} + \epsilon_{ij} \end{aligned} \quad (5.16)$$

com $i = 1, \dots, 3$ configurações ; $j = 1, \dots, 19$ equipamentos

$$\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{G}_1) \quad ; \quad \mathbf{b}_{ij} \sim N(\mathbf{0}, \mathbf{G}_2) \quad ; \quad \epsilon_{ij} \sim N(0, \sigma_e^2 |\mathbf{v}_{ij}|^{2\delta})$$

Validação do Modelo

À semelhança dos modelos anteriores, a correção da autocorrelação periódica também foi observada, e conforme mostrado na comparação da Figura 5.29, é possível verificar a correção. De igual modo, na Figura 5.29 é mostrada a correção da falta de homocedasticidade utilizando a função de variância, onde se verifica que estão dispersos em torno de um valor médio nulo.

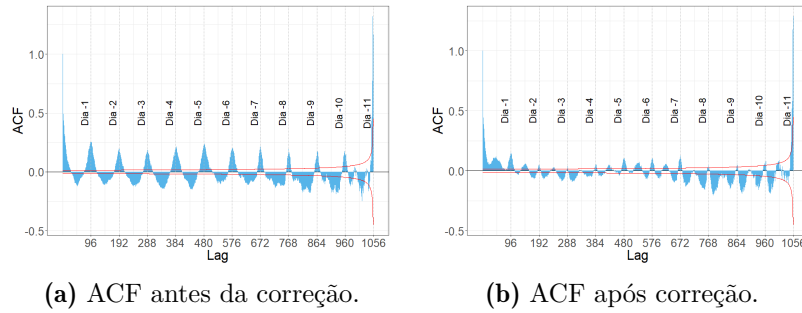


Figura 5.29: Análise residual para verificação da independência no modelo 2G/3G - *Huawei*.

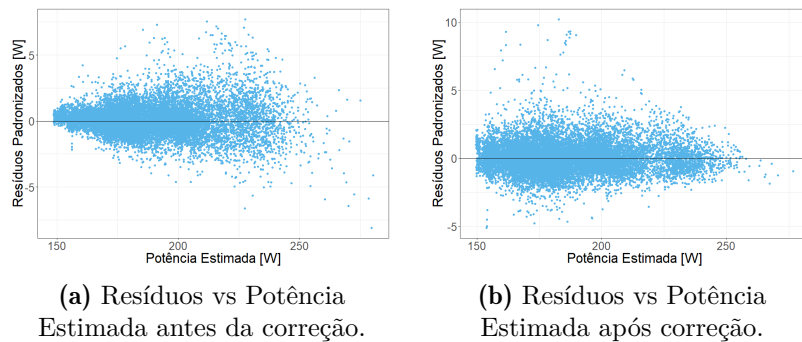


Figura 5.30: Análise residual para verificação da homocedasticidade no modelo 2G/3G - *Huawei*.

As estimativas dos coeficientes são apresentadas na Tabela 5.31. No resultado com variáveis padronizadas, o tráfego HSDPA também se mostrou mais importante, de seguida o tráfego de voz 3G e com menor importância o tráfego PS R99, enquanto que os tráfegos em 2G têm importância intermédia. A Figura 5.32 apresenta a aplicação do modelo sendo verificado um bom ajuste.

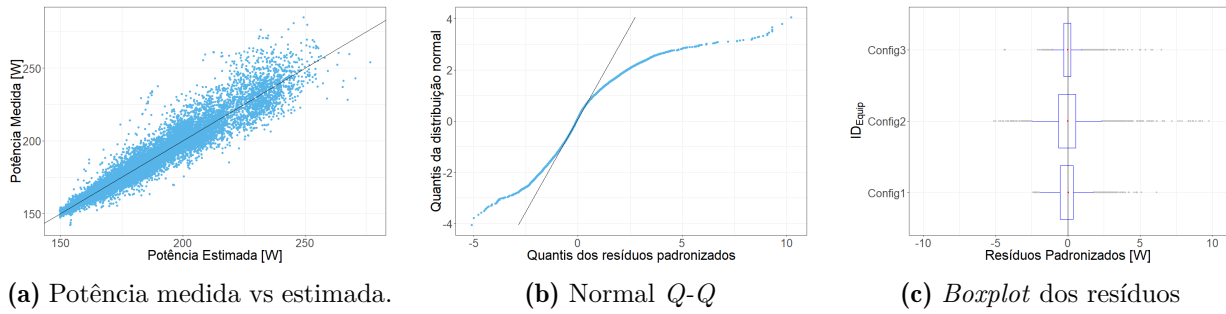


Figura 5.31: Análise residual para verificação da linearidade e normalidade no modelo 2G/3G - *Huawei*.

