

MSC

2.º
CICLO

FCUP
2016

U.PORTO

Correlação de Eventos Ocorridos em Sistemas de
Metering de Energia, Instalados em Redes de
Distribuição de Baixa Tensão

Sara Raquel da Silva Almeida

FC

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Correlação de Eventos Ocorridos em Sistemas de Metering de Energia, Instalados em Redes de Distribuição de Baixa Tensão

Sara Raquel da Silva Almeida

Mestrado em Engenharia Matemática
Departamento de Matemática

2016

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Correlação de Eventos Ocorridos em Sistemas de Metering de Energia, Instalados em Redes de Distribuição de Baixa Tensão

Sara Raquel da Silva Almeida

Mestrado em Engenharia Matemática

Departamento de Matemática

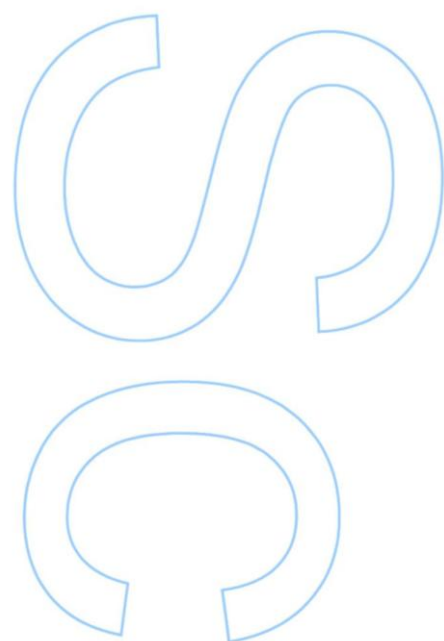
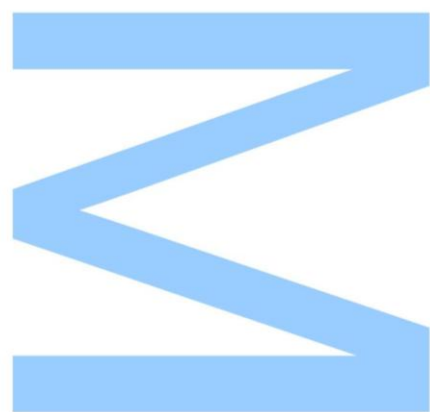
2016

Orientador

Prof^a Doutora Maria do Carmo Vaz Miranda Guedes, FCUP

Coorientador

Prof^a Doutora Rute Alexandra Borges de Almeida, FCUP

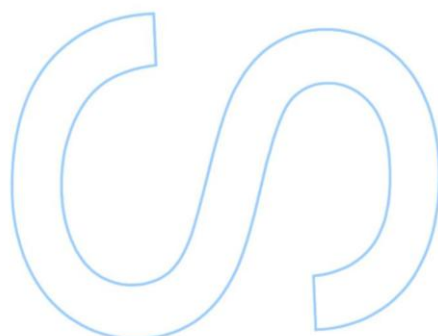
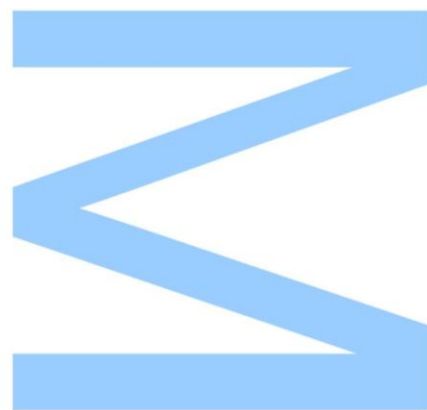




Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ____/____/____



Agradecimentos

Um grande e especial obrigada...

- Para os meus pais, que sempre me apoiaram e que sempre estiveram lá quando mais precisava. Por todos os conselhos e ensinamentos que fazem de mim quem sou hoje. Por toda a força, orientação e apoio nos melhores e piores momentos.
- Para os meus irmãos que estão sempre lá quando é preciso, apoiando-me incondicionalmente. Por todas as palavras amigas de encorajamento no momento em que mais preciso e por toda a felicidade que comigo partilham todos os dias.
- Para as minhas amigas de coração: Bruna e Inês que sempre estiveram comigo nos piores e melhores momentos. Por tanto me ouvirem sem se queixarem e por todos os conselhos que me deram, sem nunca pedir nada em troca.
- Para a minha orientadora, Maria do Carmo Guedes e coorientadora, Rute Almeida, por serem minhas Professoras e conselheiras e me ajudarem a crescer na viagem que foi o estágio curricular.
- Para o meu orientador, Alberto Jorge Bernardo, que tão bem me acolheu, que me fez crescer e que me forneceu outra visão do que é o mundo empresarial.
- Para todos os meus colegas de empresa por acreditarem no meu trabalho e me acolherem tão bem.
- Para a Efacec, a empresa que me acolheu e me fez crescer.

Resumo

É hoje essencial a otimização da eficiência energética, uma vez que o consumo de eletricidade se está a tornar mais exigente com a evolução da tecnologia e com o crescimento populacional, as pessoas têm mais aparelhos elétricos do que alguma vez foi pensado. Para além disso é possível também produzir-se eletricidade, como por exemplo, com a instalação de painéis solares fotovoltaicos, processo a que se chama de microprodução, que causam fenómenos de sobretensão (tensão acima dos valores máximos tolerados), os quais podem prejudicar os dispositivos domésticos diminuindo-lhes o tempo de vida ou mesmo avariando-os. Desta forma aquilo que se considerou ser uma distribuição energética ótima, está desadequada e nesse sentido são necessárias mudanças.

Há necessidade de compreender melhor o comportamento da rede de distribuição de baixa tensão, para melhor a planear e operar, motivo pelo qual a tendência é instalar dispositivos de monitorização nas redes de distribuição de baixa tensão, os contadores de energia.

Finalmente, a ligação à rede dos utilizadores finais (clientes ou microprodutores) e, por consequência, dos contadores de energia respetivos, é efetuada sem memória, ou seja, as equipas que os conectam e instalam, selecionam arbitrariamente a fase na qual intervêm. Não há cadastro dessa atividade pelo que é necessário criar ferramentas informáticas capazes de identificar as fases às quais os contadores de energia estão conectados, tema alvo deste estágio. Esse mapeamento é necessário para que se possa implementar soluções de software capazes de monitorizar a rede de distribuição de baixa tensão, em prol de um melhor desempenho do ponto de vista da eficiência energética e da sua resiliência.

Um dos objetivos a serem cumpridos nesse sentido é a identificação da(s) fase(s), dependendo de ser um contador monofásico ou trifásico, dos contadores numa rede de distribuição de baixa tensão. Vai ser apresentada uma possível resolução do problema pela aplicação de diferentes abordagens, passando pela utilização de métodos de modelação de séries temporais, deteção da associação entre diferentes conjuntos de dados e análise de estatística descritiva para as diferenças entre contadores e fases. Para além dos dados fornecidos foi necessário também a criação de outros conjuntos que permitissem melhorar a comparação e associação da morfologia das curvas de tensão dos contadores e das fases.

Por fim foi criado e implementado um programa que automaticamente forneça, a partir dos critérios escolhidos, a fase ou fases a que o contador foi ligado.

Este trabalho, efetuado em estágio em contexto curricular na empresa líder de mercado Efacec, permite dar um passo em frente na otimização da eficiência energética e espera-se que possa ter uma utilidade mais ampla para além daquilo que foi trabalhado.

Abstract

Today the optimization of the energetic efficiency is fundamental, since the consumption of electricity is becoming more and more demanding with the evolution of technology and the rapid population growth, people have more electronic devices than what it was ever thought. Furthermore it is also possible to produce electricity from the installation of photo-voltaic panels, for example, a process called micro-generation that causes values of voltage higher then the maximum tolerated, which can damage the electronic devises, diminishing their life time. In this way what was once thought to be an optimized electric distribution, is now inadequate and in that sense requires changes.

There is a need to better understand the low voltage distribution grid behavior, in order to better plan it and operate it, and for this monitoring devises, called meters, are being settled in the low voltage distribution grid..

Finally, the user's connection to the grid, and by consequence, the meters's connection is made by teams of specialized engineers that install them, selecting arbitrarily a phase in which it intervenes. Since there is no record of that activity it is necessary to create computational tools that can identify the phase to which those meters are connected, and that is the main subject of this work. The mapping of the grid is necessary to implement other software solutions capable of motorizing the grid in benefit of a better performance from the electric efficiency and resilience point of view.

So, phase identification is one of the main goals that need to be fulfilled in order to achieve electric efficiency, and in this report it will be presented a possible solution for that problem, using different approaches such as modeling methods for time series, association detection between different data sets and statistical description for series of differences between meters and phases. Besides the provided data sets, there was a need to create other data sets that could allow a better understanding of the association of the morphology voltage curves from meters and phases.

Finally it was implemented a program that automatically supplies from the chosen criteria, the phase to which the meter is connected.

This work done in a curricular context in the Efacec company, a market leader, allows us to give a

step forward in the electronic efficiency optimization and it is expected that this work might bring a more amplified use then what it was done in this report.

Conteúdo

Agradecimentos	4
Resumo	5
Abstract	7
Lista de Figuras	11
Lista de Tabelas	13
1 Introdução	16
2 Processos Estocásticos, Coeficientes e Testes Estatísticos: Elementos de Teoria	19
2.1 Processos Estocásticos	19
2.1.1 Processos Estacionários	19
2.1.2 Séries Temporais	20
2.1.3 Modelos Clássicos	22
2.1.4 Processos Não Estacionários	24
2.2 Coeficientes e Testes Estatísticos	27
2.2.1 Coeficientes de Correlação	28
2.2.2 Testes de Hipótese para a Homogeneidade de Variâncias	29
2.3 Outros Testes de Hipótese Usados	30
3 Descrição e Análise dos Dados	32
3.1 Descrição dos Dados	32
3.2 Análise dos Dados	33
3.3 Novos Conjuntos de Dados	37
3.4 Seleção do Melhor Modelo	40
3.4.1 Melhor Modelo Para a Tensão Média da Fase 1	40
3.4.2 Estimação dos Coeficientes dos Modelos sobre as Tensões dos Contadores Monofásicos	46

3.5	CrITÉrios, Resultados e ValidaÇ�o	49
3.5.1	Desvio Padr�o para as S�ries das Diferen�as e para as S�ries das Diferen�as das Derivadas	49
3.5.2	Coeficiente de Correla��o de Spearman	52
3.5.3	CrITÉrios Propostos	55
3.5.4	Aplica��o dos CrITÉrios e Classifica��o	57
3.5.5	Valida��o para a Rede 3	59
3.6	Elementos Computacionais	60
4	Conclus�o e Trabalho Futuro	62
	Bibliografia	63
	Anexos	66

Lista de Tabelas

2.1	Comportamento das funções ACF e PACF para a decisão dos modelos ARMA	24
3.1	Aplicação do teste de Ljung - Box para testar a correlação dos resíduos	44
3.2	Aplicação do teste de Shapiro para a normalidade dos resíduos	46
3.3	Valor do AIC	46
3.4	Aplicação do teste de Ljung - Box aos resíduos do modelo ótimo para o contador 1	47
3.5	Aplicação do teste de Shapiro aos resíduos do modelo ótimo para o contador 1	49
3.6	Medidas de dispersão para as três séries de diferenças do contador 1	49
3.7	Teste de hipótese de Brown-Forsythe para as séries das diferenças	51
3.8	Medidas de dispersão para as séries de diferenças das derivadas do contador 1	52
3.9	Teste de hipótese de Brown-Forsythe para as séries das diferenças das derivadas	52
3.10	Coeficiente de correlação de Spearman entre o Posto de Transformação e contadores . . .	53
3.11	Coeficiente de correlação de Spearman entre as séries das derivadas do Posto de Trans- formação e dos contadores	54
3.12	Resultados dos 5 critérios para a rede 1	55
3.13	Sistema de votação para a rede 1	56
3.14	Classificação obtida dos 5 critérios dos dados das 18h25 até às 2h20 do primeiro conjunto de dados	57
3.15	Classificação obtida dos 5 critérios dos dados das 2h25 até às 10h20 do primeiro conjunto de dados	57
3.16	Sistema de votação para as duas metades da rede 1	58
3.17	Porcentagem de acertos da classificação obtida dos 5 critérios dos dados da rede 2	59
3.18	Classificação obtida do critério escolhido aplicado aos dados da rede 3	59
4.1	Aplicação do teste de Ljung - Box para testar a correlação dos resíduos do modelo para a fase 2	105
4.2	Aplicação do teste de Shapiro para a normalidade dos resíduos do modelo para a fase 2 .	106

4.3	Aplicação do teste de Ljung - Box para testar a correlação dos resíduos do modelo para a fase 3	107
4.4	Aplicação do teste de Shapiro para a normalidade dos resíduos do modelo para a fase 3	109
4.5	Aplicação do Teste de Ljung - Box aos resíduos para a rede 1	109
4.7	Medidas de dispersão para as três séries de diferenças dos 36 contadores	111
4.12	Teste de hipótese de Brown-Forsythe para as séries das diferenças dos 36 contadores	116
4.14	Medidas de dispersão para as três séries de diferenças das derivadas dos 36 contadores	118
4.18	Teste de hipótese de Brown-Forsythe para as séries de diferenças das derivadas dos 36 contadores	122
4.20	Coeficiente de correlação de Spearman entre as séries de tensão para os 36 contadores	124
4.22	Coeficiente de correlação de Spearman entre as séries das derivadas para os 36 contadores	126
4.24	Resultados dos critérios para a rede 1	128
4.26	Resultados dos critérios para a 1 ^a metade da rede 1	130
4.28	Resultados dos critérios para a 2 ^a metade da rede 1	132
4.30	Sistema de votação para a rede 1 e ambas as Metades	134

Lista de Figuras

3.1	Diagrama de tensão média das três fases do Posto de Transformação	34
3.2	Diagrama de tensões médias do contador 1 e das três fases do Posto de Transformação . .	34
3.3	Diagrama de tensões médias do contador 21 e das três fases do Posto de Transformação .	35
3.4	Diagrama de tensão média dos contadores 7, 10, 29, 32 e 34	36
3.5	Diagrama de tensão média dos contadores 8, 9, 17, 22, 24 e 28	36
3.6	Diagrama da diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 1	37
3.7	Diagrama das derivadas da tensão média das fases 1, 2 e 3 no Posto de Transformação . .	38
3.8	Diagrama das derivadas da tensão média do contador 1	39
3.9	Diagrama da diferença entre as derivadas do contador 1 e as derivadas das 3 fases no Posto de Transformação	39
3.10	Função de autocorrelação da série da tensão média da fase 1 do Posto de Transformação .	40
3.11	Tendência polinomial de grau 6 da série de tensão média da fase 1 do Posto de Transformação	41
3.12	Função de autocorrelação da série de tensão média da fase 1 do Posto de Transformação da qual foi subtraída a tendência	41
3.13	Função de autocorrelação da série de tensão média da fase 1 do Posto de Transformação diferenciada	42
3.14	Função de autocorrelação parcial da série de tensão média da fase 1 do Posto de Trans- formação diferenciada	42
3.15	Função de autocorrelação dos resíduos do modelo para a fase 1 do Posto de Transformação	44
3.16	Histograma dos resíduos do modelo para a fase 1 do Posto de Transformação	45
3.17	Q - Q plot dos resíduos do modelo da fase 1	45
3.18	Função de autocorrelação dos resíduos do modelo ótimo para o contador 1	47
3.19	Histograma dos resíduos do modelo ótimo para o contador 1	48
3.20	Q - Q plot dos resíduos do modelo ótimo para o contador 1	48
4.4	Diagrama de tensões médias do contador 5 e das três fases do Posto de Transformação . .	68

4.6	Diagrama de tensões médias do contador 7 e das três fases do Posto de Transformação . .	69
4.9	Diagrama de tensões médias do contador 10 e das três fases do Posto de Transformação .	70
4.16	Diagrama de tensões médias do contador 17 e das três fases do Posto de Transformação .	75
4.27	Diagrama de tensões médias do contador 29 e das três fases do Posto de Transformação .	80
4.36	Diagrama da diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 5	85
4.37	Diagrama da diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 7	85
4.38	Diagrama da diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 10	86
4.40	Diagrama da diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 17	87
4.43	Diagrama da diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 29	88
4.46	Diagrama das derivadas da tensão média do contador 5	90
4.47	Diagrama das derivadas da tensão média do contador 7	91
4.48	Diagrama das derivadas da tensão média do contador 10	91
4.50	Diagrama das derivadas da tensão média do contador 17	92
4.53	Diagrama das derivadas da tensão média do contador 29	94
4.56	Diagrama da diferença entre as derivadas do contador 5 e as derivadas das 3 fases no Posto de Transformação	95
4.57	Diagrama da diferença entre as derivadas do contador 7 e as derivadas das 3 fases no Posto de Transformação	96
4.58	Diagrama da diferença entre as derivadas do contador 10 e as derivadas das 3 fases no Posto de Transformação	96
4.60	Diagrama da diferença entre as derivadas do contador 17 e as derivadas das 3 fases no Posto de Transformação	97
4.65	Tendência exponencial da série de tensão média da fase 1 do Posto de Transformação . . .	100
4.66	Tendência logarítmica da série de tensão média da fase 1 do Posto de Transformação . . .	101
4.67	Tendência potencial da série de tensão média da fase 1 do Posto de Transformação	101
4.68	Tendência polinomial de grau 2 da série de tensão média da fase 1 do Posto de Transformação	102
4.69	Tendência polinomial de grau 3 da série de tensão média da fase 1 do Posto de Transformação	102
4.70	Tendência polinomial de grau 4 da série de tensão média da fase 1 do Posto de Transformação	103
4.71	Tendência polinomial de grau 5 da série de tensão média da fase 1 do Posto de Transformação	103
4.72	Tendência linear da série de tensão média da fase 1 do Posto de Transformação	104

4.73	Função de autocorrelação dos resíduos do modelo para a fase 2 do Posto de Transformação	104
4.74	Histograma dos resíduos do modelo para a fase 2 do Posto de Transformação	105
4.75	Q - Q plot dos resíduos do modelo da fase 2	106
4.76	Função de autocorrelação dos resíduos do modelo para a fase 3 do Posto de Transformação	107
4.77	Histograma dos resíduos do modelo para a fase 3 do Posto de Transformação	108
4.78	Q - Q plot dos resíduos do modelo da fase 3	108

Capítulo 1

Introdução

Hoje em dia existe a necessidade de recolha de dados de instalação elétrica que permitam a tomada de decisões no que diz respeito à otimização da eficiência energética, sendo a procura de um equilíbrio da carga numa rede de distribuição de baixa tensão um dos muitos passos que podem ser dados nessa direção. Nesse sentido é necessário perceber primeiro alguns conceitos, nomeadamente, o de rede de distribuição de baixa tensão, posto de transformação, contadores e linhas de baixa tensão.

Uma rede de distribuição de baixa tensão é constituída por um Posto de Transformação, as linhas de baixa tensão e os contadores que estão ligados às linhas de baixa tensão.

O Posto de Transformação é o responsável por transformar a média tensão em baixa tensão que pode depois ser utilizada pelo consumidor final doméstico, pequeno industrial ou comercial.

As linhas de baixa tensão levam a energia elétrica desde os postos de transformação até aos locais onde é consumida em baixa tensão. Cada linha apresenta três fases, que foram denominadas por fase 1, 2 e 3.

Os contadores encontram-se nos locais de consumo ligados às linhas de baixa tensão e são uma ferramenta essencial e obrigatória para a monitorização contínua do estado da instalação elétrica que disponibilizam os consumos totais de energia bem como outras grandezas elétricas como a tensão, corrente, potência ativa. Os contadores podem ser trifásicos se se encontrarem ligados às três fases, ou monofásicos, se se encontrarem ligados a apenas uma fase. O contador que se situa no Posto de Transformação é sempre um contador trifásico e é o único do qual se tem sempre acesso à identificação das três fases.

São muitos os contadores que se encontram ligados a apenas uma das três fases de uma linha de baixa tensão, existindo uma distribuição não equitativa dos utilizadores finais, que têm contadores, pelas fases o que provoca aquilo a que se chama um desequilíbrio de carga. A razão da sua ocorrência deve-se essencialmente à instalação arbitrária dos contadores numa das três fases, no caso de ser monofásico.

Depois de identificada a fase a que um contador está ligado é possível proceder ao mapeamento da rede de distribuição de baixa tensão. Assim, a resolução deste problema permite solucionar outras situações como, por exemplo, o desequilíbrio de carga da rede, pelo facto de ser possível saber quantos e quais os contadores ligados a cada uma das fases; no caso de curto circuito perceber quais os contadores afetados e resolver este problema mais eficazmente; detetar fraudes identificando quais os contadores e portanto os consumidores que cometeram a fraude; possibilitar o aviso de problemas diretamente com os clientes afetados sem avisar incorretamente outros clientes, aumentando assim a satisfação do cliente; detetar a localização das áreas com mais perdas na rede, entre outros (Ramos et al., 2015).

Neste trabalho procura-se definir uma estratégia matemática fundamentada, que permita criar um método automático com o qual seja possível a identificação da fase a que um contador foi ligado, com base apenas nas curvas de tensão do mesmo e do Posto de Transformação respetivo. Nesse sentido procurou-se relacionar o comportamento da curva de tensão dos contadores com as curvas de tensão do contador no Posto de Transformação, uma vez que a curva do contador deverá apresentar um comportamento parecido com o comportamento da curva para a fase a que o contador estiver ligado.

Existem já algumas abordagens, nomeadamente, na identificação manual e empírica, por engenheiros especializados, observando os gráficos das séries obtidas pelo contador, por interrupção da passagem de energia ou na injeção de sinais elétricos em duas das fases alternadamente e observar quais os contadores que são afetados, classificando-os de acordo com o obtido (Weatherhead et al., 2015). Estas duas últimas abordagens afetam a comodidade do cliente, não sendo por isso métodos favoráveis.

Para o trabalho foram disponibilizados dados de três redes, denominadas por rede 1, 2 e 3 respetivamente, nas quais não existe micro-geração. A rede 1 tem 36 contadores mais o contador do Posto de Transformação, sendo que nenhum deles é trifásico com a exceção do contador no Posto de Transformação. A rede 2 tem 16 contadores mais o contador do Posto de Transformação, sendo que 9 deles são trifásicos, além do contador trifásico do Posto de Transformação. A rede 3 tem 43 contadores mais o contador do Posto de Transformação, onde nenhum deles é trifásico à exceção do contador do Posto de Transformação.

Para a rede 2 foi fornecida também a classificação obtida da primeira abordagem, de identificação visual e para a rede 3 foi fornecida a classificação obtida pelo método da injeção de sinal na fase a que está conectado contador e por deteção desse sinal no Posto de Transformação.

Os dados foram fornecidos num formato xml e para a análise exploratória inicial foram trabalhados com a ferramenta Excel do Microsoft Office. O resto do trabalho foi realizado usando o software estatístico R (R Core Team, 2014).

A estrutura da tese é a que se escreve seguidamente. No capítulo 2 é feita uma descrição teórica que inclui noções de modelação de séries temporais, bem como os métodos de diagnóstico usados para a validação do modelo, depois de escolhido como sendo o que melhor se adequa aos dados. São também descritos os coeficientes de correlação mais usados para a deteção de associação de variáveis e os testes de hipótese para a homogeneidade de variâncias, que vão ser depois necessários para o estudo que vai ser feito na secção 3.5.

No capítulo 3 é feita uma descrição, das três redes trabalhadas. Para os dados da rede 1 são apresentados os resultados da análise exploratória e descritiva, modelação das séries dos dados de tensão e é estudada a associação entre os dados de tensão das fases e contadores e entre as derivadas dos dados de tensão das fases e dos contadores.

Usando depois a rede 2 são aplicados todos os critérios e é escolhido aquele que melhor resolve o problema, sendo depois efetuada uma validação para os dados da rede 3. É ainda feita uma explicação dos problemas encontrados e resolvidos no processo de criação do programa que foi desenvolvido para a automatização do problema.

A conclusão e trabalho futuro são apresentados no capítulo 4.

Capítulo 2

Processos Estocásticos, Coeficientes e Testes Estatísticos: Elementos de Teoria

Neste capítulo serão apresentados alguns conceitos de teoria sobre processos estocásticos, testes estatísticos e coeficientes que foram usados ao longo do trabalho.

2.1 Processos Estocásticos

A modelação de séries temporais pode ser feita considerando as observações como uma realização de um processo estocástico. Para simplificação denotaremos o processo estocástico por Y_t . Sendo este uma família de variáveis aleatórias $\{Y_t : t \in T\}$.

2.1.1 Processos Estacionários

Sejam

- $\mu_{Y_t} = E(Y_t)$,
- $\gamma_Y(k) = C_{Y_t}(t, k) = Cov(Y_t, Y_{t-k}) = E[(Y_t - \mu_{Y_t})(Y_{t-k} - \mu_{Y_{t-k}})]$
- $\gamma_Y(0) = \sigma_{Y_t}^2 = Var(Y_t) = E[(Y_t - \mu_{Y_t})^2]$

a média, função de autocovariância e a variância, respetivamente, de um processo estocástico.

Para a modelação de séries temporais um conceito importante tem de estar presente. Esse conceito é o de estacionariedade.

Tem-se estacionariedade estrita quando para todo o n e todo o $t = t + 1, t + 2, \dots, t + n$, a distribuição conjunta de $Y_{t+1}, Y_{t+2}, \dots, Y_{t+n}$ é igual à distribuição conjunta de $Y_{t+1+h}, Y_{t+2+h}, \dots, Y_{t+n+h}$ para $|h| \geq 0$, ou seja, a distribuição conjunta de uma sequência de variáveis mantém-se constante ao longo do tempo.

Na prática, a estacionariedade estrita não pode ser verificada, pelo que se define estacionariedade fraca e se usa esta como padrão. A estacionariedade fraca ocorre quando as seguintes condições são satisfeitas:

- $E(Y_t) = \mu$, onde μ é constante,
- $\sigma_{Y_t}^2 = \sigma^2$, onde σ^2 é constante,
- $E(Y_t - E(Y_t))E(Y_\tau - E(Y_\tau))$, depende apenas de $t - \tau$.

Pode-se ver que na estacionariedade fraca não é a distribuição como um todo que é considerada mas apenas a média, a variância e a função de autocovariância e portanto, mais fraca nesse sentido (Burke and Hunter, 2005). A estacionariedade estrita implica a estacionariedade fraca, mas o contrário só é verdade para o caso gaussiano.

2.1.2 Séries Temporais

Uma série temporal pode ser vista como uma realização de um processo estocástico, indexado no tempo.

Supondo que existem T observações, estas serão denotadas por $y_t, t = 1, \dots, T$.

Um dos objetivos mais comuns da análise de séries temporais é a possibilidade de se poder fazer previsões para o futuro, através de um modelo baseado nos valores observados das séries temporais. No entanto, para além da previsão, o modelo obtido pode também ser usado para descrever as características mais importantes no padrão das séries temporais. O modelo pode ser formulado à custa de componentes estocásticas e determinísticas a partir dos termos da série.

Função de Autocorrelação

A função de autocorrelação, para um desfasamento, lag , h é dada por:

$$\rho_Y(h) = \frac{\gamma_Y(h)}{\gamma_Y(0)} \quad (2.1)$$

Esta função mede a correlação entre y_t e y_{t+h} , além disso têm-se que $\rho_Y(0) = 1$ e $\rho_Y(k) = \rho_Y(-k)$ (Box-Steffensmeier et al., 2014).

Função de Autocorrelação Parcial

A função de autocorrelação parcial, para um desfasamento h mede a correlação entre y_t e y_{t+h} removendo o efeito de valores de y intermédios (Box-Steffensmeier et al., 2014). Assim, para um desfasamento h maior ou igual a 2, tem-se:

$$\psi_{hh} = \frac{\rho_Y(h) - \sum_{j=1}^{h-1} \psi_{h-1,j} \rho_Y(h-j)}{1 - \sum_{j=1}^{h-1} \psi_{h-1,j} \rho_Y(j)} \quad (2.2)$$

onde $\psi_{hj} = \psi_{h-1,j} - \psi_{hh} \psi_{h-1,h-j}$ e T é o número de observações da série temporal. (Kočenda et al., 2015).

As funções referidas acima são funções teóricas que refletem as características do processo estocástico. No entanto, não é possível ter acesso a todas as realizações completas do processo, mas sim de uma amostra de uma só realização, a partir da qual se vai querer estimar o processo adequado, assumindo ergodicidade. Deste modo interessa definir as funções de autocorrelação e autocorrelação parcial amostrais da seguinte forma:

Função de Autocorrelação Amostral

$$\hat{\rho}_Y(h) = \frac{\sum_{t=s+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (2.3)$$

onde $\bar{y} = \frac{1}{T} \sum_{t=1}^T y_t$ é a média amostral, sendo que os restantes parâmetros mantêm o mesmo significado (Kočenda et al., 2015).

Função de Autocorrelação Parcial Amostral

Esta função é obtida substituindo o valor teórico da função de autocorrelação pela função de autocorrelação amostral, $\hat{\rho}_Y(h)$ na equação 2.2 (Kočenda et al., 2015), ou seja, ao longo do trabalho serão usadas as funções de autocorrelação amostral e autocorrelação parcial amostral, às quais foram atribuídas a sigla ACF e PACF respetivamente.

Perante uma série temporal há determinadas questões que se podem colocar, como por exemplo:

- Existe alguma tendência nos dados, ou seja, estes tendem a crescer ou a decrescer ao longo do tempo?
- A série apresenta sazonalidade, ou seja, existe a repetição de um padrão ao longo de um espaço temporal como um ano, mês, dia ou estações do ano, entre outros?
- Existem *outliers*, dados que estejam deslocados em relação aos outros dados?

- Ocorre alguma variação abrupta na série?

Todos estes tópicos devem ser considerados quando uma análise é efetuada à série temporal e devem ser tidos em conta para a escolha de um modelo adequado.

2.1.3 Modelos Clássicos

Sejam tendência, T_t , sazonalidade, S_t , e ruído aleatório, I_t , as componentes consideradas para a determinação dos modelos aditivo e multiplicativo, que resultam da combinação destas três componentes da seguinte forma:

- Modelo aditivo: $Y_t = T_t + S_t + I_t$
- Modelo Multiplicativo: $Y_t = T_t \cdot S_t \cdot I_t$

O modelo aditivo deve ser escolhido quando a variação sazonal é relativamente constante ao longo do tempo e não é proporcional ao nível de tendência. Quando a variação sazonal cresce com a tendência, então deve ser usado o modelo multiplicativo (Mimmac et al., 2001). Facilmente se transforma um modelo multiplicativo num aditivo pela aplicação da função logaritmo.

O modelo mais simples é aquele que apenas tem em consideração a tendência da série:

- $Y_t = T_t + \varepsilon_t$

onde ε_t é ruído branco, ou seja, uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média zero e variância constante.

A construção e estudo de outros modelos como o modelo cíclico, de tendência linear, modelo de regressão, em que a variável Y_t pode ser explicada à custa das componentes tendência e sazonalidade, pode ser consultada em (Harvey, 2001).

No entanto, modelos como o modelo de regressão acabam por não se tornar modelos apropriados, no caso em estudo neste trabalho, uma vez que a componente de tendência e sazonal são determinísticas e atribuem o mesmo peso às observações futuras. De modo a que isto não aconteça, constrói-se um modelo em que as componentes variam com o tempo.

Processo Autorregressivo, AR

O modelo autorregressivo de ordem p , $AR(p)$, é um modelo de processos estacionários que relaciona o valor atual y_t da série temporal com os valores do passado mais um termo de ruído branco. Está definido pela seguinte equação:

$$y_t = \sum_{i=1}^p \Phi_i y_{t-i} + \varepsilon_t \quad (2.4)$$

onde $\Phi_1, \Phi_2, \dots, \Phi_p$ são constantes e $\{\varepsilon_t\}$ é o ruído branco (Burke and Hunter, 2005).

A equação (2.4) pode ser reescrita na seguinte forma:

$$y_t - \sum_{i=1}^p \Phi_i y_{t-i} = \varepsilon_t \quad (2.5)$$

Introduzindo o operador de desfasamento L , tal que,

$$Ly_t = y_{t-1}$$

$$L^n y_t = y_{t-n}$$

a equação (2.5) pode ser escrita como

$$y_t - \sum_{i=1}^p \Phi_i L^i y_t = \varepsilon_t \quad (2.6)$$

ou ainda,

$$(1 - \sum_{i=1}^p \Phi_i L^i) y_t = \varepsilon_t \quad (2.7)$$

onde $1 - \sum_{i=1}^p \Phi_i L^i$ é uma função polinomial de L de grau p , denominada por $\Phi(L)$. Logo, a equação (2.7) é simplificada deste modo:

$$\Phi(L) y_t = \varepsilon_t \quad (2.8)$$

Prova-se que quando as raízes λ_i da equação

$$\Phi(\lambda_i) = 0 \quad (2.9)$$

para $\forall i = 1, \dots, p$, em módulo, são maiores do que um, o processo AR é estacionário (Burke and Hunter, 2005).

Processo de Média Móvel, MA

O modelo média móvel de ordem q , $MA(q)$, é um modelo que só pode ser aplicado no caso de se ter uma série temporal estacionária e que contém apenas componentes estocásticas. O seu nome provém do facto de y_t ser obtido pela aplicação de pesos, componentes de média móvel, às variáveis $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ e depois mover esses pesos e aplicar às variáveis $\varepsilon_{t+1}, \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}$ para a obtenção de y_{t+1} e assim sucessivamente. Este modelo fica definido pela seguinte equação:

$$y_t = \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2.10)$$

onde ε_t é ruído branco e $1, -\theta_1, -\theta_2, \dots, -\theta_q$ são as componentes de média móvel (Cryer and Chan, 2008).

Analogamente ao que foi feito para o modelo AR definido à custa de um polinómio de grau p , o modelo MA pode ser definido à custa da função polinomial de L de grau q como:

$$\Theta(L) = 1 - \sum_{i=1}^q \theta_i L^i \quad (2.11)$$

E que usando a função (2.11) então:

$$y_t = \Theta(L)\varepsilon_t \quad (2.12)$$

Quando combinados os dois processos anteriores obtém-se o chamado processo autorregressivo de média móvel.

Processo Autorregressivo de Média Móvel, ARMA

O modelo autorregressivo de média móvel de ordem p e q, ARMA(p,q), está definido pela seguinte equação:

$$y_t = \sum_{i=1}^p \Phi_i y_{t-i} + \varepsilon_t - \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2.13)$$

O modelo ARMA pode ainda ser reescrito à custa das funções $\Phi(L)$ e $\Theta(L)$ de uma forma mais compacta,

$$\Phi(L)y_t = \Theta(L)\varepsilon_t \quad (2.14)$$

ou ainda,

$$y_t = \frac{\Theta(L)}{\Phi(L)}\varepsilon_t \quad (2.15)$$

Note-se que a condição de estacionariedade é a mesma que no processo autorregressivo, ou seja, a série é estacionária se as raízes de $\Phi(L)$ forem superiores a 1, não sendo colocada nenhuma restrição à função $\Theta(L)$ (Harvey, 2001).

Com a análise do comportamento das funções de autocorrelação e autocorrelação parcial, ACF e PACF, de uma série temporal observada, é possível decidir qual dos modelos AR, MA ou ARMA se deve usar. Essa informação encontra-se na tabela 2.1 (Cryer and Chan, 2008).

	AR(p)	MA(q)	ARMA(p,q)
ACF	decaimento gradual	trunca para desfasamento > q	decaimento gradual
PACF	trunca para desfasamento > p	decaimento gradual	decaimento gradual

Tabela 2.1: Comportamento das funções ACF e PACF para a decisão dos modelos ARMA

2.1.4 Processos Não Estacionários

Os modelos apresentados anteriormente são alguns dos modelos que funcionam muito bem quando temos uma série estacionária, mas para a estimação de tais modelos é necessário antes de mais averiguar à cerca da estacionariedade da série em análise.

Existem vários métodos e testes que podem ser usados para a avaliação de estacionariedade, no entanto um processo mais rápido de observação da não estacionariedade da série é através da sua função de autocorrelação.

Um decaimento da ACF muito lento denuncia não estacionariedade relevante, que tem de ser considerada na modelação (Cryer and Chan, 2008, Yaffee and McGee, 2000).

Quando a situação é de não estacionariedade da série, os dados deverão ser submetidos a transformações até se obter a estacionariedade ou aproximação dela. Algumas técnicas mais frequentes são a diferenciação a um passo ou a diferenciação sazonal ou de tendência.

A diferenciação a um passo é obtida fazendo $y_t - y_{t-1}$, sendo esta a primeira diferença, denotada por Δy_t e o processo diz-se integrado de ordem 1. Uma segunda diferenciação obtém-se calculando $\Delta(y_t - y_{t-1})$ e o processo diz-se integrado de ordem 2. A diferenciação pode ser aplicada tantas vezes quantas necessárias até a obtenção da estacionariedade, no entanto a diferenciação excessiva pode causar depois problemas na estimação dos parâmetros (Cryer and Chan, 2008), pelo que se deve ir controlando esta técnica usando um método de verificação de estacionariedade da série, como a observação da ACF, e parar quando for obtida.

Um modelo que inclui a diferenciação é o modelo autorregressivo integrado de média móvel, que é aplicado a séries não estacionárias.

Processo Autorregressivo Integrado de Média Móvel, ARIMA

O modelo autorregressivo integrado de média móvel, $ARIMA(p,d,q)$, tem como parâmetros p do processo autorregressivo, q do processo média móvel e d que representa o número de vezes que a série foi diferenciada, denominando-se um processo integrado de ordem d . Este modelo é obtido supondo que a série segue um modelo ARMA estacionário depois de diferenciada. Está definido pela seguinte equação:

$$\Phi(L)\Delta^d y_t = \Theta(L)\varepsilon_t \quad (2.16)$$

usando as funções $\Phi(L)$ e $\Theta(L)$ definidas anteriormente pelas equações 2.7 e 2.11 (Harvey, 2001).

É de notar que um modelo $ARIMA(p, 0, q) = ARMA(p, q)$.

O processo ARIMA tem como pressupostos o igual espaçamento entre os dados e a não existência da falta de dados na série (Yaffee and McGee, 2000).

Assim, dependendo do facto da série ser ou não estacionária, diferentes modelos podem ser aplicados. Alguns dos modelos mais comuns são, caso se esteja perante uma série estacionária, o modelo autorregressivo, o modelo média móvel e o autorregressivo de média móvel. Caso a série seja não estacionária, então pode ser aplicado um modelo autorregressivo integrado de média móvel.

Box e Jenkins fornecem os passos básicos para a escolha do modelo ideal para os dados (Wang, 2006):

- identificação do modelo;
- estimação dos parâmetros do modelo;
- diagnóstico e crítica ao modelo, onde a análise aos resíduos é efetuada de modo a validar o modelo e até sugerir um modelo alternativo.

Estes passos devem ser repetidos até se encontrar um modelo satisfatório.

Para a identificação do melhor modelo, duas das medidas mais comumente usadas são o Critério de Informação de Akaike e o Critério de Informação Bayesiana. Para além de servirem para a escolha dos parâmetros podem ser usados para a escolha dos modelos em si.

Critério de Informação de Akaike

O critério de informação de Akaike, AIC, diz que o modelo escolhido deve ser aquele que minimiza a equação:

$$AIC(k) = -2\log(ML) + 2k \quad (2.17)$$

onde ML é máxima verosimilhança e k o número de parâmetros independentemente ajustados dentro do modelo. Se o modelo contiver um termo constante então $k = p + q + 1$, senão $k = p + q$, sendo p e q os parâmetros do modelo (Cryer and Chan, 2008).

No caso do modelo $ARMA(p, q)$ tem-se que $k = p + q$ e o valor de AIC pode ser calculado através da seguinte equação:

$$AIC(p, q) = n\log(\sigma^2) + 2(p + q) \quad (2.18)$$

onde σ^2 é a variância dos resíduos do modelo e n o tamanho da amostra (Wang, 2006).

Critério de Informação Bayesiana

O critério de informação Bayesiana, BIC, é definido por

$$BIC(k) = -2\log(ML) + k\log(n) \quad (2.19)$$

onde n é o tamanho da amostra e as outras variáveis mantêm também o seu significado.

Finalmente, depois de escolhido o modelo, é necessário analisar os seus resíduos de modo a perceber se de facto o modelo escolhido é o modelo que melhor se adequa aos dados. Esta análise perfaz a validação do modelo.

Um dos testes mais importantes para que um modelo se possa dizer adequado é que os resíduos sejam independentes, uma vez que isto garante que o modelo escolhido é adequado, no entanto na prática aceita-se um modelo como adequado com base apenas na não correlação dos resíduos. Assim, uma das maneiras mais simples de verificar o pressuposto da não correlação é observar a função de autocorrelação dos resíduos do modelo. Isto é importante porque uma vez que os resíduos são uma componente da variável que se está a modelar, e caso os resíduos pareçam ser correlacionados, então significa que existe um padrão que não foi tomado em consideração pelos termos AR e MA. Se estes forem autocorrelacionados, então não são ruído branco, e deve ser procurado outro modelo (Pankratz, 2009).

Para testar analiticamente a correlação dos resíduos pode ser usado o teste de Ljung - Box, como é referido na secção 2.3.

Assim como foi já referido, na prática aceita-se um modelo como adequado apenas com base na não correlação dos resíduos e para assegurar que não correlação significa independência de facto, os resíduos devem também ser normalmente distribuídos. Para os modelos AR, MA, ARMA ou ARIMA assume-se habitualmente o pressuposto de que as componentes ε_t devem ser independentes e identicamente distribuídas e que seguem uma distribuição Gaussiana (Hipel and McLeod, 1994). O pressuposto da normalidade pode ser verificado visualmente através do histograma dos resíduos ou do gráfico quantil-quantil, $Q - Q$ plot (Ekstrom and Sørensen, 2014).

Quando o histograma dos resíduos apresenta uma forma que seja semelhante à função densidade de uma distribuição normal de média zero e variância 1, portanto em forma de sino, tal indica a normalidade dos resíduos.

Outro indicador são os gráficos $Q - Q$ plot, que no caso em que os dados são normalmente distribuídos devem ter a forma de uma linha reta.

Analiticamente o teste de Shapiro pode ser usado para testar a normalidade dos resíduos.

2.2 Coeficientes e Testes Estatísticos

Um dos propósitos da Estatística, é permitir tomar decisões sobre populações, baseadas em informações como a média, variância, entre outras, que sejam resultantes dos dados amostrais dessas populações e que em Estatística se denotam por decisões estatísticas. E são usualmente usados coeficientes para perceber qual a relação entre esses dados amostrais, que permitem um melhor entendimento dos dados. Assim considere-se nos seguintes casos X como uma variável aleatória e k amostras, desta variável, de tamanho n_i , para a amostra i .

2.2.1 Coeficientes de Correlação

Dadas duas variáveis aleatórias X e Y com médias μ_X e μ_Y , a covariância entre as duas é dada por:

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] \quad (2.20)$$

Para medir o grau de associação entre as duas variáveis poder-se-ia usar a covariância, no entanto esta apresenta uma desvantagem, na medida em que depende das unidades das variáveis. De modo a que seja possível medir a associação entre variáveis mas usando uma medida que não dependa das unidades, neste caso que seja adimensional, são usados os coeficientes de correlação que, dependendo das variáveis em questão, ou das condições de aplicação, são diferentes.

Coeficiente de Correlação de Pearson

Este coeficiente mede a relação linear entre duas variáveis quantitativas, sendo que uma relação linear perfeita entre as duas variáveis apresenta um coeficiente igual a 1 ou a -1, onde -1 representa uma relação inversa, ou seja, quando uma variável cresce a outra decresce. No entanto, este coeficiente nada diz sobre outro tipo de associações, que não a relação linear. Para além, disso é um coeficiente sensível à presença de valores extremos, *outliers* (Sharma, 2012). O coeficiente é definido por:

$$\rho = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} \quad (2.21)$$

onde σ_X e σ_Y são os desvios padrão das variáveis X e Y respetivamente.

Para testar a hipótese nula onde o coeficiente de correlação de Pearson é igual a zero, é necessário que a distribuição do par das duas variáveis seja aproximadamente normal (Morgan et al., 2011), sendo este coeficiente uma estatística paramétrica.

Coeficiente de Correlação de Spearman

Este coeficiente de correlação pode ser usado quando pelo menos uma das variáveis é ordinal, ou seja, quando pode ser atribuída uma ordem ao conjunto de dados (Morgan et al., 2011), e não exige a normalidade da variáveis. Para o cálculo deste coeficiente são usadas as ordens desses valores. Este pode ser visto como um coeficiente de correlação na medida em que varia entre -1 e 1 e é interpretado da mesma maneira. A sua fórmula é dada por:

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (2.22)$$

em que d_i são as diferenças das ordens das variáveis X e Y .

2.2.2 Testes de Hipótese para a Homogeneidade de Variâncias

Assim como foi já referido, podem ser tomadas decisões sobre populações baseadas em informações das amostras. Para se tomar essas decisões, é necessário formular hipóteses que deverão ser testadas. A hipótese a ser testada é chamada de hipótese nula H_0 . Qualquer outra hipótese que defira dessa hipótese é chamada de hipótese alternativa H_1 . Os testes de hipótese, também designados por testes de significância, são os processos que nos permitem decidir rejeitar, ou não, uma hipótese.

Por exemplo, caso se queira testar a hipótese nula de igualdade de variâncias de k amostras, tem-se a hipótese H_0 :

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2 \quad (2.23)$$

e a hipótese alternativa $H_1 : \sigma_i^2 \neq \sigma_j^2$, para pelo menos um par (i, j) .

Um aspeto importante que deve ser tomado em consideração, é a normalidade das variáveis. Se houver evidência de que os dados provêm de uma variável que segue uma distribuição normal, então podem ser usados os testes de hipótese de Bartlett ou o teste de hipótese de F-Max de Hartley (Milliken and Johnson, 2009).

Teste de Hipótese de F-Max de Hartley

Para a aplicação deste teste todas as amostras devem ter o mesmo tamanho $n_i = n$ para qualquer $i = 1, \dots, k$. A estatística de teste é:

$$F_{max} = \frac{\max_i(s_i^2)}{\min_i(s_i^2)} \quad (2.24)$$

Esta estatística segue uma distribuição $F(v, t)$ e a hipótese nula é rejeitada se $F_{max} > F_{max, \alpha, v, t}$, onde $v = n - 1$ são os graus de liberdade associados a t variâncias individuais e s_i^2 é a variância da amostra i (Milliken and Johnson, 2009).

Teste de Hipótese de Bartlett

Este teste tem a vantagem de não ser necessário que as amostras tenham todas o mesmo tamanho (Milliken and Johnson, 2009). A estatística de teste é dada por:

$$U = \frac{1}{C} [v \log(s^2) - \sum_{i=1}^k v_i \log(s_i^2)] \quad (2.25)$$

onde $v_i = n_i - 1$, $v = \sum_{i=1}^k v_i$, $s^2 = \sum_{i=1}^k \frac{v_i s_i^2}{v}$ e $C = 1 + \frac{1}{3(k-1)} \sum_{i=1}^k (\frac{1}{v_i} - \frac{1}{v})$.