Tabela 5.15: Estimativas intervalares a 95% para os coeficientes do modelo 2G/3G - *Huawei*.

		Estimativa Pontual		IC (95%)			Estimativa Pontual		IC (95%)
(Intercept) [W]	$\hat{\beta}_0$	176.258	[168.842,183.674]		(Intercept) [W]	$\hat{\beta}_0$	186.147	[178.196,194.099]	
T_{CS_2G} [Erl]	$\hat{\beta}_1$	2.599	[0.814,4.384]		T_{CS_2G} [Erl]	$\hat{\beta}_1$	1.060	[0.485,1.636]	
T_{PS_2G} [Mb]	$\hat{\beta}_2$	0.134	[0.128,0.140]		T_{PS_2G} [Mb]	$\hat{\beta}_2$	1.609	[1.524,1.693]	
T_{CS_3G} [Erl]	$\hat{\beta}_3$	1.110	[1.073,1.146]		T_{CS_3G} [Erl]	$\hat{\beta}_3$	2.052	[1.929,2.176]	
T_{PS_R99} [Mb]	$\hat{\beta}_4$	0.608	[0.425,0.792]		T_{PS_R99} [Mb]	$\hat{\beta}_4$	0.232	[0.151,0.312]	
T_{HSDPA} [Mb]	$\hat{\beta}_5$	0.016	[0.015,0.017]		T_{HSDPA} [Mb]	$\hat{\beta}_5$	3.822	[3.734,3.910]	

(a) Estimativas com variáveis não-padronizadas.

(b) Estimativas com variáveis padronizadas.

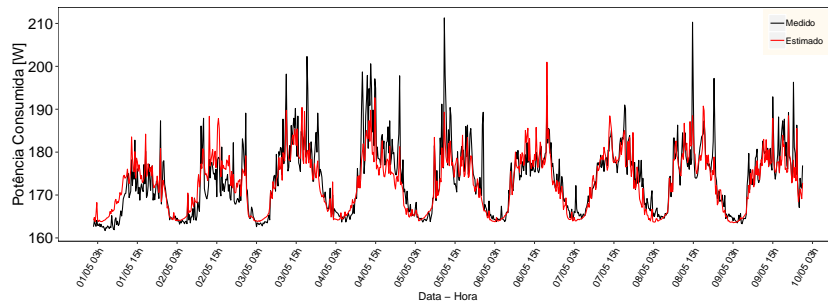


Figura 5.32: Aplicação do modelo partilhado 2G/3G ao equipamento n° 08 da configuração 1 - *Huawei*.

5.3 Estimativas e Métricas dos Modelos

De forma a generalizar e obter as estimativas do consumo energético em novos equipamentos deve-se utilizar somente a equação da reta representada pelos efeitos fixos. A Tabela 5.16 apresenta as métricas e as equações dos modelos desenvolvidos. Observa-se que tanto o modelo 2G como o de rádio partilhado apresentam as menores métricas de R_a^2 . Este facto pode ser devido aos equipamentos da tecnologia 2G apresentarem baixas variações do consumo energético. No modelo de rádio partilhado torna-se difícil separar a influência de cada tecnologia, uma vez que apesar de estarem configuradas no mesmo equipamento rádio, operam em amplificadores distintos. Entre os rádios que operam em modo dedicado, a tecnologia 4G possui maior consumo *baseline*.

Tabela 5.16: Equação de regressão linear estimada dos modelos - *Ericsson*.

Modelo	R_a^2	RMSE [W]	MAPE [%]	Equação de regressão linear estimada
4G	0.99	6.14	1.65	$P \sim 194.816 + 6.110 T_{4G} - 25.634 F_{MicroSleep} - 96.152 F_{Sleep}$
3G	0.99	3.55	0.92	$P \sim 168.027 + 0.084 T_{CS_3G} + 1.176 T_{PS_R99} + 0.011 T_{HSDPA}$
2G	0.88	1.04	0.38	$P \sim 151.989 + 2.065 T_{CS_2G} + 0.034 T_{PS_2G}$
2G/3G	0.80	6.84	1.49	$P \sim 262.601 + 3.108 T_{CS_2G} + 0.048 T_{PS_2G} + 0.066 T_{CS_3G} + 0.758 T_{PS_R99} + 0.018 T_{HSDPA}$

Como se pode observar na Tabela 5.17 as métricas do modelo de banda base, sendo o valor máximo do MAPE é de 8.49%. O modelo 2G possui baixo RMSE devido ao consumo médio dos equipamentos ser próximo.

Tabela 5.17: Métricas dos modelos de Banda Base - *Ericsson*.

	R_a^2	RMSE [W]	MAPE [%]
Modelo 2G	-	2.66	4.73
Modelo 3G	0.90	15.25	6.21
Modelo 4G	0.50	12.65	8.49

As estimativas dos equipamentos *Huawei* são apresentadas na Tabela 5.18. Devido ao tamanho da amostra ser reduzido, e os equipamentos possuírem características semelhantes, o modelo não consegue ser representativo da diversidade real de uma rede.

Tabela 5.18: Equação de regressão linear estimada dos modelos - *Huawei*.

Modelo	R_a^2	RMSE [W]	MAPE [%]	Equação de regressão linear estimada
4G	0.99	9.10	3.70	$P \sim 140.858 + 7.891 T_{4G} - 40.262 F_{RF_OFF} - 87.276 F_{Sleep}$
3G	0.99	2.35	1.10	$P \sim 130.240 + 1.116 T_{CS_3G} + 1.338 T_{PS_R99} + 0.010 T_{HSDPA} - 80.298 F_{Sleep}$
2G/3G	0.93	5.78	1.88	$P \sim 176.258 + 2.599 T_{CS_2G} + 0.134 T_{PS_2G} + 1.110 T_{CS_3G} + 0.608 T_{PS_R99} + 0.016 T_{HSDPA}$

5.4 Extrapolação do Consumo Energético

Esta secção detalha um caso de extrapolação para o consumo anual de energia de um operador real com equipamentos RRUs e BBUs *Ericsson*. Considerou-se o consumo de energia *baseline* estimados pelos modelos apresentados anteriormente, sem tráfego, e também a *busy hour* (95% do tráfego). A quantidade de CO_2 emitida baseia-se na Equação 5.17, onde E é a energia necessária para a operação da rede. O carvão foi assumido como o combustível fóssil usado para gerar essa energia, com um factor F_c de 94600 kg/TJ, [58].

$$CO_2 = E \times F_c \quad (5.17)$$

De acordo com o resultado apresentado na Tabela 5.19, o consumo anual de energia considerando o tráfego é de 90.2 GWh, com 30728.72 toneladas de CO_2 emitidas para a atmosfera. A rede analisada é composta por 4145 estações base (*sites*) com diferentes equipamentos, de modo que cada site consome em média 1591.51 kWh por mês. Considerando uma taxa de 0.125 €/kWh, o custo mensal é de 198.94€ por *site*.

Tabela 5.19: Extrapolação do consumo energético anual.

	4G	3G	2G	2G/3G
Nº BBUs	4242	4763	4146	0
Nº RRUs	15711	16074	7676	3534
Potência por Equipamento				
BBU <i>Baseline</i> [W]	121.4	106.2	40.5	0.0
BBU Média [W]	137.1	183.26	40.50	0.0
RRU <i>Baseline</i> [W]	194.8	168.0	151.9	262.6
RRU 95% Tráfego [W]	220.2	186.1	154.9	293.6
Consumo Energético Total e Pegada Ecológica (<i>Baseline</i>)				
Consumo Anual [GWh]	31.3	28.1	11.7	8.1
Emissões Anual de CO_2 [t]	10659.5	9569.7	3984.5	2758.5
Consumo Energético Total e Pegada Ecológica (95% Tráfego)				
Consumo Anual [GWh]	35.4	33.9	11.9	9.1
Emissões Anual de CO_2 [t]	12055.8	11545.0	4052.6	3099.1

Dada a evolução das gerações de comunicações móveis e o possível *switch-off* de tecnologias, na Tabela 5.20 é realizado uma estimativa da redução do consumo de energia ao desligar as tecnologias 2G e/ou 3G. Para isso, utilizou-se os resultados obtidos do consumo energético anual considerando o tráfego 95%, sem considerar os impactos devido à transferência de tráfego entre tecnologias. Em relação aos rádios partilhados 2G/3G, não é possível através do modelo estimar a redução do consumo energético caso alguma tecnologia fosse desligada. Neste caso, o *switch-off* de uma tecnologia iria alterar o valor do consumo *baseline*. A situação ideal seria se existissem amostras com o rádio a operar em modo dedicado e partilhado, podendo assim usar uma variável nominal com essa informação e estimar a redução. Posto isto, assumiu-se através da documentação do fabricante, que 30% de energia em rádios partilhados podem ser poupados, desativando uma das tecnologias.