A estatística U segue uma distribuição $\chi^2(k-1)$ e a hipótese de igualdade de variâncias é rejeitada se $U > \chi_{\alpha, k-1}^2$ (Milliken and Johnson, 2009).

Como foi referido estes testes têm como pressuposto a normalidade dos dados e quando esta não é verificada devem ser usados outros testes que não sejam tão sensíveis à ausência da normalidade. Este é o caso do teste de hipótese de Levene e uma das suas variantes, o teste de hipótese de Brown - Forsythe (Ware et al., 2013).

Teste de Hipótese de Levene e Brown-Forsythe

Este teste tem como estatística de teste

$$F = \frac{N - k}{k - 1} \frac{\sum_{i=1}^k (\bar{d}_{i.} - \bar{d}_{..})^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} n_i (d_{ij} - \bar{d}_{i.})^2} \quad (2.26)$$

onde $d_{ij} = |x_{ij} - \bar{x}_{i.}|$, ou seja a diferença absoluta entre cada observação e a média do seu grupo para $i = 1, \dots, k$ e $j = 1, \dots, n_i$. Os valores $\bar{d}_{i.}$ são as médias dos grupos de d_{ij} , $\bar{d}_{..}$ é a média total dos valores d_{ij} , $N = \sum_{i=1}^k n_i$.

A estatística de teste segue uma distribuição $F(k - 1, N - k)$ e a hipótese nula é rejeitada, com um nível de significância α , quando $F > F_{\alpha, k-1, N-k}$ (Gastwirth et al., 2009).

Uma variante do teste de Levene é o teste de Brown-Forsythe, que é mais robusto no sentido em que usa a mediana em vez da média (Gastwirth et al., 2009), ou seja, onde em vez de d_{ij} usa $z_{ij} = |x_{ij} - \bar{x}_{med}|$.

O teste de Brown-Forsythe controla melhor o risco do que o teste de hipótese de Levene, de se poder declarar que as variâncias são diferentes, quando estas são de facto iguais, ou seja controla melhor um erro de tipo I - rejeitar uma hipótese nula verdadeira (Schlotzhauer, 2007).

2.3 Outros Testes de Hipótese Usados

Para além dos testes de hipótese usados para a homogeneidade das variâncias, outros testes são usados dependendo da hipótese que se quiser testar. É comum usarem-se os seguintes testes quando, depois de obtida a modelação das séries temporais, se está a verificar se os modelos em questão são ou não adequados para os dados.

Teste de Hipótese de Ljung - Box

Analicamente pode ser usado o teste de Ljung - box, para testar a correlação dos resíduos, cuja hipótese nula corresponde a que as primeiras L autocorrelações sejam estatisticamente nulas. A estatística do teste é dada por:

$$Q = N(N + 2) \sum_{k=1}^L \frac{\rho_k^2(\varepsilon)}{N - k} \quad (2.27)$$

onde N é o tamanho da amostra e $\rho_k^2(\varepsilon)$ é a autocorrelação amostral da série dos resíduos $\varepsilon\{t\}$ no desfasamento k. A estatística Q segue uma distribuição $\chi^2(L - p - q)$ (Wang, 2006) e rejeita a hipótese

nula, com um nível de significância α , quando $Q > \chi_{1-\alpha, L-p-q}^2$.

Teste de Shapiro - Wilk para a Normalidade

Este teste tem como hipótese nula a normalidade da população de onde veio a amostra e a estatística de teste é dada por:

$$W = \frac{\sum_{i=1}^k a_i (x_{n-i+1} - x_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.28)$$

onde n é o tamanho da amostra, $k = \frac{(n+1)}{2}$, se n for ímpar e $k = \frac{n}{2}$ se for par, os coeficientes a_i têm de ser consultados em tabelas próprias. A hipótese nula é rejeitada, com um nível de significância α , comparando o valor da estatística W obtido com os valores de uma tabela própria para este teste.

No capítulo seguinte é efetuada a aplicação destes conceitos na situação prática do trabalho em questão.

Capítulo 3

Descrição e Análise dos Dados

Foram trabalhados dados de três redes designadas por 1, 2 e 3.

3.1 Descrição dos Dados

De cada um dos contadores das três redes foi obtida a seguinte informação:

- **Fh** : - rede 1: tempo em segundos desde as 18h25 do dia 22 de Abril de 2015 até às 10h20 do dia seguinte, amostrado de 5 em 5 minutos, perfazendo um total de 192 tempos.
- rede 2: tempo em segundos desde as 12h do dia 11 de Novembro de 2015 até às 11h30 do dia 12 de Novembro de 2015, amostrados de 5 em 5 minutos, perfazendo um total de 95 tempos
- rede 3: tempo em segundos desde as 0h00 de 31 de agosto de 2015 até 23h45 do dia 23 de Setembro de 2015; dados diários a partir da 0h00, dos dias 19 de Julho, 12 de Agosto, 23 de Setembro, 11 de Outubro, 2 de Novembro e 25 de Dezembro, amostrados de 15 em 15 minutos, perfazendo um total de 96 tempos para cada dia do conjunto dos dados diários e 2304 tempos no outro conjunto.
- **Id2(t)** : Fornece o nome do posto de transformação e os nomes dos contadores.
- **Altot(t)** : total de energia ativa importada pelo contador para cada instante temporal Fh , ou seja, a energia ativa importada acumulada até ao momento em análise, em watt-hora.
- **Ali(t)** : incremento de energia ativa importada pelo contador em cada instante temporal Fh , ou seja, fornece a diferença entre $Altot(t + 1) - Altot(t)$, em watt-hora.
- **LAvgVLi(t)** : tensão média na fase i , para $i = 1, 2, 3$ do contador no Posto de Transformação, ou $i = a, b, c$ para cada uma das três fases diferentes mas desconhecidas, num contador trifásico, em volt (V).
- **LAvgV(t)** : tensão média num contador monofásico, correspondente a uma única fase desconhecida, em volt.

Os nomes dos contadores da rede 1 serão identificados e codificados como contador 1, 2, ..., 36 e contador no Posto de Transformação.

Os nomes dos contadores da rede 2 foram codificados da seguinte forma, para um contador trifásico acrescentou-se o sufixo .1, .2 e .3 ao respetivo contador, por exemplo se o contador 1 é trifásico passou a chamar-se de contador 1.1, contador 1.2, contador 1.3, os restantes que não forem trifásicos e o contador do Posto de Transformação mantém a mesma notação criada para a rede 1

Para a rede 3 foram fornecidos dados desde julho de 2015 até janeiro de 2016; deste conjunto foram escolhidos os dados referidos anteriormente que vão ser usados para testar os critérios escolhidos e validar a decisão, uma vez que se tem a solução real para o problema que pode ser consultada na secção 3.5. Para a criação de um algoritmo que fornecesse automaticamente a classificação, foram apenas usados os dados diários, este trabalho será depois desenvolvido na secção 3.6. Os nomes dos contadores foram codificados da mesma maneira que para a rede 1, uma vez que não existem contadores trifásicos, ou seja, contador 1, 2, 3, ..., 43 e contador no Posto de Transformação.

3.2 Análise dos Dados

Toda a análise referida nesta secção foi efetuada para os dados de tensão da rede 1.

Os dados de tensão fornecidos para esta rede encontravam-se em vários ficheiros xml, por esse motivo, foi construída uma única folha Excel que contivesse toda a sua informação. Para além disso foram construídos gráficos para análise do contador no Posto de Transformação e dos restantes contadores, uma vez que se pretende comparar os comportamentos das curvas de tensão do contador no posto de transformação e dos outros contadores.

Aqui apenas serão representados os gráficos das tensões das três fases do Posto de Transformação e de dois contadores, que representam os dois comportamentos diferentes observados, sendo que os restantes gráficos poderão ser consultados no Anexo A.

As variáveis que foram usadas no decorrer do trabalho foram as referentes à tensão média, nomeadamente, $LAvgVL1(t)$, $LAvgVL2(t)$, $LAvgVL3(t)$ e $LAvgV(t)$ e à variável temporal Fh . Como valores de referência tem-se os de tensão das três fases registadas do contador no Posto de Transformação, uma vez que são as únicas curvas de tensão das quais se sabe a fase.

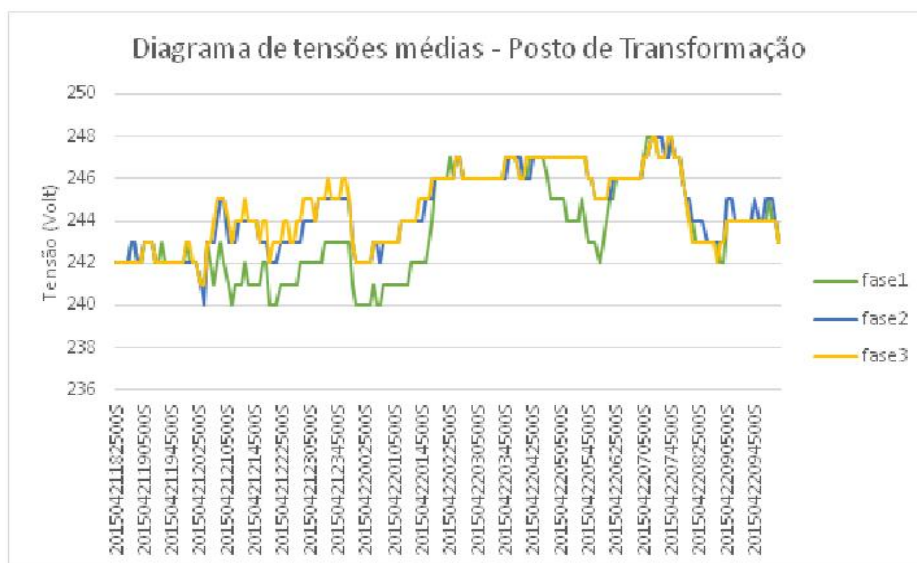


Figura 3.1: Dados da tensão média das três fases no Posto de Transformação

Na figura 3.1 pode-se ver que as curvas de tensão média das fases 2 e 3 no Posto de Transformação são muito parecidas tornando-se difícil a sua distinção e complicada a comparação da morfologia das curvas e associação de uma fase para cada um dos contadores trabalhados. O valor máximo dos dados das três tensões é de 248 V.



Figura 3.2: Dados da tensão média do contador 1 e das três fases do contador no Posto de Transformação

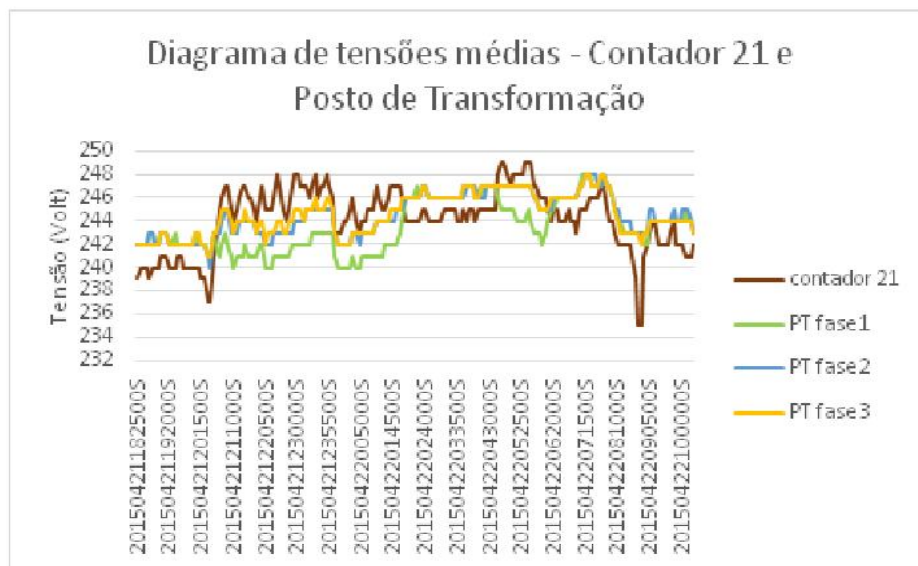


Figura 3.3: Dados da tensão média do contador 21 e das três fases do contador no Posto de Transformação

Quanto aos contadores foram registados dois tipos de comportamentos, que podem ser observados nas figuras 3.2 e 3.3. Na figura 3.2 a tensão média do contador 1, representada a castanho, encontra-se sempre abaixo da tensão de qualquer uma das três fases, tendo como valor máximo de tensão 248 V. Este é o comportamento normal da tensão média de qualquer contador, uma vez que ao longo do cabo elétrico existe decaimento de tensão, isto quer dizer que se a tensão começa com o valor máximo de 248 V numa das fases então, no mesmo instante de tempo em qualquer outro contador, o valor deve ser registado abaixo ou igual aos 248 V, com um erro máximo de 1 a 2 V. Mas, como se pode ver na figura 3.3 isso não acontece, o que pode representar erro no contador ou um erro de leitura por parte do contador.

O mesmo se passa em mais 19 contadores e como tal, a primeira análise foi unicamente feita aos contadores cuja tensão apresenta o normal funcionamento, de modo a não influenciar resultados.

Os contadores trabalhados foram: contador 1, 4, 5, 7, 8, 9, 10, 11, 15, 17, 22, 24, 28, 29, 32, 33 e 34.

Decorrida a análise aos contadores e não se tendo verificado nenhuma dependência entre os critérios e os contadores escolhidos, acabou depois por ser feita uma análise a todos os contadores.

Como já foi referido o objetivo é perceber com que fase é que a tensão registada por um determinado contador é mais relacionada. Neste sentido foram também construídos dois gráficos, que agrupassem determinados contadores do conjunto em cima escolhido cujas curvas de tensão média parecessem ter o mesmo comportamento; este agrupamento foi feito visualmente, mas permite perceber se os resultados diferem muito do que intuitivamente seria decidido.

Por uma questão de simplicidade o tempo foi considerado de zero até 192, ou seja, a hora 18h25 do dia 22 de Abril de 2015 corresponde agora ao valor 0 e a hora 10h20 do dia seguinte corresponde ao valor

192, não influenciando os resultados.

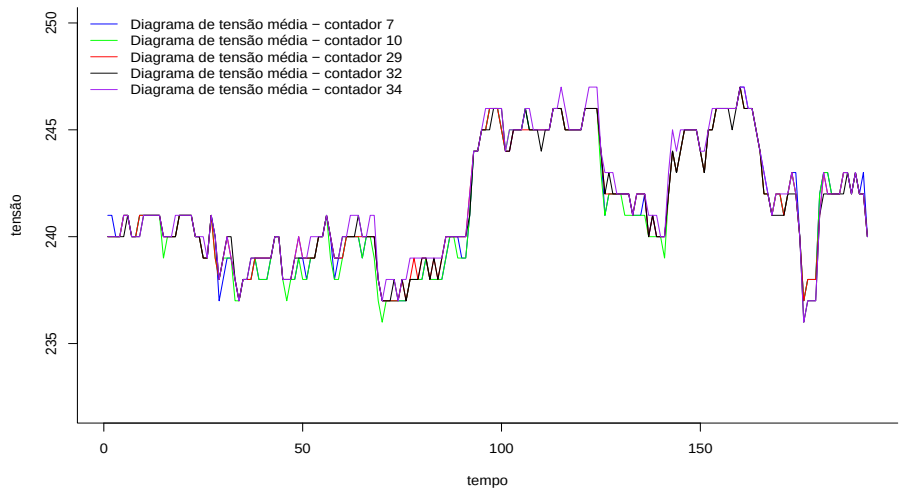


Figura 3.4: Agrupamento dos dados de tensão média dos contador 7, 10, 29, 32 e 34

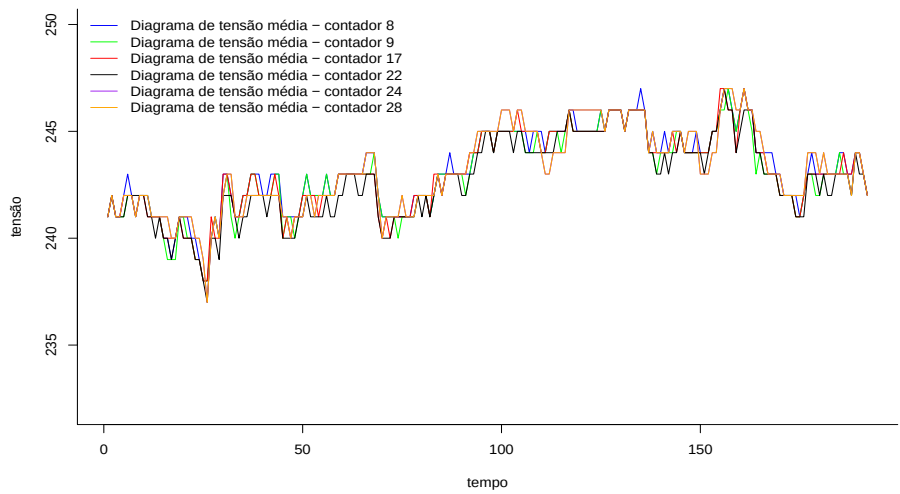


Figura 3.5: Agrupamento dos dados de tensão média dos contador 8, 9, 17, 22, 24 e 28

Pode-se ver na figura 3.4 que existe um momento temporal, por volta dos 180, que corresponde mais ou menos às 9h do dia 23, em que todos os contadores agrupados apresentam a mesma queda de tensão, onde em princípio se pode afirmar, não esquecendo que se está unicamente a usar informação visual, que, seja qual for a fase a que o contador 7 pertence, os outros quatro contadores também deverão pertencer. O mesmo se passa na figura 3.5 onde se vê a semelhança entres os contadores 8, 9, 17, 22, 24 e 28 e mais uma vez se pode dizer que estes devem pertencer todos à mesma fase, em princípio diferente da fase a que os contadores da figura 3.4 estiverem ligados.

3.3 Novos Conjuntos de Dados

Sabe-se que existindo decaimento de tensão, a comparação entre a tensão média registada no contador e as tensões registadas no contador do Posto de Transformação não deve ser feita comparando valores, mas sim variação, ou seja, se um contador estiver na fase 1, a curva da tensão registada no contador deve ser parecida com curva da tensão na fase 1 registada no contador do Posto de Transformação.

De modo a estudar melhor a relação morfológica entre as séries de tensões foram criadas três novas séries de tensão para cada contador, as quais se denominou de séries das diferenças, correspondendo às séries das diferenças entre as tensões em cada uma das três fases no Posto de Transformação e a tensão no contador a ser avaliado.

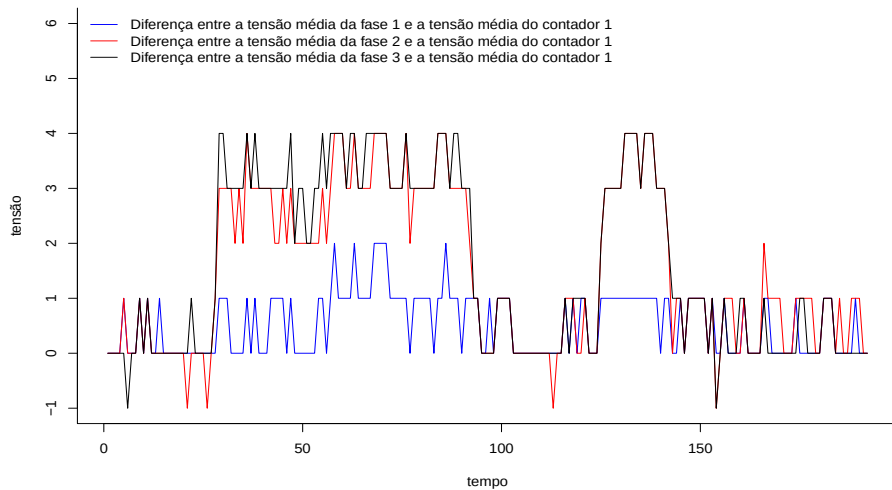


Figura 3.6: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 1

A figura 3.6 contém as séries das diferenças calculadas para o contador 1 e são um exemplo das séries que foram geradas para todos os outros contadores, que podem ser encontradas no Anexo B.

Caso um contador pertença a uma determinada fase, então a diferença entre a tensão dessa fase no Posto de Transformação e a registada no contador deveria ser próximo de uma constante a menos dos erros de medição (máximo 2 V), pois em teoria as duas curvas deverão ter o mesmo comportamento. Observando a figura 3.6 podemos ver que é a fase 1 que apresenta uma curva de tensão média mais constante, logo o contador 1 provavelmente está ligado à fase 1, no entanto na prática nem sempre esta decisão será assim tão simples, uma vez que esta clara distinção entre as três séries das diferenças não ocorre; por outro lado não é suficiente a observação gráfica, mesmo que esta pareça clara para se tirar uma conclusão definitiva, assim nas secções 3.4 e 3.5 serão abordadas as técnicas que foram usadas para esse fim.

Como já anteriormente foi referido, espera-se que a curva de tensão registada num contador tenha

as mesmas características morfológicas fundamentais que a curva de tensão de uma das fases registadas no contador do Posto de Transformação, o que implica que, se existem momentos em que a tensão das fases decai mais bruscamente e outros momentos em que a tensão vai aumentando, mas mais lentamente o mesmo irá acontecer à tensão registada no contador que estiver nessa fase. Para melhor poder comparar esta relação foram construídas para cada tempo t , séries contendo as derivadas instantâneas dos dados de tensão. Assim:

- Para a tensão da fase 1: $\frac{LAvgVL1(t+1) - LAvgVL1(t)}{Fh(t+1) - Fh(t)}$
- tensão da fase 2: $\frac{LAvgVL2(t+1) - LAvgVL2(t)}{Fh(t+1) - Fh(t)}$
- tensão da fase 3: $\frac{LAvgVL3(t+1) - LAvgVL3(t)}{Fh(t+1) - Fh(t)}$
- para a tensão de um contador monofásico: $\frac{LAvgV(t+1) - LAvgV(t)}{Fh(t+1) - Fh(t)}$

No entanto, uma vez que os dados são de cinco em cinco minutos e as derivadas vão ser sempre calculadas para valores seguidos de tensão, a diferença $Fh(t+1) - Fh(t)$ vai ser sempre igual a 5. Se se usar apenas a diferença entre tensão obtém-se uma série temporal cujos novos valores estão numa escala diferente da série das derivadas. Uma vez que só vai interessar comparar estes valores e não propriamente o seu valor numérico, decidiu-se trabalhar com o valor das diferenças, que ainda assim serão denominadas de séries das derivadas.

As novas séries são apresentadas nas figuras 3.7 e 3.8.

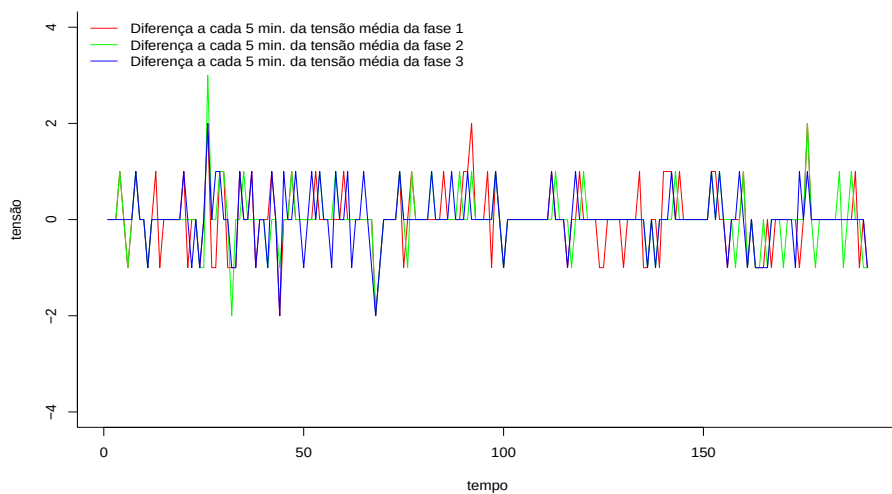


Figura 3.7: Séries que contêm as derivadas da tensão média das fases 1, 2 e 3 no Posto de Transformação

Os gráficos das derivadas para os outros contadores podem também ser consultados no Anexo B.