Tabela 5.20: Impacto económico e pegada ecológica com o *switch-off* da rede 2G e/ou 3G.

	Live	Switch-Off 2G	Switch-Off 3G	Switch-Off 2G e 3G
Total Consumo Anual [GWh]	90.23	75.62	53.65	35.41
Custos [M€]	11.28	9.45	6.71	4.43
Emissões CO_2 [t]	30728.72	25753.14	18271.04	12059.2
Poupança [M€]	-	1.83 (-17%)	4.57 (-40%)	6.85 (-60%)
Poupança CO_2 [t]	-	4975.58	12457.77	18669.5

Na situação do *switch-off* da rede 2G, 1.83 M€ seriam poupados. Em contraste, o 3G tem mais impacto, já que 4.57 M€ seriam economizados, e evitando a emissão de 12457.77 toneladas de CO_2 ao longo de um ano.

Por razões históricas, a tecnologia 2G, por ser a mais antiga, seria a primeira a ser desligada, mas ainda assim é uma rede confiável com boa cobertura de voz. Assim, desligar a rede 3G é também um cenário realista, e o tráfego de voz poderia ser redistribuído para 2G e o tráfego de dados para 4G. Na situação extrema de ter apenas a rede 4G, seriam poupados 6.85 M€. Nesse caso, o tráfego de voz poderia ser transportado por VoLTE.

6

Conclusões

6.1 Síntese	74
6.2 Trabalho Futuro	75

6.1 Síntese

Nesta dissertação apresentaram-se modelos de consumo energético para os principais equipamentos que constituem uma estação base, nomeadamente RRUs e BBUs para as tecnologias atualmente massificadas no mercado de telecomunicações, 2G, 3G e 4G, e para os *vendors Ericsson e Huawei*.

Para a caracterização descritiva dos dados utilizaram-se medidas estatísticas e análise de componentes principais. Esta análise inicial dos dados permitiu concluir que o consumo energético tem uma forte dependência com o tráfego gerado nas estações base, tanto voz como dados. Esta etapa foi imprescindível para compreender como os dados se encontravam estruturados e assim definir o método de inferência estatística a utilizar na modelação dos mesmos.

No modelo de consumo energético para os equipamentos de rádio, a primeira abordagem foi baseada no modelo de regressão linear de efeitos fixos. Devido à estrutura dos dados, surgiu a necessidade de utilizar os modelos de regressão de efeitos mistos, pois a introdução dos termos de efeitos aleatórios permitiu controlar a variabilidade da amostra proveniente da estrutura de agrupamento presente neste tipo de dados.

Para a seleção dos modelos utilizaram-se as métricas de AIC e TRV, permitindo deste modo uma comparação entre os mesmos. Adicionalmente, foram analisados os pressupostos, e em caso de falha, foram utilizadas transformações com o intuito de minorar o efeito da não verificação de alguns pressupostos.

O modelo 4G, em ambos os *vendors*, utiliza o volume tráfego como principal variável preditora e, uma vez que as *features* de poupança energética estavam ativas nos equipamentos rádio, foram também consideradas como variáveis independentes no modelo.

Nos modelos 2G, 3G e de rádio partilhado (2G/3G banda dos 900 MHz), utilizaram-se os tráfegos de voz e de dados, correspondente a cada tecnologia. Paralelamente, os mesmos modelos foram estimados com as variáveis padronizadas de modo a identificar qual o tipo tráfego com maior influência no consumo energético. Assim, foi verificado tanto no modelo 3G de rádio dedicado como no modelo de rádio partilhado, que as variações do tráfego HSDPA possuem maior influência, pois requerem maior processamento para garantir o elevado ritmo de transmissão. Nos equipamentos 2G destaca-se maior influência do tráfego de voz, no entanto as estimativas dos coeficientes do modelo são baixas, o que é expectável devido ao facto de se verificar pouca variabilidade no consumo energético.