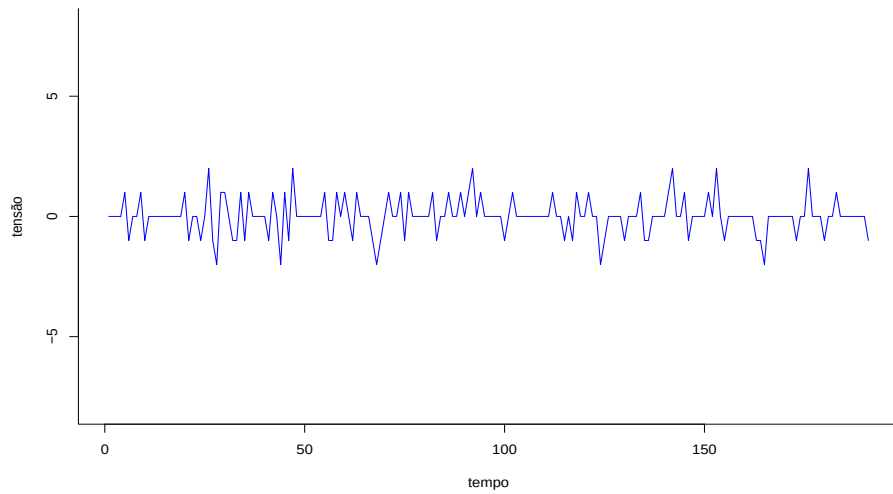


Figura 3.8: Série que contém as derivadas da tensão média do contador 1

Construídas estas séries, uma maneira de comparar se o comportamento da tensão registrada num contador é parecido com o comportamento registrado no contador do Posto de Transformação, é usar a mesma metodologia usada anteriormente e construir outras novas séries temporais que contenham a diferença das séries anteriores sempre entre o contador e as três fases do Posto de Transformação.

Assim, mais uma vez vão ser criadas três novas séries para cada contador que vão conter a diferença da série das derivadas, com uma outra escala, da tensão registrada no contador e a série das derivadas da tensão de cada uma das fases que foi registrada no contador do Posto de Transformação, que é a nossa referência. Estas séries são de agora em diante chamadas de séries das diferenças das derivadas. Para o contador 1 as séries são representadas na figura 3.9:

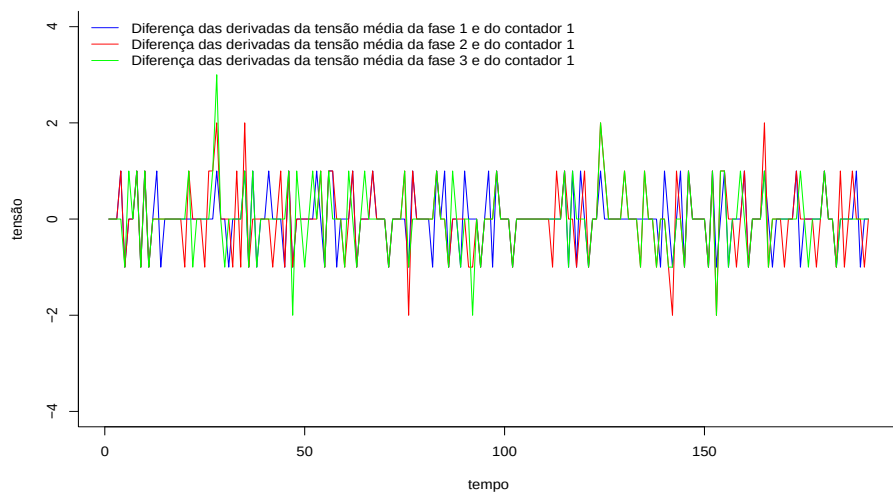


Figura 3.9: Série que contém as diferenças entre as derivadas da tensão média do contador 1 e as derivadas da tensão média das 3 fases no Posto de Transformação

As séries das diferenças das derivadas para os restantes contadores podem ser consultadas em Anexo B.

3.4 Seleção do Melhor Modelo

A primeira tentativa para o desenvolvimento de um método para atribuição de um contador a uma fase consistiu em obter um modelo ótimo para a tensão média das três fases do Posto de Transformação e estimar um modelo com a mesma estrutura para a tensão média dos contadores. Posteriormente, pretende-se efetuar uma análise estatística dos modelos para decidir qual o modelo que melhor se adequa à tensão média do contador. Isto permitirá criar um critério de escolha da possível fase para esse contador.

3.4.1 Melhor Modelo Para a Tensão Média da Fase 1

O primeiro passo foi analisar a ACF da série de tensão média da fase 1. A figura 3.10 mostra a representação gráfica dessa funções para diferentes desfasamentos.

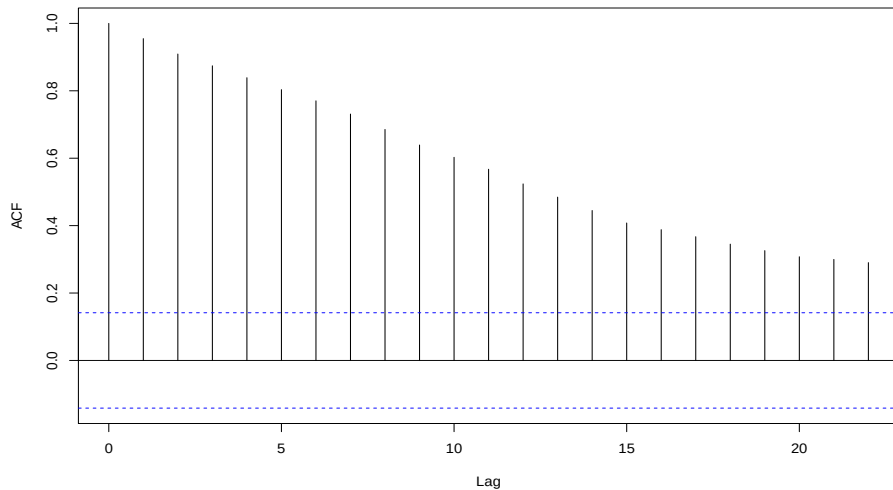


Figura 3.10: Função de autocorrelação da série da tensão média da fase 1 do Posto de Transformação

Observando a função de autocorrelação na figura 3.10 pode-se suspeitar que os dados não são estacionários, uma vez que para cada desfasamento se vê o valor da função acima das bandas de confiança e que estes decrescem muito lentamente para zero. Recorreu-se então à primeira transformação, retirar à série da fase 1 a tendência da mesma. Foram estimadas várias linhas de tendência, linear, exponencial, logarítmica, potencial e polinomial, que podem ser consultadas em Anexo C, tendo-se escolhido aquela que apresenta-se maior valor do coeficiente de determinação, r^2 , que indica o quanto o modelo consegue explicar melhor os valores observados.

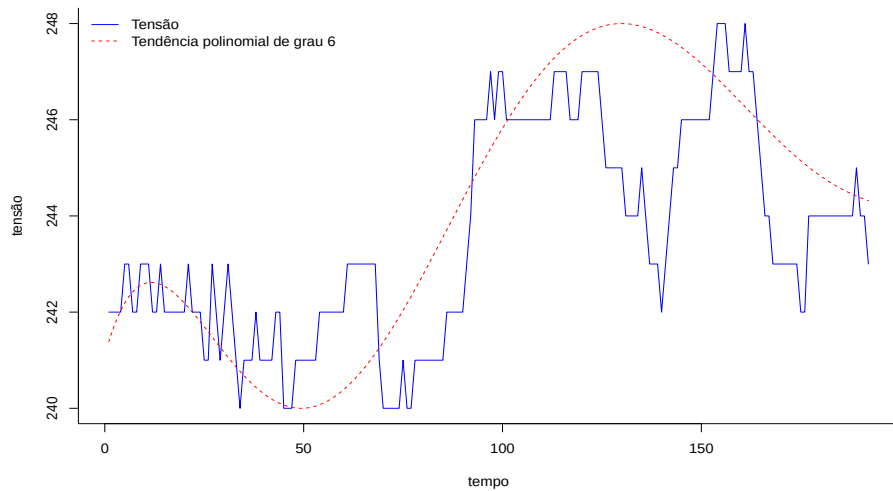


Figura 3.11: Tendência polinomial de grau 6 da série de tensão média da fase 1 do Posto de Transformação

A tendência polinomial de grau 6, figura 3.11 foi a que melhor se adequou à série; assim, foi feita a diferenciação da série subtraindo-lhe a tendência polinomial e analisou-se novamente a função de autocorrelação para perceber se a nova série seria aproximadamente estacionária ou não.

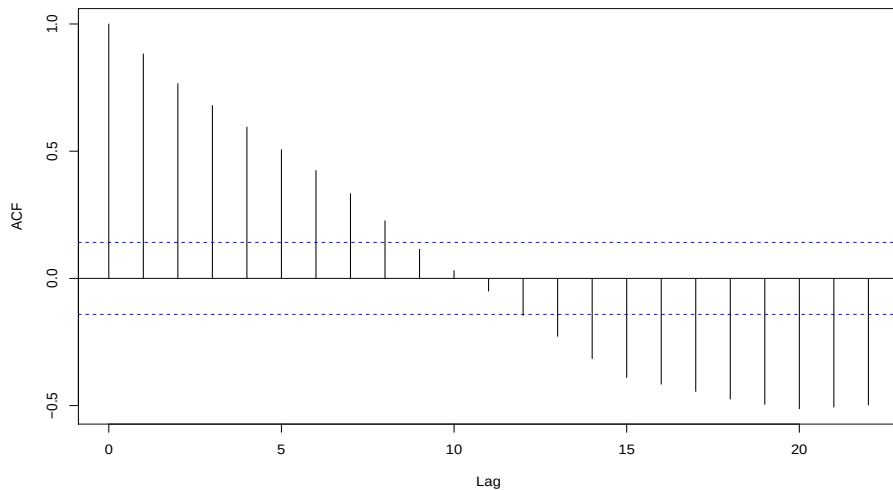


Figura 3.12: Função de autocorrelação da série de tensão média da fase 1 do Posto de Transformação à qual foi subtraída a tendência polinomial de grau 6

Pela análise gráfica da figura 3.12 pode-se ver que, apesar dos dados se encontrarem subtraídos da tendência, a função de autocorrelação desta nova série decai novamente lentamente para zero o que significa mais uma vez a não estacionaridade dos dados. Alternativamente considerou-se então a diferenciação a um passo da série inicial da tensão média da fase 1 do Posto de Transformação, que corresponde ao que foi chamada anteriormente de série de derivadas.

A sua função de autocorrelação pode ser vista na figura 3.13.

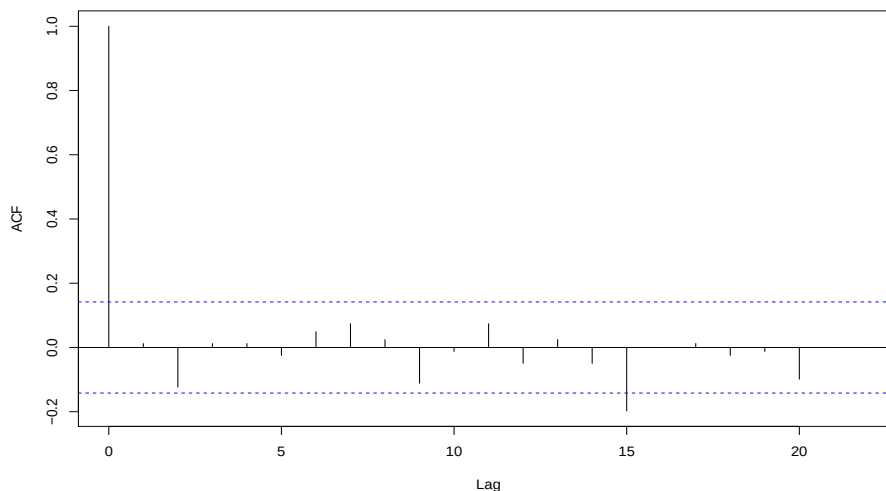


Figura 3.13: Função de autocorrelação da série de tensão média da fase 1 do Posto de Transformação diferenciada a um passo

Uma vez que a série pode ser considerada aproximadamente estacionária pode-se passar à etapa de identificação dos parâmetros p e q do modelo ARIMA $(p,1,q)$, onde $d = 1$ pois a série foi diferenciada uma vez, caso não se obtivesse a estacionariedade da série depois de diferenciada poder-se-ia ter de recorrer a uma nova diferenciação.

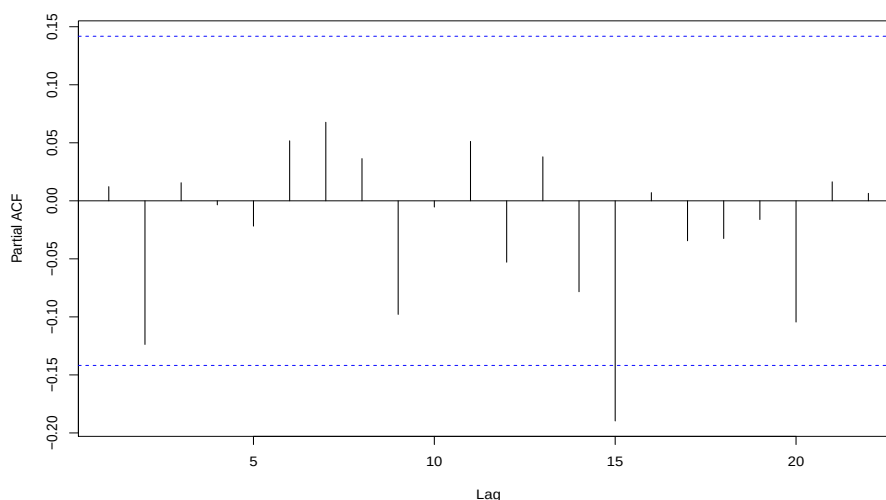


Figura 3.14: Função de autocorrelação parcial da série de tensão média da fase 1 do Posto de Transformação diferenciada a um passo

Pela análise das duas funções, a função de autocorrelação e autocorrelação parcial, das figuras 3.13 e 3.14, vemos que ambas decaem sinusoidalmente para zero o que significa que o melhor modelo a aplicar à série já diferenciada será um modelo ARMA e não um modelo AR ou MA (Cryer and Chan, 2008).

Uma vez que, a análise gráfica nem sempre é correta e suficiente para se justificar uma decisão, neste caso a escolha dos parâmetros do modelo, aliada ao facto de não se encontrarem referências com um suporte teórico sólido aos métodos gráficos de apoio à decisão do modelo, foram testadas várias ordens para os parâmetros e escolheu-se aquele que apresentasse melhor resultado, o modelo de menor valor de AIC. Para a identificação dos parâmetros em falta p e q , foram testados, para ambos, os valores de 0 a 20. É de notar que quando $p=0$ está a ser testado o modelo $MA(q)$ e quando $q=0$ então está a ser testado um modelo $AR(p)$.

O modelo obtido, com um menor valor de AIC, foi um modelo ARIMA (4,0,2), obtido da série diferenciada, cujo $AIC=382.38$.

Estimado o modelo é necessário validá-lo, ou seja, verificar se este modelo é um modelo que se adequa aos dados. Para isso foi feita uma análise visual e estatística aos resíduos do modelo.

Esta análise consistiu nos seguintes passos:

- Análise da função de autocorrelação dos resíduos;
- Teste de Ljung - Box para a correlação;
- Análise do histograma dos resíduos;
- Análise dos gráficos Q-Q plot;
- Teste de Shapiro para a normalidade.

Para testar a correlação dos resíduos foi efetuada a análise da função de autocorrelação dos resíduos e efetuado o teste de hipótese de Ljung - Box.

Para testar a normalidade foram calculados o histograma dos resíduos e o gráfico Q-Q plot que permitem ter uma noção visual da normalidade dos resíduos, a qual pode ser verificada através do teste estatístico de Shapiro-Wilk, cuja hipótese nula neste caso é a normalidade dos resíduos.

A função de autocorrelação dos resíduos está representada na figura seguinte:

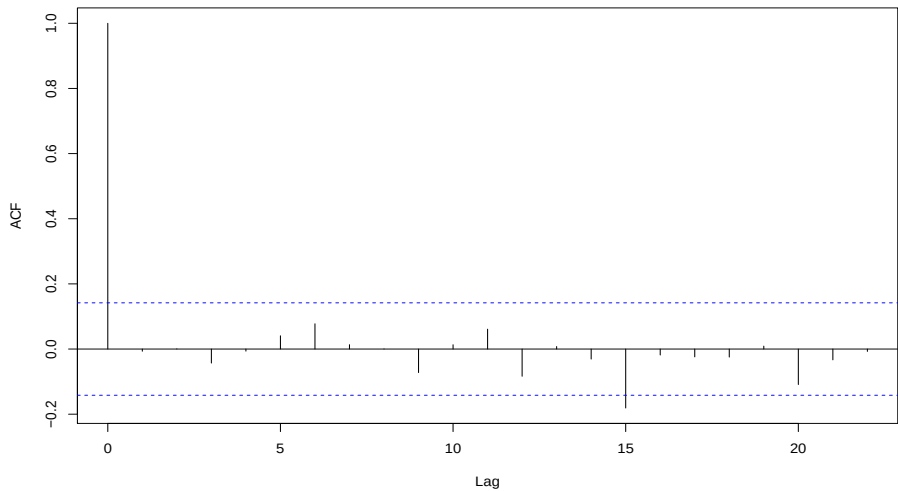


Figura 3.15: Função de autocorrelação dos resíduos do modelo para a fase 1 do Posto de Transformação

Os valores da ACF da figura 3.15 encontram-se quase todos dentro dos limites de confiança, tal indica a não correlação dos resíduos. O valor do teste de Ljung - Box encontra-se na tabela 3.1.

Teste para a Independência dos Resíduos	
Teste de Ljung - Box	valor p = 0.9777

Tabela 3.1: Aplicação do teste de Ljung - Box para testar a correlação dos resíduos

A hipótese nula de que as primeiras L correlações são estatisticamente nulas foi aceite uma vez que o valor p obtido é maior que o valor 0.05, que foi o nível de significância usado.

Assim, o pressuposto de não correlação dos resíduos foi verificado. Os resultados dos testes para os outros contadores pode ser consultado no Anexo C.

Quanto à normalidade dos mesmos seguem-se o histograma e o $Q - Q \text{ plot}$ nas figuras 3.16 e 3.17 respetivamente.

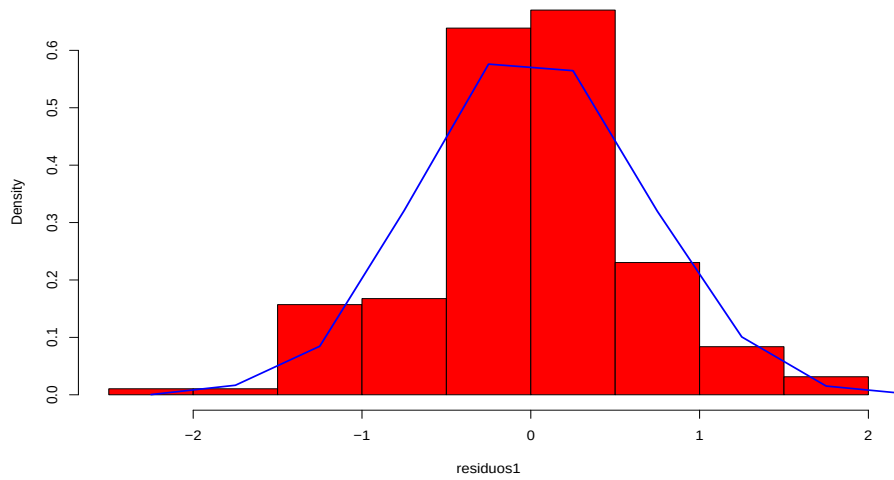


Figura 3.16: Histograma dos resíduos do modelo para a fase 1 do Posto de Transformação

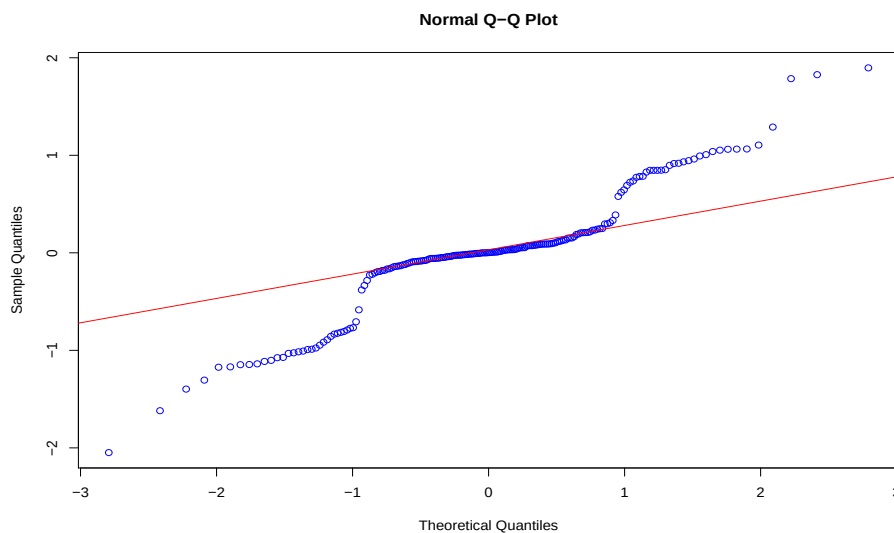


Figura 3.17: Q - Q plot dos resíduos do modelo da fase 1

Das figuras 3.16 e 3.17 pode-se ver que tanto o histograma como o gráfico Q-Q plot não parecem sugerir a normalidade dos resíduos, uma vez que, no histograma, existe uma grande quantidade dos dados que não se encontram debaixo da curva e no gráfico Q-Q plot, existe uma grande quantidade dos resíduos que não acompanham a linha reta.

O valor obtido do cálculo do teste de hipótese de Shapiro encontra-se na seguinte tabela:

Teste para a Normalidade dos Resíduos	
Teste de Shapiro - Wilk	Valor $p = 3.842e^{-8}$

Tabela 3.2: Aplicação do teste de Shapiro para a normalidade dos resíduos

A hipótese nula do teste de Shapiro-Wilk é rejeitada para uma significância de 5%, o que significa que com 95% de confiança os resíduos não são normalmente distribuídos tal como foi sugerido pelas indicações gráficas. No entanto, na prática basta que os resíduos sejam não correlacionados para que o modelo se diga adequado, (Wang, 2006).

Toda esta análise mostra que o modelo escolhido é adequado para os dados. O mesmo tipo de análise foi depois efetuada para as outras duas fases obtendo-se o modelo ARIMA (10,0,7) para a série temporal da fase 2 e o modelo ARIMA (8,0,4) para a série temporal da fase 3. Estes resultados podem ser consultados em Anexo C.

3.4.2 Estimação dos Coeficientes dos Modelos sobre as Tensões dos Contadores Monofásicos

O passo seguinte foi, a partir dos três modelos obtidos, para as tensões nas fases no Posto de Transformação escolher qual destes modela melhor a tensão média de cada contador.

Assim, por exemplo, para o contador 1 foram estimados os coeficientes dos modelos ARIMA (4,1,2), ARIMA (10,1,7) e ARIMA (8,1,4) que melhor modelam os dados não diferenciados da tensão média desse contador.

Obtendo-se assim três modelos candidatos para a modelação dos dados da tensão do contador 1. Para obter o melhor modelo, foi escolhido aquele que tivesse menor valor de AIC (Wang, 2006), desde que fosse considerado um modelo adequado, ou seja, desde que os resíduos apresentassem ser aleatórios e independentes.

AIC	
ARIMA(4,1,2)	434.9
ARIMA(10,1,7)	440.26
ARIMA(8,1,4)	442.92

Tabela 3.3: Valor do AIC

Destes três modelos obtidos, o modelo que apresenta menor valor de AIC, na tabela 3.3, é o modelo ARIMA (4,1,2), que tem a mesma estrutura que o modelo obtido para modelar a tensão da fase 1. Esta análise foi efetuada sem se verificar os próprios coeficientes obtidos dos modelos e escolheu-se apenas o modelo que apresentasse o mesmo número de coeficientes, que os modelos para as três fases, e menor valor de AIC. Logo, em princípio, o modelo escolhido para modelar a série do contador 1 será o modelo ARIMA (4,1,2).

Para verificar se de facto o modelo escolhido é um modelo adequado para a modelação da série deste contador, foi feita a mesma análise para a série dos resíduos. A função de autocorrelação para o modelo ARIMA (4,1,2) está representada na seguinte figura.

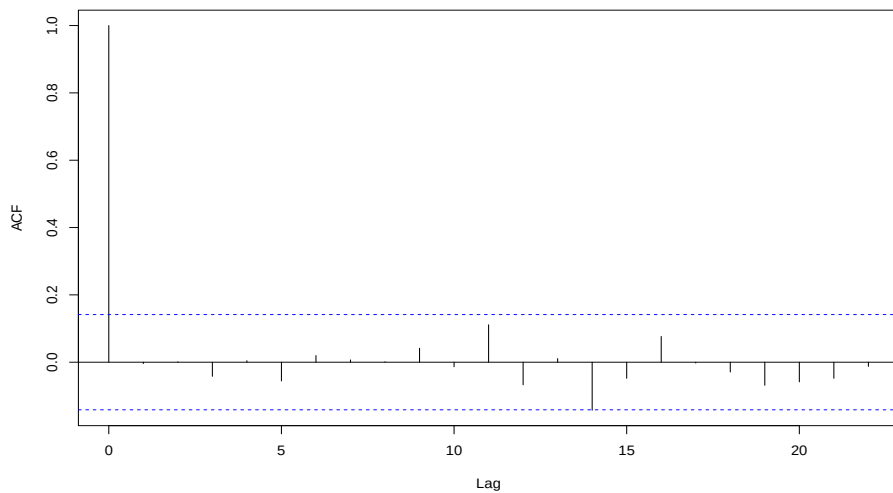


Figura 3.18: Função de autocorrelação dos resíduos do modelo ótimo para o contador 1

A função de autocorrelação dos resíduos do modelo da figura 3.18 indica a não correlação dos mesmos.

Teste para a Independência dos Resíduos do Modelo para o Contador 1

Modelo Ótimo	Teste de Ljung - Box
ARIMA(4,1,2)	valor p = 0.8184

Tabela 3.4: Aplicação do teste de Ljung - Box aos resíduos do modelo ótimo para o contador 1

Uma vez que a hipótese nula é aceite na tabela 3.4, o pressuposto da não correlação é verificado.

Os gráficos obtidos para a análise da normalidade dos resíduos foram os seguintes:

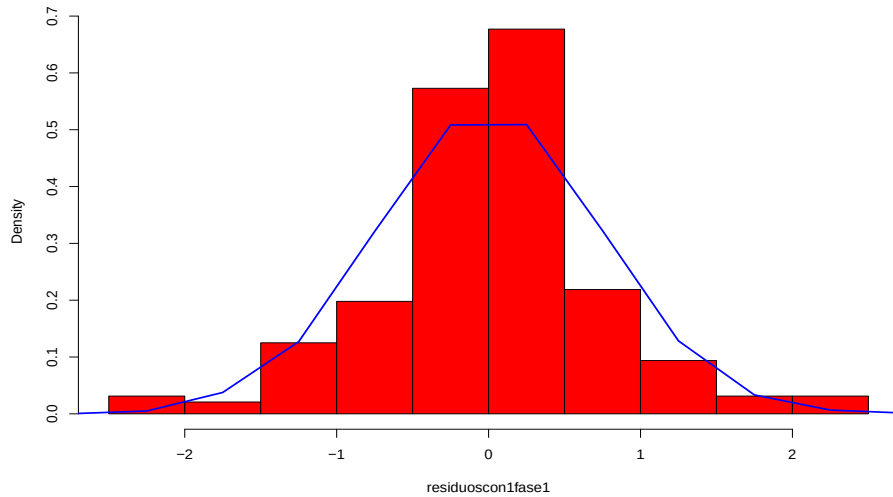


Figura 3.19: Histograma dos resíduos do modelo ótimo para o contador 1

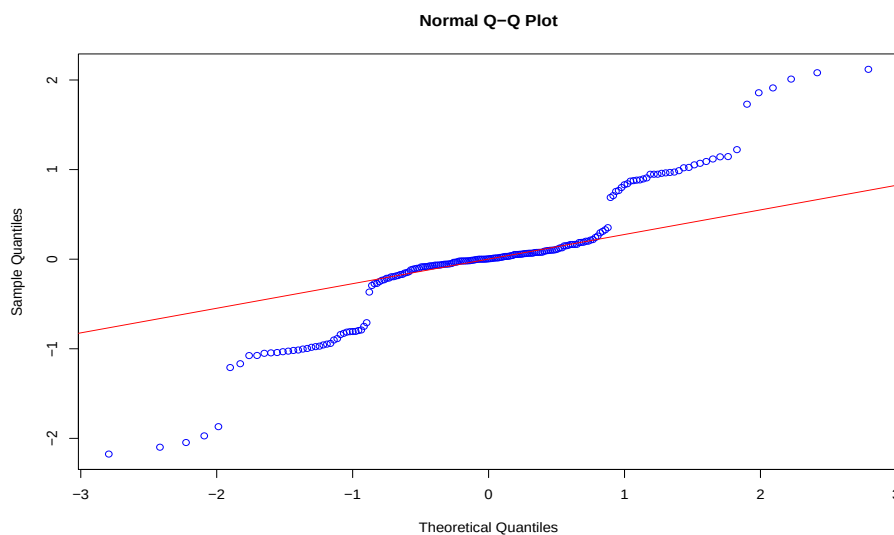


Figura 3.20: Q - Q plot dos resíduos do modelo ótimo para o contador 1

O gráfico do histograma, na figura 3.19 e o gráfico $Q - Q$ plot, na figura 3.20, parecem mais uma vez não indicar a normalidade dos resíduos.

Teste para a Normalidade dos Resíduos do Modelo para o Contador 1	
Modelo Ótimo	Teste de Shapiro
ARIMA(4,1,2)	valor $p = 2.188e^{-8}$

Tabela 3.5: Aplicação do teste de Shapiro aos resíduos do modelo ótimo para o contador 1

O valor do teste de Shapiro obtido da tabela indica que os resíduos não são normalmente distribuídos. Mas, mais uma vez, como estes são não correlacionados e na prática isso basta para que o modelo possa ser considerado adequado para os dados, temos que o modelo ARIMA que modelou a fase 1 é o que melhor se adequa a modelar a série de tensão do contador 1. Logo, de entre as três fases, a série da tensão da fase 1 é a que mais se assemelha à série da tensão do contador 1. Assim, podemos dizer que, usando o facto de termos a mesma estrutura de modelo como um critério de classificação, o contador 1 deverá estar ligado à fase 1.

3.5 Critérios, Resultados e Validação

3.5.1 Desvio Padrão para as Séries das Diferenças e para as Séries das Diferenças das Derivadas

Tomando a série das diferenças, como uma amostra de uma variável aleatória, calcularam-se as medidas de dispersão: desvio padrão e variância.

A tabela 3.6 apresenta essas medidas para os dados das três séries de diferenças do contador 1:

Diferença	Desvio Padrão	Variância
Diferença Fase 1	0.5693602	0.324171
Diferença Fase 2	1.46154	2.136098
Diferença Fase 3	1.596508	2.548838

Tabela 3.6: Medidas de Dispersão - desvio padrão e variância - para as Três Séries de Diferenças do Contador 1

Mais uma vez, o pretendido é fazer uma associação entre fases e contadores, para isso, das medidas calculadas na tabela 3.6, apenas o valor do desvio padrão vai ser analisado e a escolha irá ser feita observando unicamente esse valor para as três séries, o valor da variância será utilizado para testar a homogeneidade das variâncias obtidas, é de notar que, isto é equivalente a fazer uma análise única com o valor da variância obtido.

No Anexo D podem ser consultados os valores de desvio padrão e variância obtidos para cada contador.

Das três séries das diferenças de cada contador, aquela que apresentar o menor valor de desvio padrão

é a série mais constante. Quando isso acontece irá ser considerado que o contador pertence à fase em que isso se sucedeu. Isto porque, como foi explicado anteriormente, a morfologia da curva de tensão do contador e da curva de tensão da fase do Posto de Transformação a que ele pertence devem ser parecidas, o que significa que a diferença entre contador e a fase a que este está ligado deve ser o mais constante possível, ou seja a série do contador deve ter as mesmas oscilações no gráfico que a série da fase, se a série da fase desce o contador deve descer, se por outro lado sobe, a série do contador deve subir; se a série da fase apresenta picos elevados de decaimento de tensão então também a série do contador deve apresentar, a menos de um valor, que deveria ser constante mas que tal não se verifica na prática, devido ao decaimento de tensão ao longo do fio.

Ainda que na prática isso não aconteça, é de supor que esse valor deverá ser mais constante para a fase a que o contador estiver ligado do que para as outras duas. E para que isto aconteça, os valores da série da diferença devem dispersar o menor valor possível para a média, ou seja, apresentar o menor valor de desvio padrão. De facto, num caso extremo, em que se obtivesse uma série cujo desvio padrão fosse 0 e a sua média 1, então significava que o contador apenas se encontrava 1 Volt acima da fase em que tal acontecia, e como tal deveria estar ligado a essa fase.

Dos três valores de desvio padrão obtidos na tabela 3.6 o menor é obtido para a primeira série, ou seja, é nessa série que os seus valores estão menos dispersados em relação à média e portanto a série é mais constante, o que significa que, o contador 1, segundo o critério escolhido, está ligado à fase 1.

No entanto, pode acontecer que, mesmo que o menor valor do desvio padrão da série seja diferente dos outros dois, este possa ser estatisticamente igual a pelo menos um dos outros dois, pelo facto de duas amostras diferentes serem provenientes da mesma variável aleatória.

É então necessário aplicar um teste para se poder perceber se isto acontece ou não. De entre vários testes de hipótese, que testam a homogeneidade de variâncias, o teste de hipótese de Levene, que ao contrário de outros usados para o mesmo efeito não tem como pressuposto que a variável aleatória da qual se retirou a amostra segue uma distribuição normal e, como tal, pode ser aplicado sem ter de se verificar o pressuposto (Ritz and Streibig, 2008). Este teste de hipótese usa a média da amostra, no entanto prova-se que ao usar a mediana, se obtém um teste mais robusto; a esse teste foi dado o nome de teste de Brown-Forsythe (Ware et al., 2013), sendo este uma variante do teste de Levene, com a única diferença de se usar a mediana.

Neste trabalho foi usada a variante do teste de Levene, o teste de Brown-Forsythe. Como se tem três conjuntos de dados, quatro situações podem acontecer:

- a variância da série das diferenças para a fase 1 pode ser igual à série da diferença para a fase 2;

- a variância de série das diferenças para a fase 1 pode ser igual à série da diferença para a fase 3;
- a variância da série das diferenças para a fase 2 pode ser igual à série da diferença para a fase 3;
- todas as variâncias podem ser iguais.

Deste modo foram calculados quatro testes de hipóteses, um para cada uma destas situações, com a particularidade de que se as variâncias fossem consideradas todas iguais então não é necessário calcular os outros três testes porque iria de encontro ao resultado obtido.

Quando a hipótese nula é rejeitada, com um nível de significância de 5%, rejeita-se a hipótese de que as variâncias possam ser iguais.

O conjunto de contadores representado na tabela 3.7 foi escolhido aleatoriamente. Destes foi obtido que os contadores 5, 7, 10 e 29 estão ligados à fase 1 e o contador 17 à fase 2. Os resultados para estes e para os outros contadores podem ser consultados no Anexo D.

Contador	Todas as fases	Fase 1 e 2	Fase 1 e 3	Fase 2 e 3
Contador 1	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.1022$
Contador 5	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.06659$
Contador 7	valor $p < 2.2e^{-16}$	valor $p = 2.227e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.4781$
Contador 10	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.4221$
Contador 17	valor $p = 7.034e^{-12}$	valor $p = 7.809e^{-11}$	valor $p = 1.463e^{-06}$	valor $p = 0.03811$
Contador 29	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.3645$

Tabela 3.7: Valor do teste de hipótese de Brown-Forsythe para as séries das diferenças

Observando os valores do teste para contador 1 na tabela 3.7 apenas a hipótese nula que testa a igualdade da variância da série da diferença de tensão para a fase 2 e a variância da série da diferença de tensão para a fase 3 é que são consideradas estatisticamente iguais. Logo, pelo critério escolhido, o contador 1 encontra-se ligado à fase 1. Para os contadores 5, 7, 10 e 29 o critério indica que estes devem pertencer à fase 1 e o contador 17 de pertencer à fase 2, uma vez que não foi observada igualdade de variâncias entre a fase escolhida e outra(s).

A tabela que contém os valores do teste para todos os contadores pode ser consultada em Anexo D.

A mesma análise foi depois efetuada para séries das diferenças das derivadas construídas na secção 3.3. Tendo-se definido um outro critério, o valor mínimo do desvio padrão das três séries das diferenças das derivadas de cada contador.

Assim, para o contador 1 é apresentada a tabela 3.8.

Diferença das Derivadas	Desvio Padrão	Variância
Fase 1	0.598243	0.3578947
Fase 2	0.7181848	0.5157895
Fase 3	0.672779	0.4526316

Tabela 3.8: Medidas de dispersão - desvio padrão e variância - para as séries de diferenças das derivadas do contador 1

Segundo a tabela 3.8, sendo o critério escolhido o do menor desvio padrão, o contador 1 está ligado à fase 1 tendo sido classificado com a mesma fase usando o critério do menor desvio padrão da série das diferenças. Quanto aos contadores 5, 7, 10 e 29 foram classificados como pertencentes à fase 1, os valores das medidas de dispersão de todos os contadores e a sua classificação usando este critério podem ser consultados no Anexo D. O contador 17 obteve a classificação fase 3, diferente da obtida com o critério anterior.

A tabela seguinte apresenta o valor do teste de hipótese de Brown-Forsythe para estes contadores. Para os contadores em que o teste de hipótese que testa a hipótese nula da igualdade das três variâncias foi aceite, não foi efetuado nenhum teste para as outras hipóteses.

Contador	Todas as fases	Fase 1 e 2	Fase 1 e 3	Fase 2 e 3
Contador 1	valor p = 0.2833	-	-	-
Contador 5	valor p = 0.8021	-	-	-
Contador 7	valor p = 0.184	-	-	-
Contador 10	valor p = 0.09754	-	-	-
Contador 17	valor p = 0.7377	-	-	-
Contador 29	valor p = 0.04725	valor p = 0.01604	valor p = 0.06982	valor p = 0.5418

Tabela 3.9: Valor do teste de hipótese de Brown-Forsythe para as séries das diferenças das derivadas

O contador 1 foi classificado como pertencente à fase 1, pois era para essa fase que apresentava o menor valor do desvio padrão, no entanto o valor p obtido da tabela 3.9 correspondente a esse contador, mostram que as variâncias são todas iguais e, como tal, uma decisão tomada pelo menor desvio padrão neste caso não é correta, uma vez que a diferença entre os desvios padrão não é significativa.

3.5.2 Coeficiente de Correlação de Spearman

Uma vez que, os valores de tensão dos contadores e do Posto de Transformação se encontram diretamente relacionados, parece útil perceber qual o grau de associação entre estas variáveis. Para medir a correlação entre variáveis são usados os coeficientes de correlação que, tal como foi referido na secção

2.2.1, dependendo das variáveis em questão, ou das condições de aplicação, são diferentes.

Para medir o grau de associação entre as variáveis, foi usado o coeficiente de correlação de Spearman, que é um coeficiente que não assume hipóteses sobre a distribuição do par de variáveis, ou seja, é uma estatística não paramétrica (Morgan et al., 2011).

O coeficiente de correlação de Spearman foi calculado entre os valores de tensão do contador e os valores de tensão das fases do Posto de Transformação, obtendo-se três valores para cada contador. A tabela seguinte apresenta esses valores para um conjunto aleatório de contadores.

Os valores obtidos do cálculo do coeficiente de correlação de Spearman para todos os contadores podem ser consultados no Anexo D, na tabela 3.10 apresentam-se, como exemplo, os resultados de apenas seis contadores.

Contador	Fase 1	Fase 2	Fase 3
Contador 1	0.9658074	0.763069	0.6990243
Contador 5	0.9698603	0.8055055	0.745909
Contador 7	0.9279967	0.7043779	0.6564593
Contador 10	0.9275524	0.6968279	0.648839
Contador 17	0.8444933	0.9509836	0.9162201
Contador 29	0.9354229	0.7357393	0.6895225

Tabela 3.10: Coeficiente de correlação de Spearman entre os valores de tensão das três fases do posto de transformação e os valores de tensão dos contadores

Tomemos, por exemplo, o contador 1 da tabela (3.10), pode-se ver que neste caso os valores da tensão do contador são mais correlacionados com os valores da tensão da fase 1 do Posto de Transformação, uma vez que é o maior dos três valores. Tem-se, então, que o comportamento da curva de tensão do contador varia com grande semelhança ao decrescimento da curva de tensão da fase 1 do Posto de Transformação, pois o valor do coeficiente é positivo, o que significa que ou ambas as curvas se encontram em crescimento, ou em decrescimento, para além disso o valor elevado significa uma forte associação entre ambas as curvas, ou seja, a curva de tensão do contador apresenta o mesmo comportamento que a curva de tensão da fase 1 do Posto de Transformação, isto indicia que este contador pertence a essa fase. O mesmo se pode dizer do contador 5 ou do contador 7.

No entanto verificamos que há valores na tabela 3.10 cujos valores de correlação entre os valores de tensão do contador e os valores de tensão das fases não diferem muito, como é o caso do contador 17, onde se obteve as correlações com os valores de tensão da fase 2 e da fase 3 muito próximos de 0.9. Neste caso é difícil de decidir a qual poderá pertencer o respetivo contador.

Assim, se houver evidência de que um dos valores do coeficiente de correlação de Spearman é maior que os outros dois, o que significa uma maior relação do contador com a respetiva fase em que isso acontece, esse contador será atribuído a essa fase. Para cada contador foi então escolhida a fase a que pertence, baseado apenas no facto de ter o maior valor obtido para esse coeficiente.

Há no entanto que garantir que o valor do coeficiente tem significado estatístico, ou seja, que testando a hipótese nula do coeficiente ser igual a zero, que esta hipótese é rejeitada.

Assim este teste foi apenas calculado para o valor máximo obtido entre os valores de tensão do contador e os valores de tensão de cada uma das fases do Posto de Transformação. A hipótese nula foi rejeitada para todos os contadores e a tabela com os valores pode ser consultada em Anexo D. Assim o valor obtido do coeficiente tem significado e como tal pode ser usado para concluir com qual das três fases a curva de tensão do contador melhor se encontra mais relacionada e portanto é mais parecida morfológicamente.

Outro critério proposto consistiu na mesma análise do cálculo do coeficiente de correlação de Spearman às séries das derivadas da tensão do contador e das derivadas da tensão de cada uma das respetivas fases, sendo que a descrição da construção das séries se encontra na secção 3.3.

A tabela 3.11 apresenta os valores do coeficiente para o mesmo conjunto de contadores anteriormente escolhido. O valor p, calculado apenas para o maior valor da correlação, foi sempre inferior a 0.05 e pode ser consultado no Anexo D. Quanto aos valores para os outros contadores também estes podem ser consultados no Anexo D.

Contador	Fase 1	Fase 2	Fase 3
Contador 1	0.5753444	0.3735381	0.4851053
Contador 5	0.5529131	0.4710801	0.5053344
Contador 7	0.5724154	0.2697296	0.4227629
Contador 10	0.5438097	0.2147425	0.3631805
Contador 17	0.4520803	0.4433206	0.4790014
Contador 29	0.6212455	0.3189702	0.41157

Tabela 3.11: Coeficiente de correlação de Spearman entre os valores das séries das derivadas das três fases do posto de transformação e os valores das derivadas dos contadores

Na tabela 3.11 a distinção que antes era evidente no contador 1, 5, ou até o contador 7 já não é assim tão proeminente, nem os valores são assim tão elevados tendo diminuído quase 0.3, mas ainda assim o maior valor é na mesma obtido para a correlação com a fase 1 do Posto de Transformação. Note-se que a classificação de outros contadores mudou, como é o caso do contador 17, que na primeira situação sugere pertencer à fase 2 e na segunda à fase 3.

3.5.3 Critérios Propostos

Assim os critérios escolhidos foram:

- Mínimo do desvio padrão para as série das diferenças;
- Mínimo do desvio padrão para as série das diferenças das derivadas;
- Máximo do coeficiente de correlação de Spearman para as séries de tensão;
- Máximo do coeficiente de correlação de Spearman para as séries das derivadas;
- Modelo ARIMA.

Dos critérios escolhidos resultou a seguinte classificação para os contadores.

Resultados dos Critérios					
Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman	Mínimo do Desvio Padrão das Derivadas	Máximo do Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 1	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 5	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 7	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 10	*Fase 1	*Fase 1	*Fase 1	*Fase 1	*Fase 1
Contador 17	*Fase 2	*Fase 2	Fase 3	*Fase 3	*Fase 3
Contador 29	*Fase 1	*Fase 1	*Fase 1	*Fase 1	*Fase 1

Tabela 3.12: Classificação obtida em cada um dos critérios trabalhados para a rede 1

As classificações marcadas com um * são as classificações cujo resultado tem significância, ou seja, cujas hipóteses nulas foram rejeitadas ou aceites conforme o teste para os respetivos critérios; quanto à classificação obtida pelo estudo feito usando um modelo ARIMA apresenta * uma vez que estas foram as classificações obtidas depois de decorrida toda a análise para a escolha do melhor modelo e que pode ser consultada na secção 3.4.

Os resultados para os restantes contadores podem ser consultadas no Anexo D. Da tabela dos resultados para todos os contadores pode ser observado que o critério que mais discorda com as restantes classificações é o do modelo ARIMA ótimo, sendo que os outros mantêm a concordância das fases para quase todos os contadores, à exceção de 11 contadores em que o coeficiente de correlação de Spearman

das derivadas não apresenta uma classificação igual às outras. Estas conclusões são observáveis quando são tomados todos os contadores em consideração. De modo a ser obtida uma classificação única e final, usando os critérios escolhidos, foi construído um sistema de votação, atribuindo o mesmo peso a cada critério. Deste modo para os mesmos contadores foi obtida a tabela 3.13.