Em relação às *features* de poupança energética e à sua utilização, apenas estavam ativas na tecnologia 4G em ambos os *vendors*, e na tecnologia 3G *Huawei*. A *feature Sleep Mode*, é facilmente aplicada, mas em contrapartida, causa um maior impacto na rede. As *features Micro Sleep* e a *Radio Frequency Channel Shutdown* são boas opções e com menor impacto, no entanto, só estão disponíveis em equipamentos mais recentes.

Relativamente ao modelo de banda base, devido ao seu consumo energético ser praticamente constante ao longo do tempo, e para as BBUs 4G, utilizou-se o número de células ativas. Já para o 3G, foi utilizado o número máximo de CEs.

Dentro dos dois *vendors* apresentados, o tamanho da amostra dos equipamentos *Ericsson* era relativamente superior, permitindo obter melhores estimativas dos parâmetros populacionais. Através dos modelos desenvolvidos não é possível comparar quais modelos de *hardware* ou tipo de fabricante têm maior eficiência energética, pelo que seria necessário estarem submetidos às mesmas condições e terem uma dimensão amostral semelhante. Apenas se pode obter uma estimativa do consumo médio, e perceber o impacto das variações do tráfego no consumo energético dos equipamentos, independente do modelo de *hardware*.

A extrapolação do consumo energético para um operador real foi apresentada utilizando equipamentos *Ericsson*. Verificou-se que o consumo anual, considerando o *baseline* e com o tráfego na *busy-hour*, varia entre 79.2 GWh e 90.2 GWh, respetivamente. Desta forma, estima-se que cada estação base consome em média 1.59 MWh por mês.

Por fim, os modelos desenvolvidos utilizaram medidas reais de consumo energético a partir da monitorização de um conjunto de *sites* em locais representativos, permitindo obter estimativas confiáveis para que possam ser utilizados pelos operadores, e assim terem uma visão completa do consumo de energia da rede usando apenas contadores de volume de tráfego disponíveis.

6.2 Trabalho Futuro

À medida que a taxa de penetração para os serviços de dados móveis aumenta, é necessário mais estações base para corresponder ao crescimento do tráfego, sendo que os operadores móveis hoje estão entre os principais consumidores de energia, enfrentarão o desafio de reduzir o consumo de energia nas suas redes.

Nesta dissertação foram propostos modelos utilizando o método de regressão linear de efeitos mistos. De forma a evoluir nesta pesquisa, é possível utilizar outras abordagens para a inferência, por exemplo, o modelo linear generalizado que assume outro tipo de distribuição ou a regressão não-linear.

Estudos futuros desse tipo poderiam incluir as diferentes *features* de poupança de energia de todas as unidades de rádio, bem como mais equipamentos de rádio para melhor representar a diversidade da rede. Além disso, estimar o consumo dos demais equipamentos, como ar condicionado e unidades de transmissão seria um caminho a seguir.

Devido à atual densificação da rede móvel com micro e/ou pico células, seria interessante a monitorização do seu consumo energético, principalmente para os equipamentos da quinta geração de comunicações móveis. Ainda na sequência, desenvolver modelos de *forecast* (previsão) para estimar o aumento do consumo energético devido ao crescimento da taxa de penetração na rede que requer a instalação de novos *sites* na rede de acesso. Por fim, avaliar os impactos no consumo energético associado ao modelo de capacidade, nas situações de transferência de tráfego devido ao *switch-off* de tecnologias.