Sistema de Votação	
Contador	Nº de votos
Contador 1	Fase 1: 5
	Fase 2: 0
	Fase 3: 0
Contador 5	Fase 1: 5
	Fase 2: 0
	Fase 3: 0
Contador 7	Fase 1: 5
	Fase 2: 0
	Fase 3: 0
Contador 10	Fase 1: 5
	Fase 2: 0
	Fase 3: 0
Contador 17	Fase 1: 0
	Fase 2: 2
	Fase 3: 3
Contador 29	Fase 1: 5
	Fase 2: 0
	Fase 3: 0

Tabela 3.13: Classificação obtida do sistema de votação para a rede 1

A tabela que contém o sistema de votação para todos os contadores pode ser consultada no Anexo D.

3.5.4 Aplicação dos Critérios e Classificação

Para efeitos de validação, foi feita a mesma análise dividindo o intervalo temporal ao meio, criando-se duas novas séries uma com a tensão média das 18h25 até às 2h20, chamada de série da primeira metade e outra com a tensão média das 2h25 até às 10h20, chamada de série da segunda metade, cada uma com 96 tempos, de modo a verificar qual a classificação obtida em dois novos conjuntos de dados que desta vez apresentem menos valores. As tabelas 3.14 e 3.15 contêm os resultados obtidos dos critérios aplicados à primeira e segunda metade respetivamente. A classificação dos restantes contadores pode ser consultada no Anexo D.

Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman Máximo	Mínimo do Desvio Padrão das Derivadas	Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 1	*Fase 1	*Fase 1	Fase 1	Fase 1	*Fase 1
Contador 5	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 7	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 10	* Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 17	Fase 2	*Fase 2	Fase 2 (Fase 3)	*Fase 3	*Fase 1 (Fase 3)
Contador 29	*Fase 1	*Fase 1	*Fase 1	*Fase 1	*Fase 1

Tabela 3.14: Classificação obtida dos 5 critérios dos dados das 18h25 até às 2h20 do primeiro conjunto de dados

Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman Máximo	Mínimo do Desvio Padrão das Derivadas	Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 1	*Fase 1	*Fase 1	*Fase 1	Fase 1	*Fase 1
Contador 5	*Fase 1	*Fase 1	*Fase 1	Fase 1	*Fase 1
Contador 7	*Fase 1	*Fase 1	*Fase 1	Fase 1	*Fase 1
Contador 10	*Fase 1	*Fase 1	*Fase 1	Fase 1	*Fase 1
Contador 17	*Fase 2	Fase 2	*Fase 3	Fase 3	*Fase 1 (Fase 3)
Contador 29	*Fase 1	*Fase 1	*Fase 1	Fase 1	*Fase 1

Tabela 3.15: Classificação obtida dos 5 critérios dos dados das 2h25 até às 10h20 do primeiro conjunto de dados

Entre parênteses encontra-se a solução obtida da aplicação dos critérios à rede 1 inteira.

Os dois primeiros critérios são os que se mantêm constantes na classificação dos contadores, já o critério do modelo ARIMA é o que em mais vezes ocorreu alteração na classificação, isto pode ser melhor observado na tabela que contém os resultados para todos os contadores no Anexo D. Da informação obtida, este último critério parece ser o que pior classifica os dados. Usando novamente o sistema de votação obteve-se as classificações da tabela 3.16.

Sistema de Votação		
Contador	Nº de votos da primeira metade	Nº de Votos da segunda metade
Contador 1	Fase 1: 5	Fase 1: 5
	Fase 2: 0	Fase 2: 0
	Fase 3: 0	Fase 3: 0
Contador 5	Fase 1: 5	Fase 1: 5
	Fase 2: 0	Fase 2: 0
	Fase 3: 0	Fase 3: 0
Contador 7	Fase 1: 5	Fase 1: 5
	Fase 2: 0	Fase 2: 0
	Fase 3: 0	Fase 3: 0
Contador 10	Fase 1: 5	Fase 1: 5
	Fase 2: 0	Fase 2: 0
	Fase 3: 0	Fase 3: 0
Contador 17	Fase 1: 1	Fase 1: 1
	Fase 2: 3	Fase 2: 2
	Fase 3: 1	Fase 3: 2
Contador 29	Fase 1: 5	Fase 1: 5
	Fase 2: 0	Fase 2: 0
	Fase 3: 0	Fase 3: 0

Tabela 3.16: Classificação obtida do sistema de votação para as duas metades da rede 1

Para a rede 2 existe, como já foi referido anteriormente, uma classificação visual para os contadores, e por isso esta rede foi usada para a escolha do(s) critério(s) que melhor classificam os dados. Só existe referência da classificação para 12 dos 16 contadores, no entanto, uma vez que a classificação de uma fase de um contador trifásico conta como a classificação efetuada para um contador individual, isto significa que ao todo foram efetuadas 30 classificações em 34. A percentagem da classificação acertada encontra-se

na seguinte tabela.

	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman	Mínimo do Desvio Padrão das Derivadas	Máximo do Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Percentagem de Acertos	≈ 59%	≈ 56 %	≈ 93%	100%	≈ 43 %

Tabela 3.17: Percentagem de acertos da classificação obtida dos 5 critérios dos dados da rede 2

A tabela com as classificações obtidas para os contadores pode ser consultada no Anexo D. Depois de efetuada esta análise houve um critério que se destacou com 100% na classificação correta das fases, o máximo do coeficiente de correlação de Spearman para as séries das derivadas, sendo assim, este método vai constituir a proposta de método automático a implementar.

Note-se que estes resultados indiciam que o critério selecionado é equivalente à identificação manual e empírica, por engenheiros especializados, constituindo uma alternativa automática e mais barata.

3.5.5 Validação para a Rede 3

Escolhido o critério que ia ser trabalhado foi feita a classificação para a rede 3 através da aplicação do critério. Esta rede foi fornecida juntamente com a sua correta classificação, o que permitiu a validação final deste método. No entanto, apenas foi fornecida a classificação de 33 contadores de 43 na referência e, portanto, a percentagem de acertos calculada foi apenas para esses contadores.

Resultados dos Critérios para a rede 3	
	Máximo do Coeficiente de Correlação de Spearman das Derivadas
Percentagem de Acertos	≈ 82%

Tabela 3.18: Classificação obtida do critério escolhido aplicado aos dados da rede 3

O resultado da tabela 3.18 foi obtido da aplicação do critério a várias estâncias temporais da rede, referidas na secção 3.1. O que quer dizer que é robusto no que toca à deteção da fase para diferentes períodos de tempo da mesma rede. Uma percentagem de erro de 82% continua a ser uma boa percentagem. Para uma análise mais aprofundada do critério escolhido e de modo a perceber se de facto o critério

escolhido é ou não um bom critério, deveriam ser estudadas outras redes de distribuição diferentes das quais houvesse uma referência e, que por outro lado, apresentassem mais contadores com referência.

3.6 Elementos Computacionais

Nesta secção apenas serão apresentados os passos principais da implementação de um algoritmo, escrito em R e as dificuldades encontradas durante o seu desenvolvimento, devido a questões de confidencialidade. O algoritmo resolve o problema que se tinha em mãos de classificação dos contadores numa rede de distribuição de baixa tensão. E permite ao utilizador neste caso a empresa de obter a classificação de todos os contadores da rede automaticamente.

Assim escolhido o critério e as séries necessárias para a sua aplicação, nomeadamente as séries das derivadas, foi pedido que fosse criado um programa que, automaticamente, dado um qualquer conjunto de dados fornecesse a fase dos contadores para a rede em questão.

O programa segue os seguintes passos:

- são eliminados os erros de cada um dos ficheiros de dados fornecidos;
- são procurados, em todos os ficheiros exceto o primeiro e acrescentados, no primeiro ficheiro, os dados dos contadores que tiverem equivalência no primeiro ficheiro;
- são eliminados os contadores, no primeiro ficheiro, que não tiveram nenhuma equivalência;
- são eliminados os dados que estiverem a mais, ou seja, dados que existem em alguns contadores mas não noutros, uma vez que é necessário o mesmo tamanho da amostra dos contadores para o cálculo do coeficiente de correlação de Spearman;
- é criada uma tabela com os dados de tensão média dos contadores por coluna, obtendo-se assim uma coluna de dados para cada contador;
- é criada uma outra tabela que contenha os valores das derivadas;
- é calculado o coeficiente de Correlação de Spearman entre as derivadas das fases do Posto de Transformação e as derivadas do contador;
- os contadores são classificados.

Neste processo foram encontradas algumas dificuldades:

- primeiro por norma os dados vem apresentados em mais do que um ficheiro de dados e como tal o programa foi criado para ler mais do que um ficheiro ao mesmo tempo;

- Os dados podiam não ter todos o mesmo tamanho e por isso era essencial que os dados que se encontrassem a mais, num ou mais contadores, fossem eliminados para se poder aplicar o coeficiente de correlação de Spearman.

Foi no último ponto que foi encontrada maior dificuldade, pois sempre que algum valor estivesse em falta num contador era necessário eliminar esse valor do contador do Posto de Transformação e posteriormente eliminar dos restantes contadores, e voltar ao princípio, para nova comparação dos dados entre Posto de Transformação e contadores.

Mesmo com estas dificuldades o programa ficou concluído e a funcionar corretamente atribuindo com sucesso uma classificação para cada contador no conjunto de dados fornecido desde que esse contador estivesse presente em todos os outros ficheiros, por causa da exigência de que as amostras dos valores dos contadores devem ter o mesmo tamanho.

Com o desenvolvimento e implementação deste método a empresa passa a dispor de uma ferramenta automática para o mapeamento da rede.

Capítulo 4

Conclusão e Trabalho Futuro

O problema da identificação da fase nos contadores numa rede de distribuição de baixa tensão é ainda um problema pouco estudado, no entanto foi aqui apresentada uma possível abordagem para a resolução deste problema.

Com a aplicação de apenas um critério, obteve-se um sucesso de 82%. O critério do máximo do coeficiente de Correlação de Spearman para as séries das derivadas foi o que provou ser o mais adequado para o problema em questão, depois de analisados outros possíveis critérios, como o mínimo do desvio padrão das séries das diferenças ou das diferenças das derivadas. Apesar da quantidade de dados fornecida, foi necessária a criação de algumas séries que ajudassem a perceber melhor a relação entre o comportamento da tensão de um contador e da tensão das fases do Posto de Transformação.

Das variáveis fornecidas apenas as de tensão é que foram trabalhadas o que mostra que existe uma grande margem para melhoria do método criado, quer através da utilização das outras variáveis e possíveis combinações entre elas, quer pela vasta quantidade de métodos, de diversas áreas, que aqui podiam ser aplicados, como por exemplo, na área de data mining.

Trabalho Futuro

Com o estudo das outras variáveis fornecidas e da aplicação de métodos de outras áreas com a área de data mining poderia ainda ser feita uma tentativa de melhoria da solução encontrada.

Neste trabalho não foram tratados Postos de Transformação que apresentassem micro-geração, que pela introdução de energia na rede alteram os diagramas de carga de modo que estes já não vão estar em constante decaimento, como os que foram trabalhados nesta tese, mas vão estar em constante modificação. Assim este trabalho poderia agora ser estendido para os contadores considerados mais problemáticos da rede de distribuição, para os quais certamente necessitarão de uma outra análise complementar à que

aqui foi referida.

Depois de identificados os contadores por fase, e com as referências geográficas dos contadores, poderia ser feito o mapeamento desses contadores, ou seja encontrar a sua localização exata na rede, o que permitiria a resolução de muitos problemas como a identificação de fraudes, a otimização da rede distribuição, o seu equilíbrio energético, entre outros.

Bibliografia

- Box-Steffensmeier, J. M., Freeman, J. R., Hitt, M. P., and Pevehouse, J. C. W. (2014). *Time Series Analysis for the Social Sciences*.
- Burke, S. and Hunter, J. (2005). *Modelling Non-Stationary Economic Time Series: A Multivariate Approach*.
- Cryer, J. D. and Chan, K.-S. (2008). *Time Series Analysis: With Applications in R*.
- Ekstrom, C. T. and Sørensen, H. (2014). *Introduction to Statistical Data Analysis for the Life Sciences*.
- Gastwirth, J. L., Gel, Y. R., and W.Miao (2009). *The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice*.
- Harvey, A. C. (2001). *Forecasting, Strutural Time Series Models and the Kalman Filter*.
- Hipel, K. W. and McLeod, A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*.
- Kočenda, Evžen, Černý, and Alexandr (2015). *Elements of Time Series Econometrics: an Applied Approach*.
- Milliken, G. A. and Johnson, D. E. (2009). *Analysis of Messy Data Volume 1: Designed Experiments*.
- Mimmac, G. M., Meyer, D. H., and Manas, G. J. (2001). *Introductory Statistics for Business*.
- Morgan, G. A., Leech, N. L., Gloeckner, G. W., and Barrett, K. C. (2011). *IBM SPSS for Introductory Statistics: Use and Interpretation*.
- Pankratz, A. (2009). *Forecasting with Univariate Box - Jenkins Models: Concepts and Cases*.
- Ramos, E., Pampliega, D., and Deschamps, P. (2015). *From Data Collecting to Business Intelligence and Data Mining*.
- Ritz, C. and Streibig, J. C. (2008). *Nonlinear Regression with R*.

Schlotzhauer, S. D. (2007). *Elementary Statistics Using JMP*.

Sharma, J. K. (2012). *Business Statistics*.

Wang, W. (2006). *Stochasticity, Nonlinearity and Forecasting of Streamflow*.

Ware, W. B., Ferron, J. M., and Miller, B. M. (2013). *Introductory Statistics: A conceptual Approach using R*.

Weatherhead, S., Higgins, C., and Collinson, A. (2015). *Characterisation of Phase Current Imbalance on Three-Phase LV Feeders to Identify Opportunities for Rebalancing*.

Yaffee, R. A. and McGee, M. (2000). *An Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS*.

Anexos

Anexo A

Neste anexo são apresentadas as figuras obtidas para os restantes contadores referentes à secção 3.2.

Gráficos conjuntos dos diagramas de tensão dos contadores e das três fases do PT.

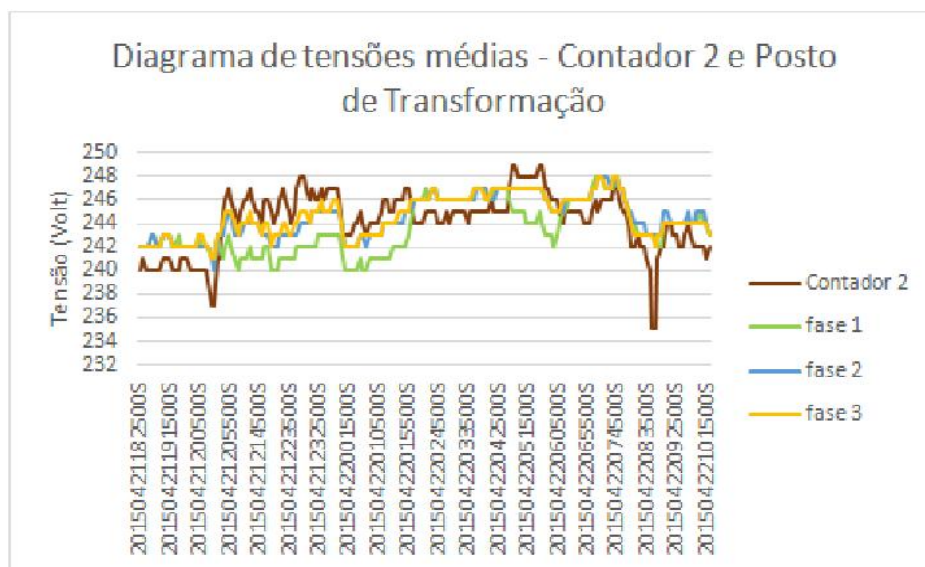


Figura 4.1: Dados da tensão média do contador 2 e das três fases do contador no Posto de Transformação

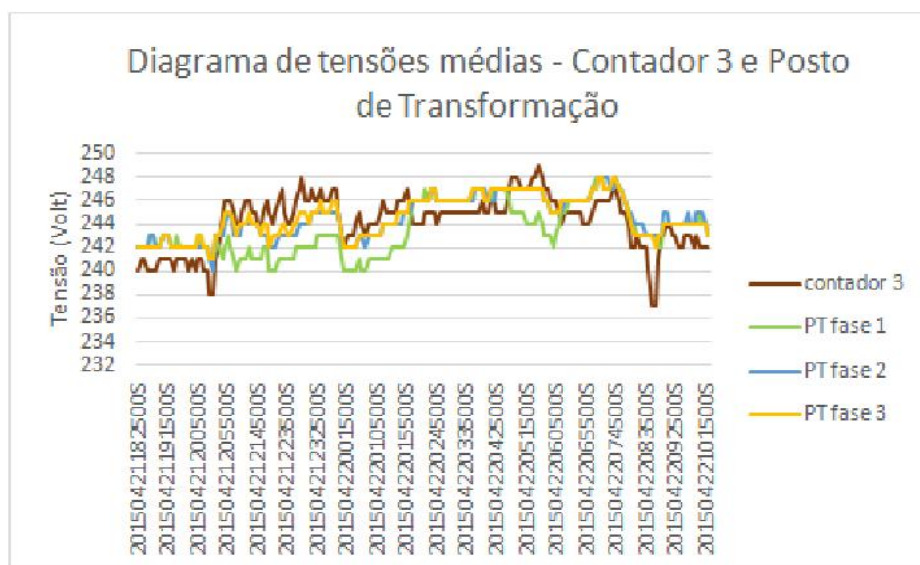


Figura 4.2: Dados da tensão média do contador 3 e das três fases do contador no Posto de Transformação

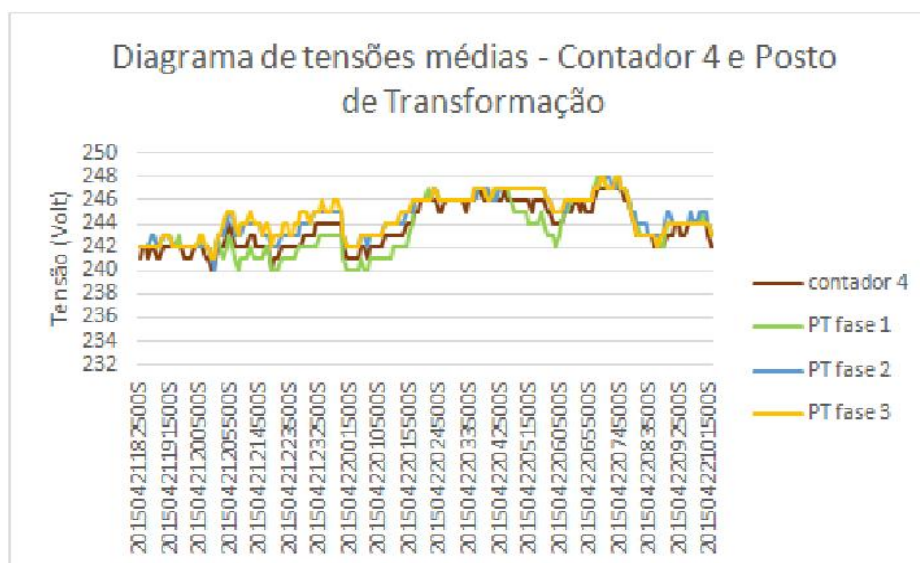


Figura 4.3: Dados da tensão média do contador 4 e das três fases do contador no Posto de Transformação

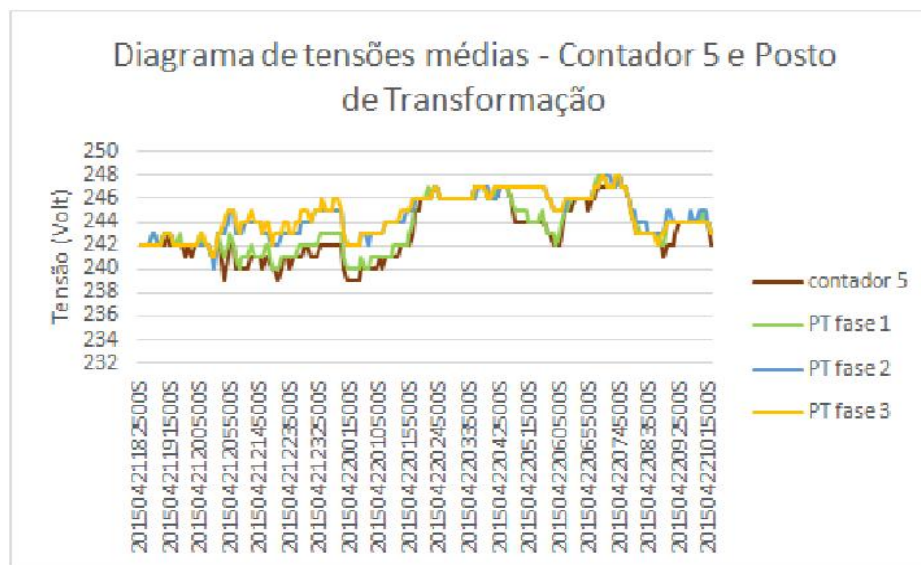


Figura 4.4: Dados da tensão média do contador 5 e das três fases do contador no Posto de Transformação

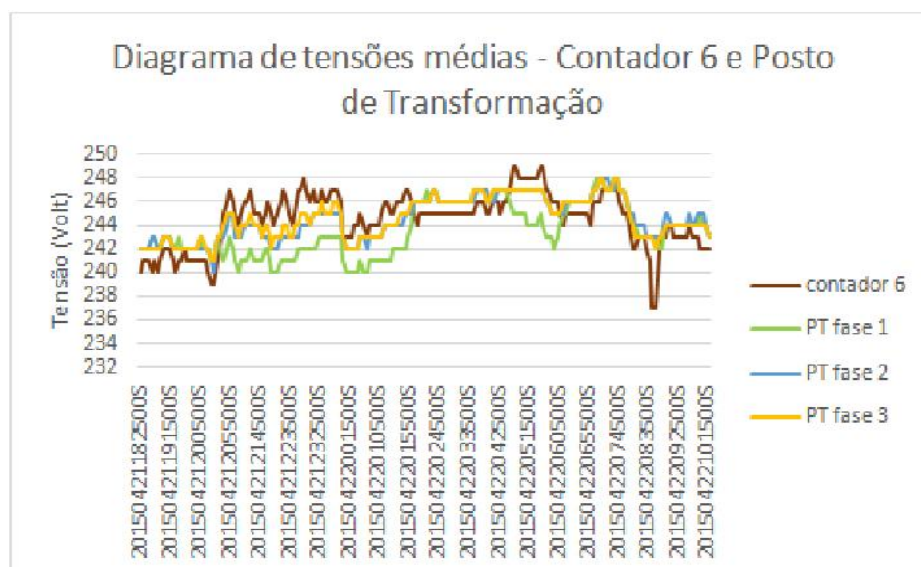


Figura 4.5: Dados da tensão média do contador 6 e das três fases do contador no Posto de Transformação

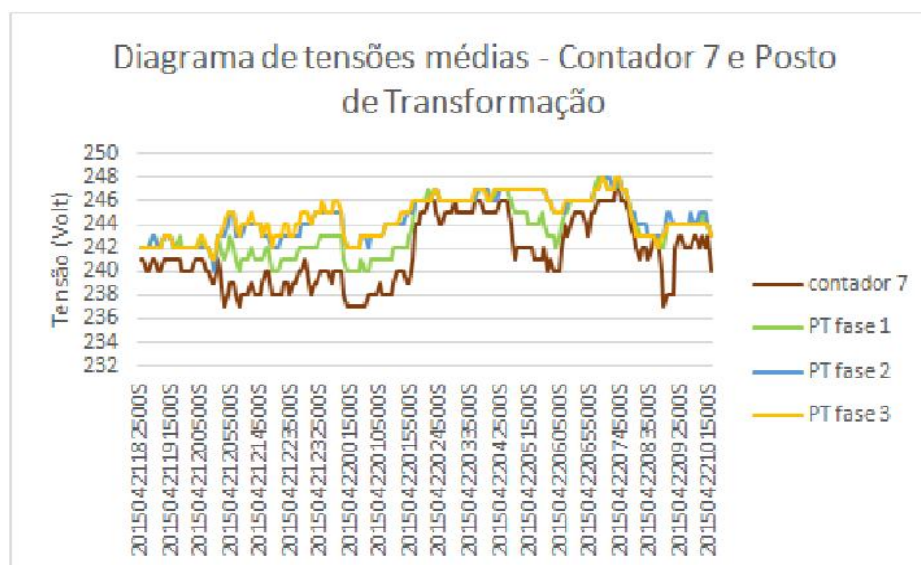


Figura 4.6: Dados da tensão média do contador 7 e das três fases do contador no Posto de Transformação

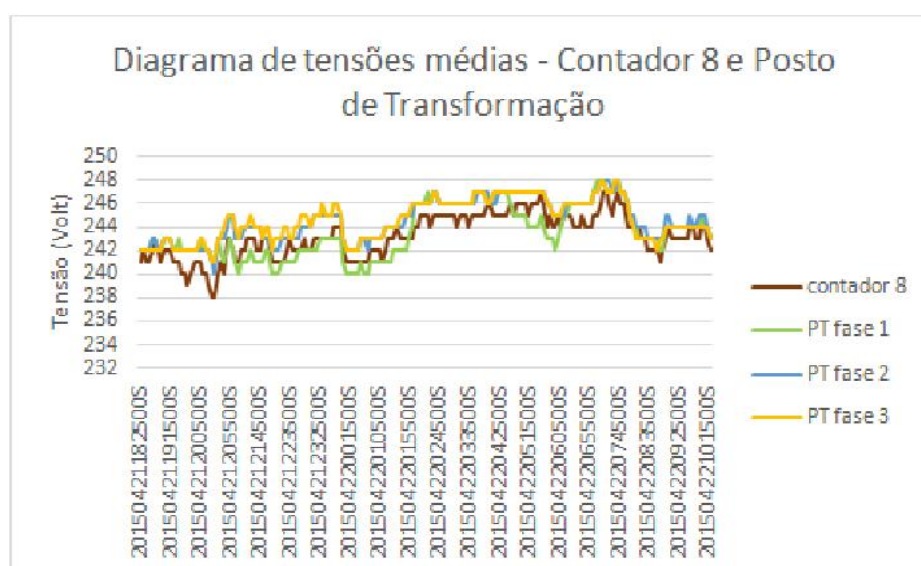


Figura 4.7: Dados da tensão média do contador 8 e das três fases do contador no Posto de Transformação

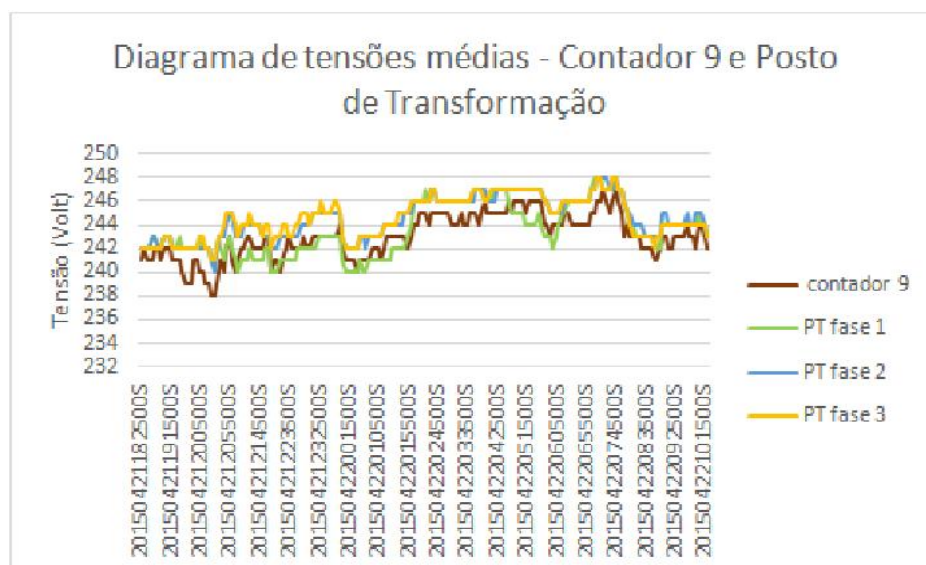


Figura 4.8: Dados da tensão média do contador 9 e das três fases do contador no Posto de Transformação

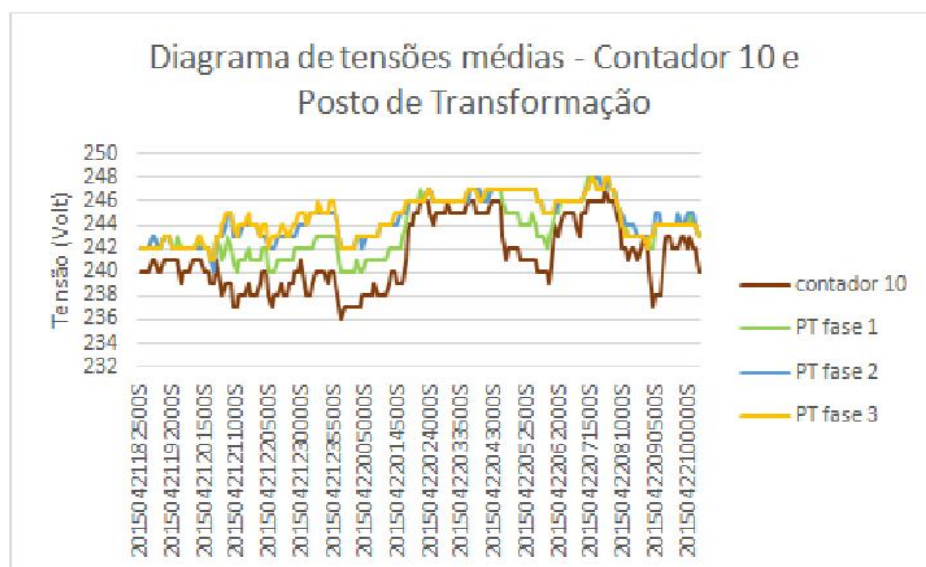


Figura 4.9: Dados da tensão média do contador 10 e das três fases do contador no Posto de Transformação

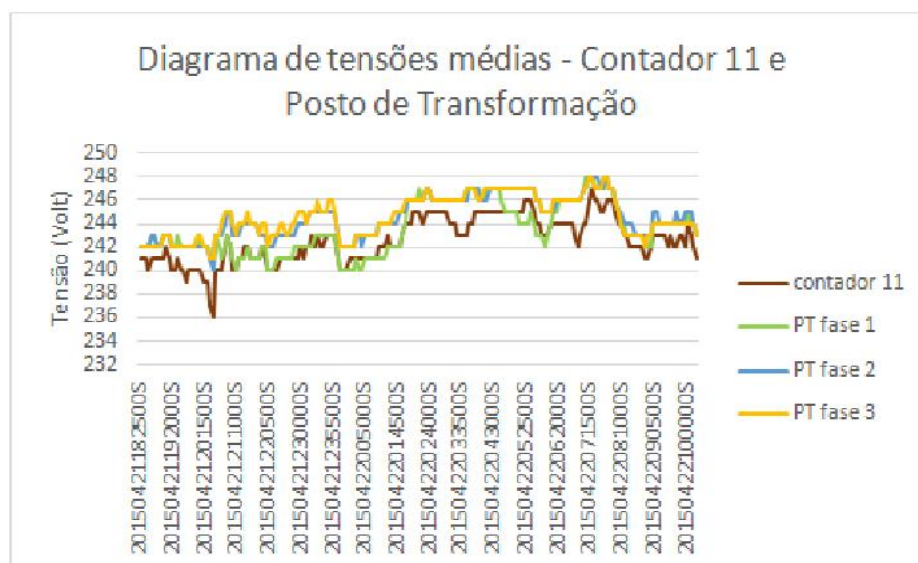


Figura 4.10: Dados da tensão média do contador 11 e das três fases do contador no Posto de Transformação

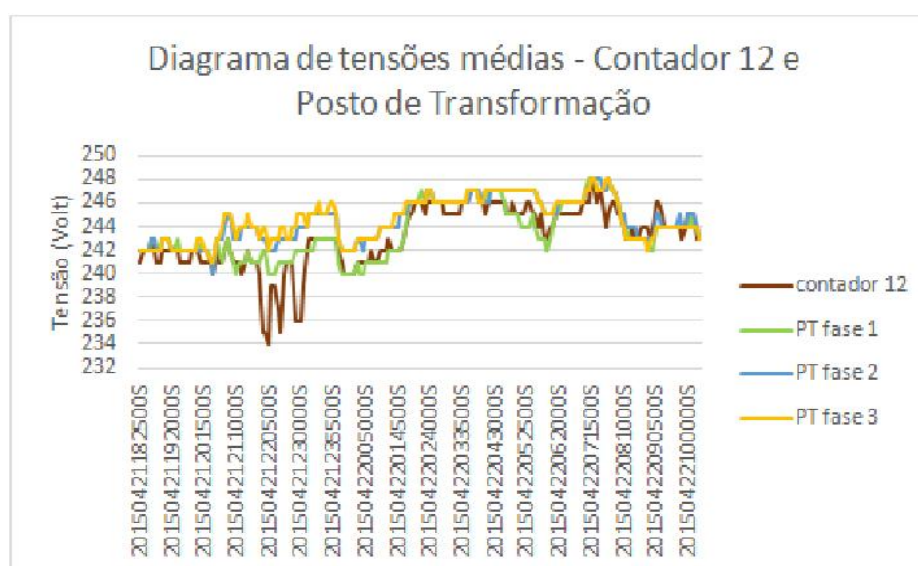


Figura 4.11: Dados da tensão média do contador 12 e das três fases do contador no Posto de Transformação

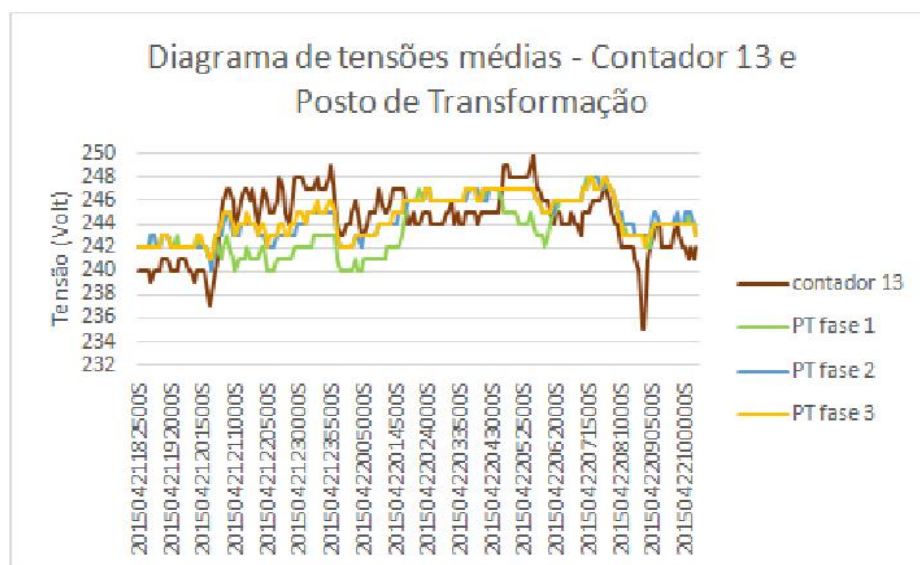


Figura 4.12: Dados da tensão média do contador 13 e das três fases do contador no Posto de Transformação

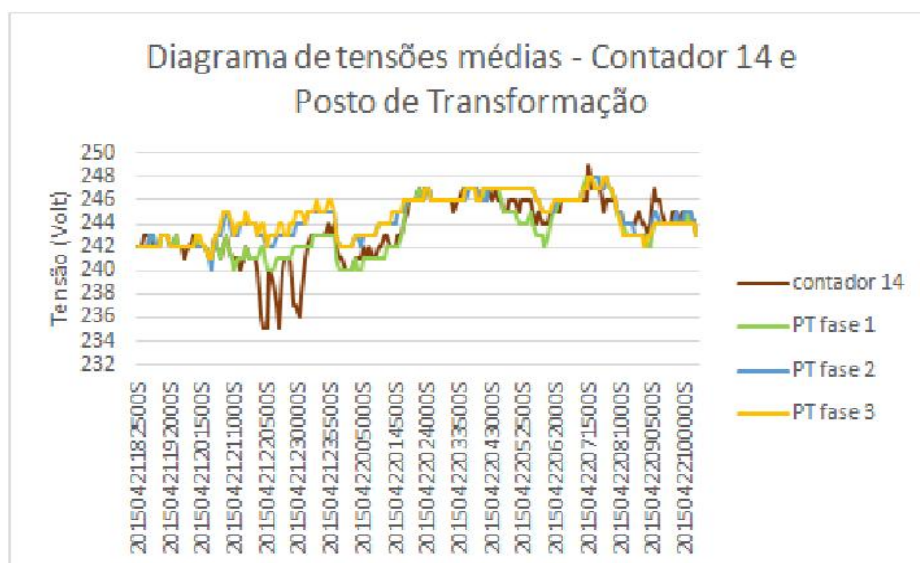


Figura 4.13: Dados da tensão média do contador 14 e das três fases do contador no Posto de Transformação

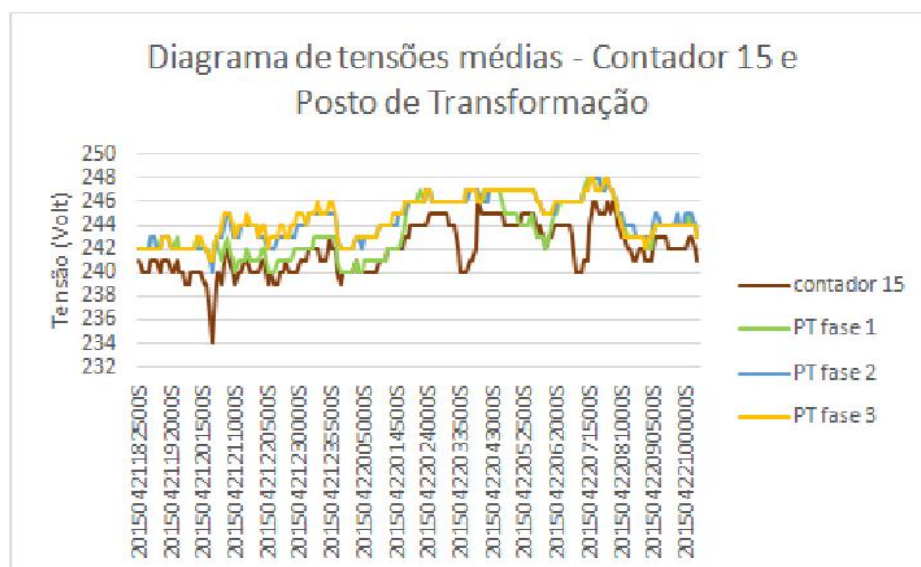


Figura 4.14: Dados da tensão média do contador 15 e das três fases do contador no Posto de Transformação

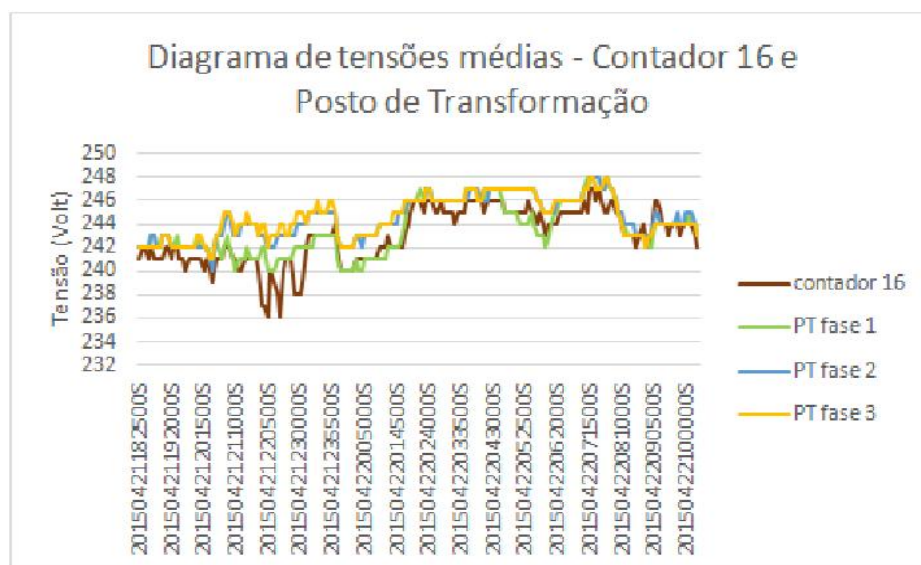


Figura 4.15: Dados da tensão média do contador 16 e das três fases do contador no Posto de Transformação

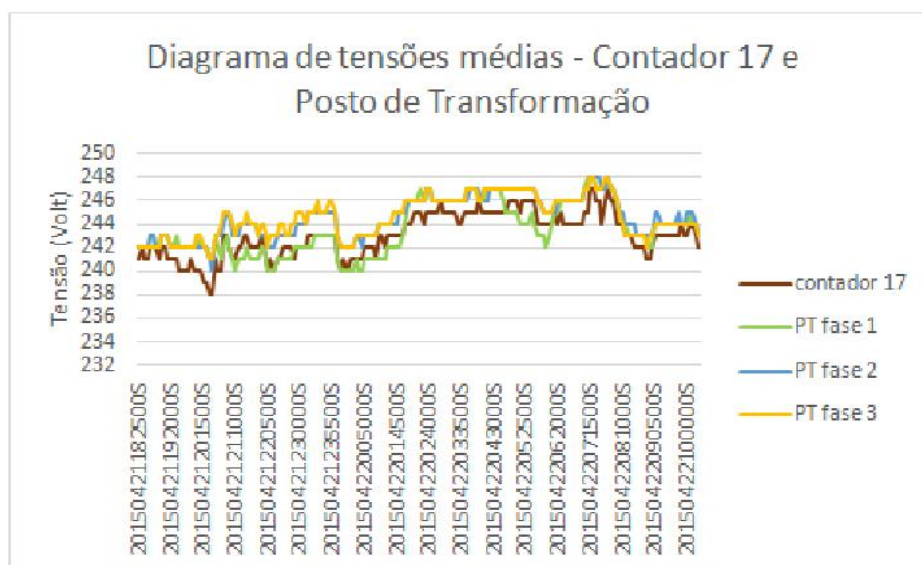


Figura 4.16: Dados da tensão média do contador 17 e das três fases do contador no Posto de Transformação

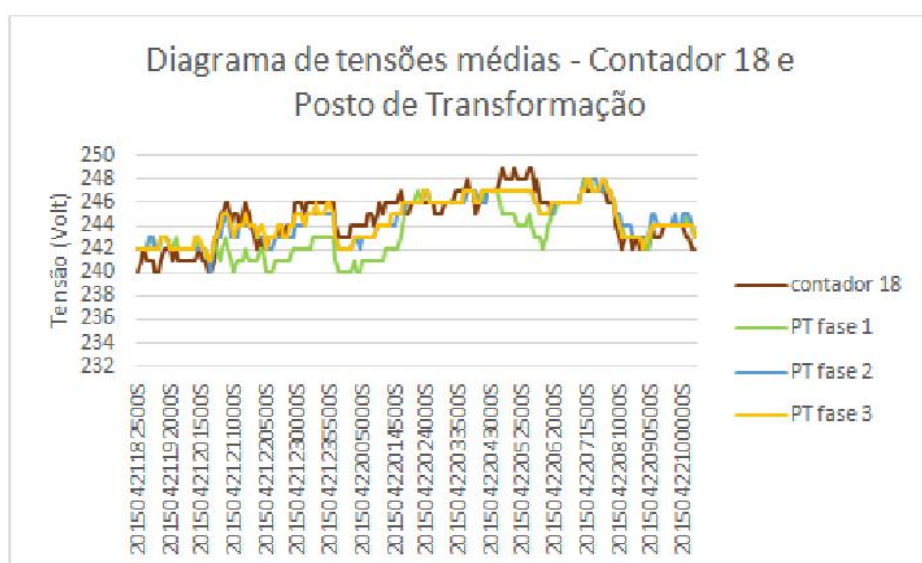


Figura 4.17: Dados da tensão média do contador 18 e das três fases do contador no Posto de Transformação

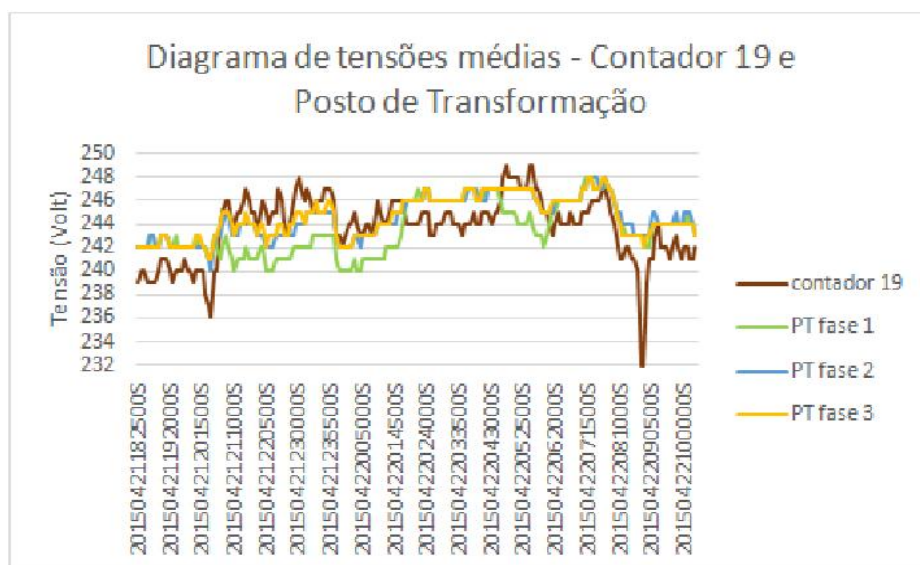


Figura 4.18: Dados da tensão média do contador 19 e das três fases do contador no Posto de Transformação

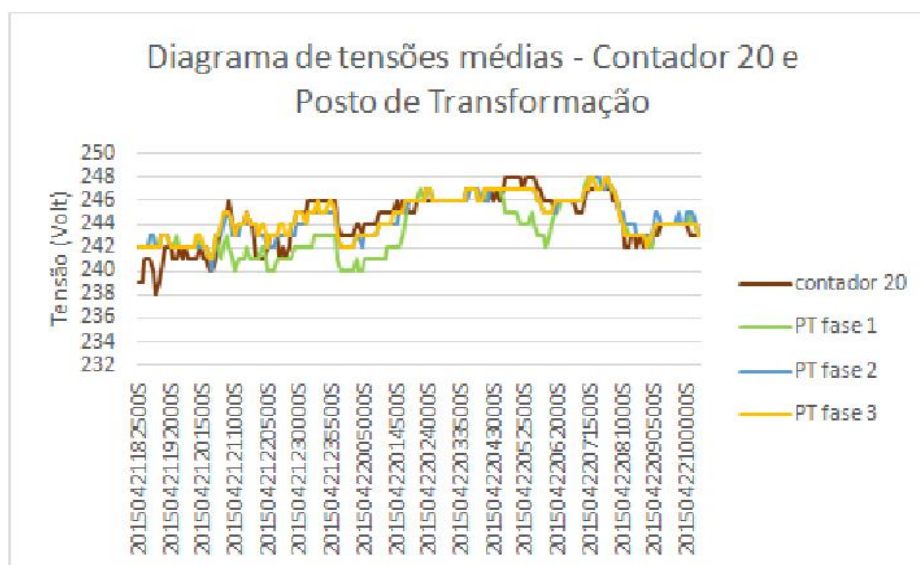


Figura 4.19: Dados da tensão média do contador 20 e das três fases do contador no Posto de Transformação