Referências

- [1] Ericsson, “Ericsson Mobility Report,” 2019. [Online]. Available: <https://www.ericsson.com/assets/local/mobility-report/documents/2019/ericsson-mobility-report-june-2019.pdf>
- [2] “Report on Climate Change, International Telecommunication Union (ITU),” 2008. [Online]. Available: <https://www.uncclearn.org/sites/default/files/inventory/itu10.pdf>
- [3] I. Humar, X. Ge, L. Xiang, M. Jo, M. Chen, and J. Zhang, “Rethinking energy efficiency models of cellular networks with embodied energy,” *IEEE Network*, vol. 25, 2011.
- [4] M. Alsharif, R. Nordin, and M. Ismail, “Survey of green radio communications networks: Techniques and recent advances,” *Journal of Computer Networks and Communications*, 2013.
- [5] P. Giri and S. Singh, “An analytical survey of power consumption and modeling in different areas of ict networks,” *International Journal of Computer Science and Information Technologies*, vol. 6, pp. 866–870, 02 2015.
- [6] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. V. Der Perre, “Challenges and enabling technologies for energy aware mobile radio networks,” *IEEE Communications Magazine*, vol. 48, no. 11, pp. 66–72, 2010.
- [7] P. Frenger, C. Friberg, Y. Jading, M. Olsson, and O. Persson, “Radio network energy performance: shifting focus from power to precision,” *Ericsson Review*, 2014.
- [8] A. Abrol and R. Kumar Jha, “Power optimization in 5G networks: A step towards green communication,” *IEEE Access*, vol. 4, pp. 1–1, 01 2016.
- [9] T. Berglund and H. Huynh, “Energy efficiency of radio units and its impact on RAN energy consumption,” Master’s thesis, Lund University Faculty of Engineering, Lund, 2017.
- [10] “Number of 2G mobile subscribers worldwide from 2008 to 2020.” [Online]. Available: <https://www.statista.com/statistics/226085/global-2g-mobile-subscriber-forecast/>
- [11] I. Fernandes, “Tecnologias celulares.” [Online]. Available: <http://paginas.fe.up.pt/~ee99207/Tecnologias/celulares/celulares.html>
- [12] J. G. Sempere, “An overview of the GSM system.” [Online]. Available: https://web.fe.up.pt/~mleitao/CMOV/Tecnico/GSM_Sempere.html

- [13] National Instruments, “Introduction to UMTS device testing transmitter and receiver measurements for WCDMA devices.” [Online]. Available: http://download.ni.com/evaluation/rtf/Introduction_to_UMTS_Device_Testing.pdf
- [14] H. Holma and A. Toskala, “WCDMA for UMTS: HSPA Evolution and LTE,” 2007.
- [15] 3rd Generation Partnership Project 3GPP, “Technical Specification Group Radio Access Network. UMTS 900 MHz Work Item Technical Report,” no. 25.816, 2009, version 8.0.0.
- [16] J. Wannstrom, “LTE-Advanced,” 2013. [Online]. Available: <http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced>
- [17] M. B. S. Stefania, T. Issam, “LTE - The UMTS Long Term Evolution: From Theory to Practice.” Wiley Publishing, 2009.
- [18] G. Abed, M. Ismail, and K. Jumari, “The Evolution to 4G Cellular Systems: Architecture and Key Features of LTE-Advanced Networks,” *International Journal of Computer Networks and Wireless Communications*, vol. 2, 2012.
- [19] Ericsson, “Long Term Evolution (LTE): as introduction,” 2007.
- [20] A. Mourato, “Desenvolvimento de um modelo de consumo energético para redes de acesso rádio 2G/3G/4G com base em dados reais,” Master’s thesis, Instituto Superior de Engenharia de Lisboa, Lisboa, 2017.
- [21] D. Gadze, S. Aboagye, and K. Agyekum, “Real time traffic base station power consumption model for telcos in ghana,” 2016.
- [22] J. Lorincz and T. Garma, “Measurements and modelling of base station power consumption under real traffic loads,” *Sensors (Basel, Switzerland)*, vol. 12, pp. 4181–310, 12 2012.
- [23] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume, and A. Fehske, “How much energy is needed to run a wireless network?” *IEEE Wireless Communications*, vol. 18, no. 5, pp. 40–49, 2011.
- [24] A. Capone, S. D’Elia, I. Filippini, A. Redondi, and M. Zangani, “Modeling energy consumption of mobile radio networks: An operator perspective,” *IEEE Wireless Communications*, 2017.
- [25] Nokia, “Nokia has ambitious plans to reduce network power consumption.” [Online]. Available: <https://www.nokia.com/blog/nokia-has-ambitious-plans-reduce-network-power-consumption/>
- [26] J. Wu, Y. Zhang, M. Zukerman, and E. Kai-Ning Yung, “Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey,” *IEEE Communications Surveys and Tutorials*, vol. 17, pp. 1–24, 04 2015.
- [27] Ericsson, “Power consumption guideline for rbs 6000,” 2009.
- [28] —, “Reduced Power Consumption in GSM RAN,” 2015.

- [29] S. Correa, *Probabilidade e estatística*, 2nd ed., Belo Horizonte, Brasil: PUC Minas Virtual, 2003.
- [30] “Intervalo de confiança.” [Online]. Available: <http://www.portalaction.com.br/inferencia/intervalo-de-confianca>
- [31] N. A. Weiss, *Introductory Statistics*, 10th ed., Pearson Education, 2016.
- [32] “Dimensionality Reduction with PCA and t-SNE in R.” [Online]. Available: <https://medium.com/@TheDataGyan/dimensionality-reduction-with-pca-and-t-sne-in-r-2715683819>
- [33] R. Bro and A. K. Smilde, “Principal component analysis,” *Analytical Methods*, pp. 2812–2831, 2014.
- [34] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*, 1st ed., Springer Publishing Company, Incorporated, 2014.
- [35] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer New York Inc., New York, NY, USA, 2001.
- [36] A. Gelman and J. Hill, “Data analysis using regression and multilevel/hierarchical models,” 2006.
- [37] A. Rocha, “Modelos lineares mistos: uma abordagem bayesiana,” Master’s thesis, Universidade de Brasília, Brasília, 2017.
- [38] J. C. Pinheiro, “Topics in mixed effects models,” Ph.D. dissertation, Univ. of Wisconsin, Madison, 2004. [Online]. Available: <http://web.math.ku.dk/~erhansen/web/stat1/pinheiro.pdf>
- [39] L. Chang, “Mixed Effects Models in R,” 2010. [Online]. Available: https://cosanlab.com/static/papers/MM_Intro_HO.pdf
- [40] R. Kievit, W. Eduard Frankenhuis, L. Waldorp, and D. Borsboom, “Simpson’s paradox in psychological science: A practical guide,” *Frontiers in psychology*, vol. 4, p. 513, 08 2013.
- [41] N. M. Laird and J. H. Ware, “Random-effects models for longitudinal data,” *Biometrics*, vol. 38, no. 4, pp. 963–974, 1982.
- [42] J. P. Littell, R.C. and R. Natarajan, “Using mixed models in a cross-over study with repeated measurements within periods,” *Tutorial in Biostatistic*, 2000.
- [43] Wikipédia, “Likelihood function.” [Online]. Available: https://en.wikipedia.org/wiki/Likelihood_function
- [44] F. Saarinen, “Using mixed models in a cross-over study with repeated measurements within periods,” Master’s thesis, Stockholm University, Stockholm, 2004.
- [45] Action Stat, “Análise de regressão.” [Online]. Available: <http://www.portalaction.com.br/analise-de-regressao/273-algumas-consideracoes>
- [46] S. Konishi and G. Kitagawa, “Information Criteria and Statistical Modeling,” 2008.

- [47] Wikipédia, “Information theory.” [Online]. Available: https://en.wikipedia.org/wiki/Information_theory
- [48] A. Loy, H. Hofmann, and D. Cook, “Model Choice and Diagnostics for Linear Mixed-Effects Models Using Statistics on Street Corners,” *Journal of Computational and Graphical Statistics*, 2017.
- [49] J. Pinheiro and D. Bates, *Mixed-Effects Models in S and S-PLUS*, Springer New York, 2006.
- [50] G. Verbeke and G. Molenberghs, “Linear mixed models for longitudinal data,” 2001.
- [51] S. Rodrigues, “Modelo de Regressão Linear e suas Aplicações,” Master’s thesis, Universidade da Beira Interior, Covilhã, 2012.
- [52] D. Lahuis, M. J. Hartman, S. Hakoyama, and P. Clark, “Explained Variance Measures for Multilevel Models,” *Organizational Research Methods*, vol. 17, pp. 433–451, 09 2014.
- [53] S. Nakagawa and H. Schielzeth, “A general and simple method for obtaining R² from generalized linear mixed-effects models,” *Methods in Ecology and Evolution*, vol. 4, no. 2, pp. 133–142, 2013.
- [54] D. J Barr, R. Levy, C. Scheepers, and H. J Tily, “Random effects structure for confirmatory hypothesis testing: Keep it maximal,” *Journal of memory and language*, vol. 68, pp. 255–278, 01 2013.
- [55] D. Bates, R. Kliegl, S. Vasishth, and H. Baayen, “Parsimonious Mixed Models,” *arXiv*, vol. 1506, 06 2015.
- [56] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, and R Core Team, *nlme: Linear and Nonlinear Mixed Effects Models*, 2019, R package version 3.1-139. [Online]. Available: <https://CRAN.R-project.org/package=nlme>
- [57] Celfinet, “RAN Energy Optimizer, Product Description,” 2016.
- [58] IPCC, “Guidelines for National Greenhouse Gas Inventories, Stationary Combustion,” 2006.