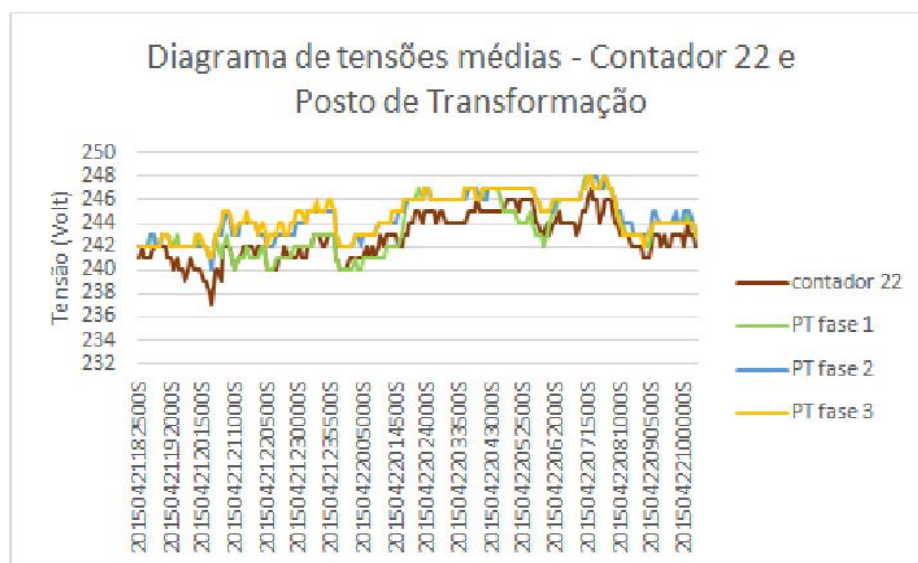


Figura 4.20: Dados da tensão média do contador 22 e das três fases do contador no Posto de Transformação

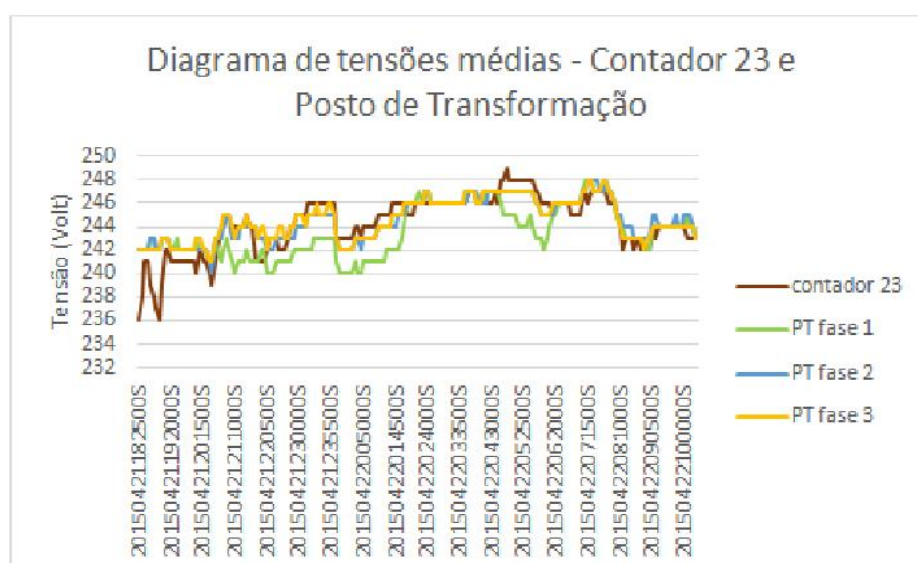


Figura 4.21: Dados da tensão média do contador 23 e das três fases do contador no Posto de Transformação

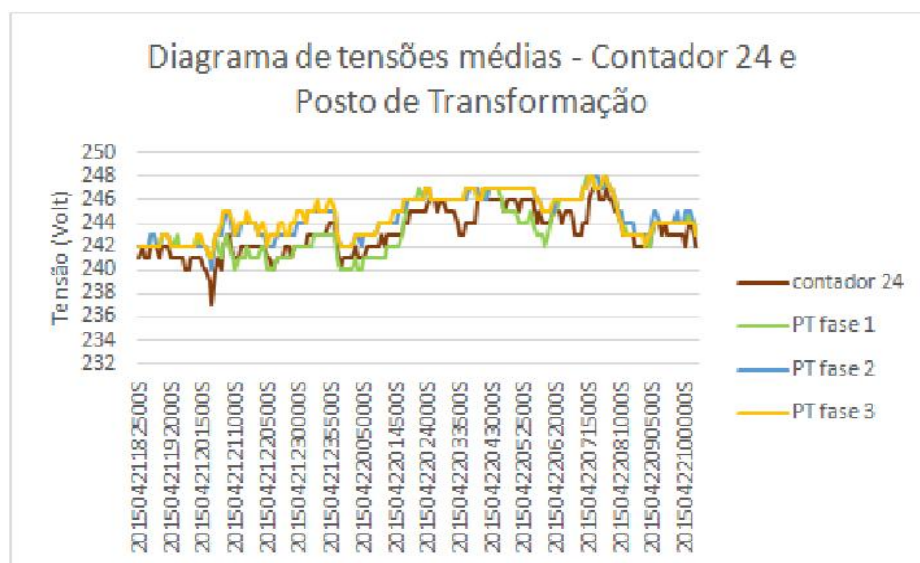


Figura 4.22: Dados da tensão média do contador 24 e das três fases do contador no Posto de Transformação

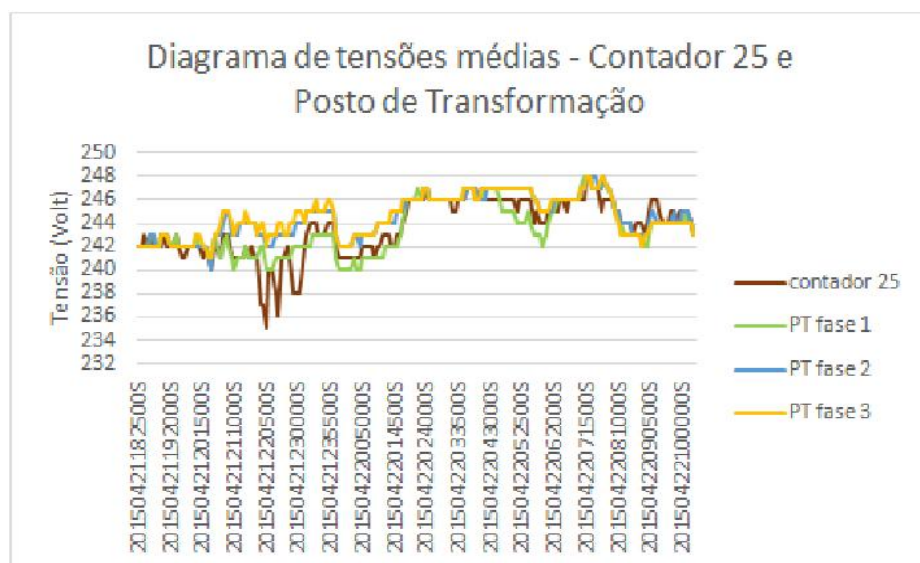


Figura 4.23: Dados da tensão média do contador 25 e das três fases do contador no Posto de Transformação

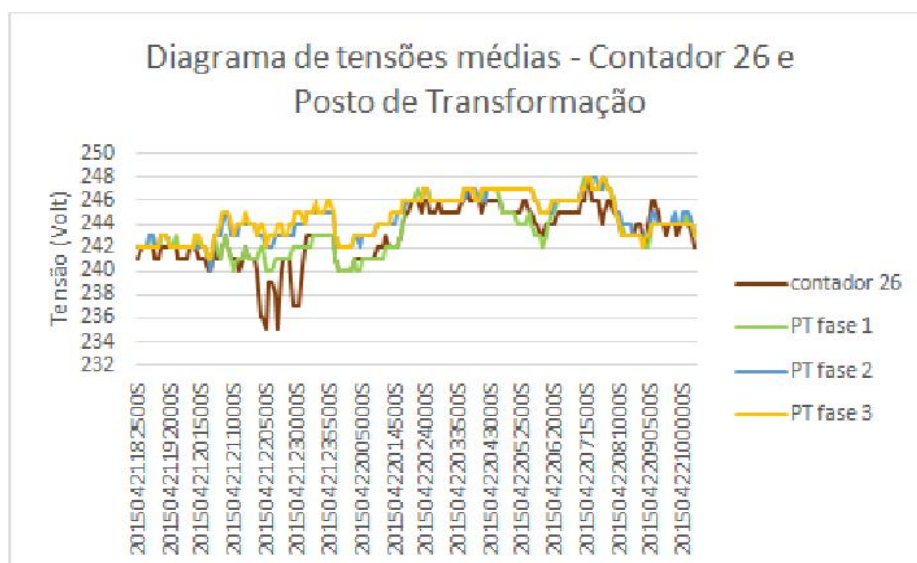


Figura 4.24: Dados da tensão média do contador 26 e das três fases do contador no Posto de Transformação

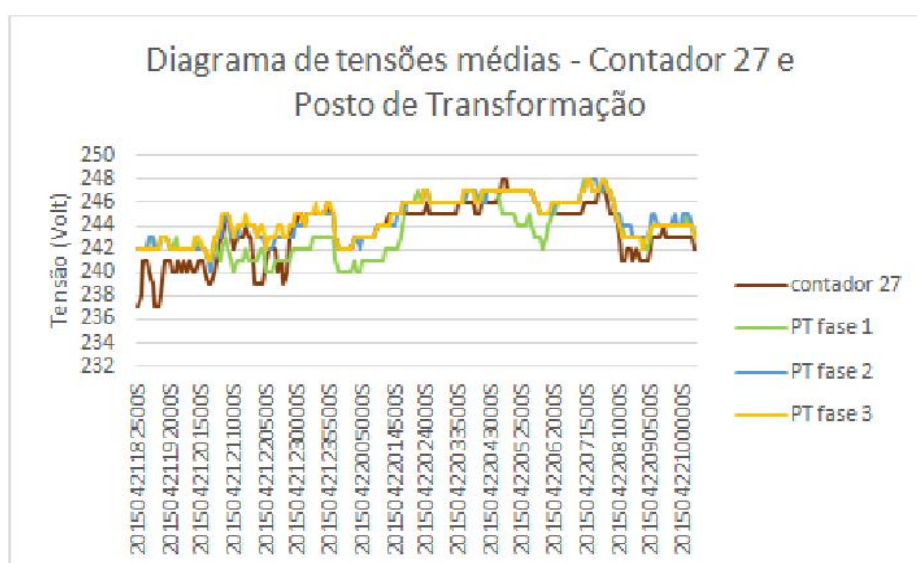


Figura 4.25: Dados da tensão média do contador 27 e das três fases do contador no Posto de Transformação

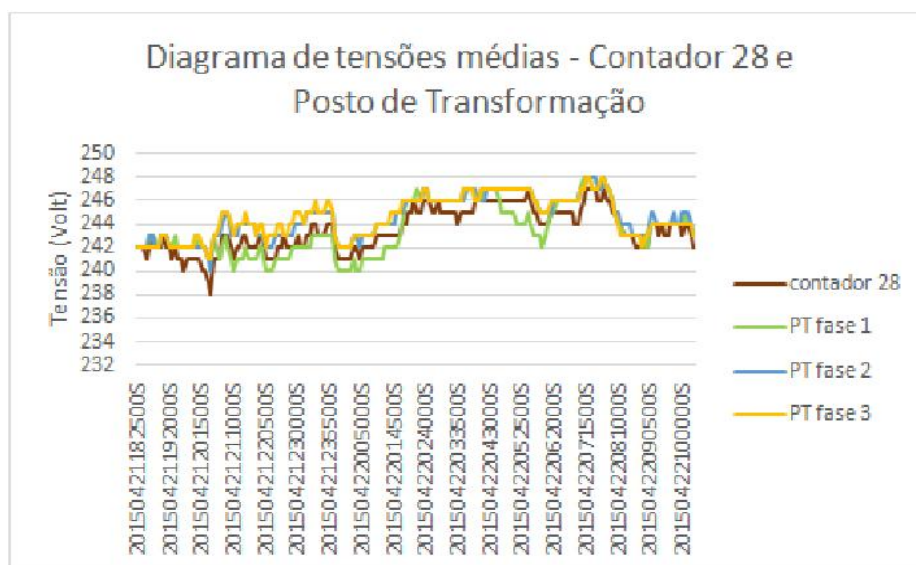


Figura 4.26: Dados da tensão média do contador 28 e das três fases do contador no Posto de Transformação

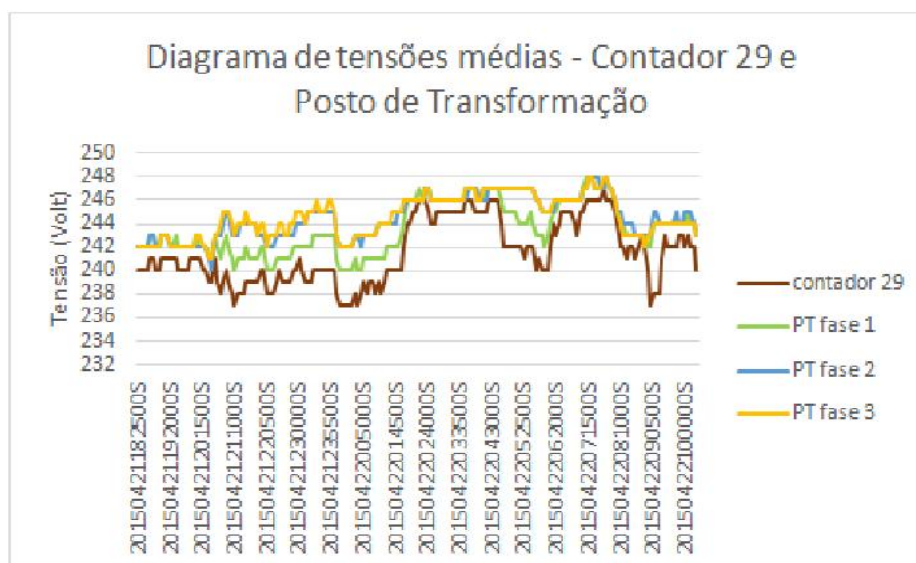


Figura 4.27: Dados da tensão média do contador 29 e das três fases do contador no Posto de Transformação

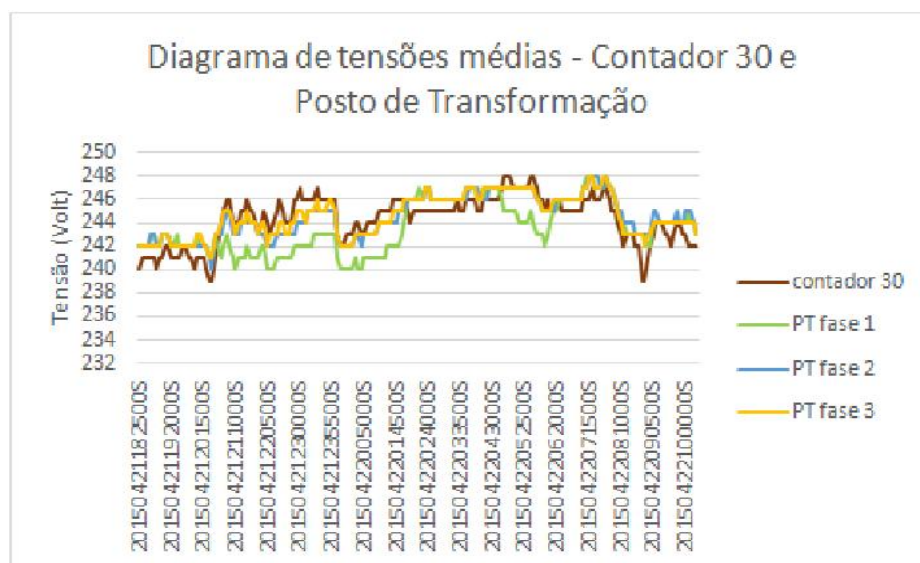


Figura 4.28: Dados da tensão média do contador 30 e das três fases do contador no Posto de Transformação

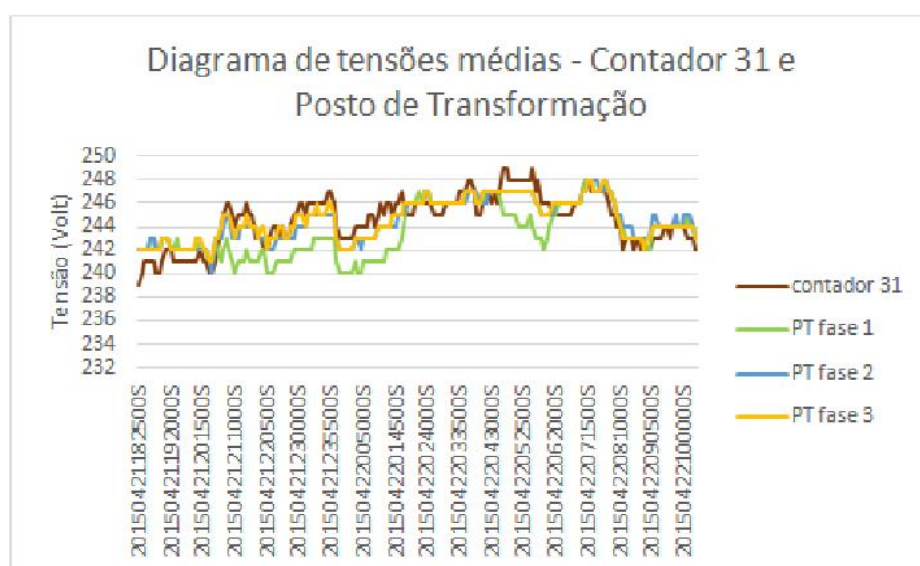


Figura 4.29: Dados da tensão média do contador 31 e das três fases do contador no Posto de Transformação

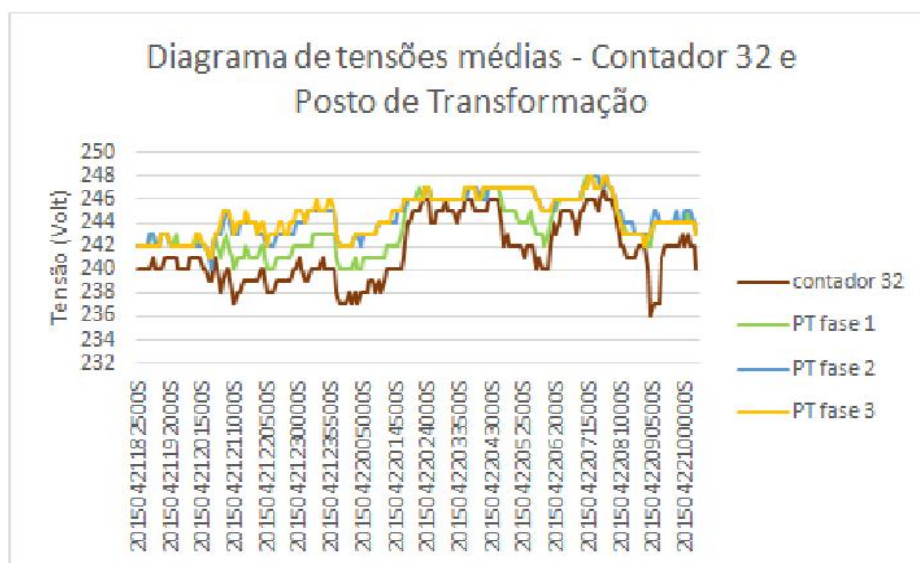


Figura 4.30: Dados da tensão média do contador 32 e das três fases do contador no Posto de Transformação

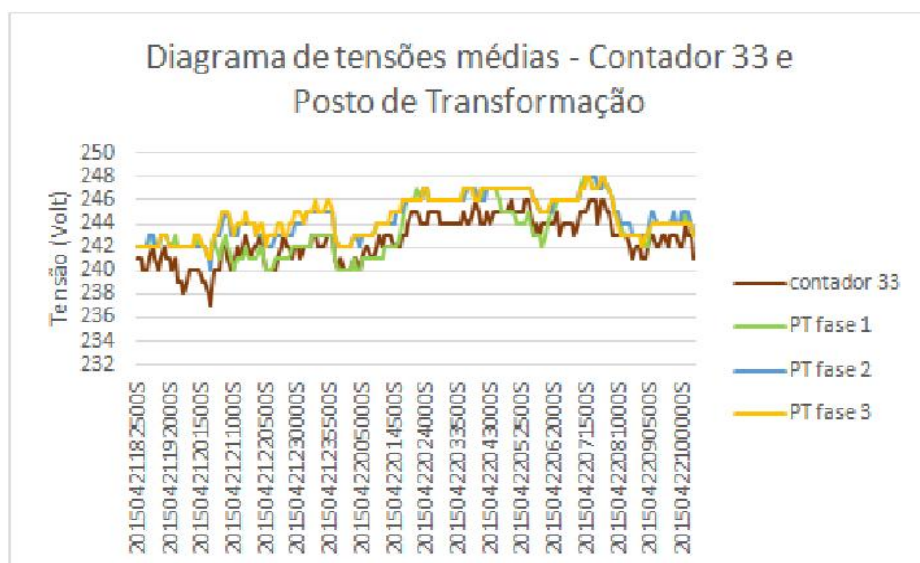


Figura 4.31: Dados da tensão média do contador 33 e das três fases do contador no Posto de Transformação

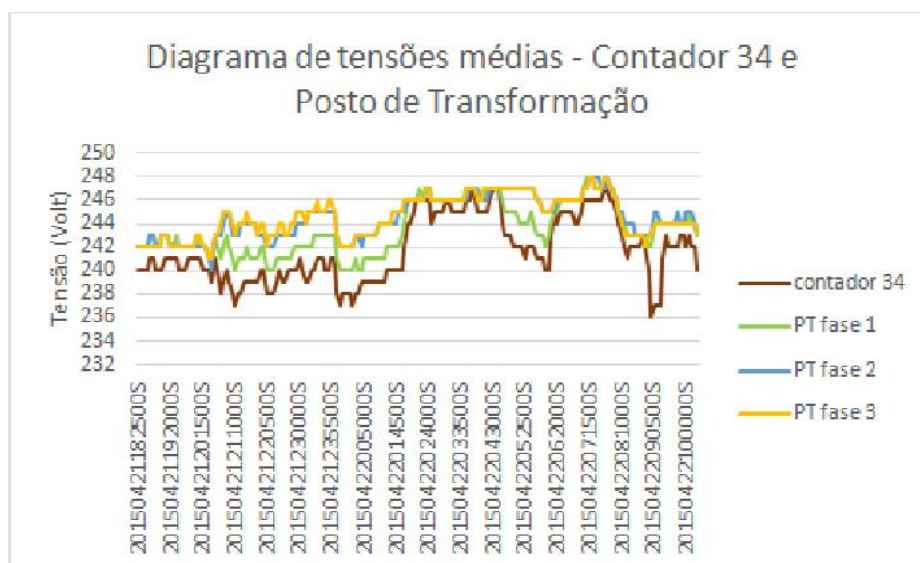


Figura 4.32: Dados da tensão média do contador 34 e das três fases do contador no Posto de Transformação

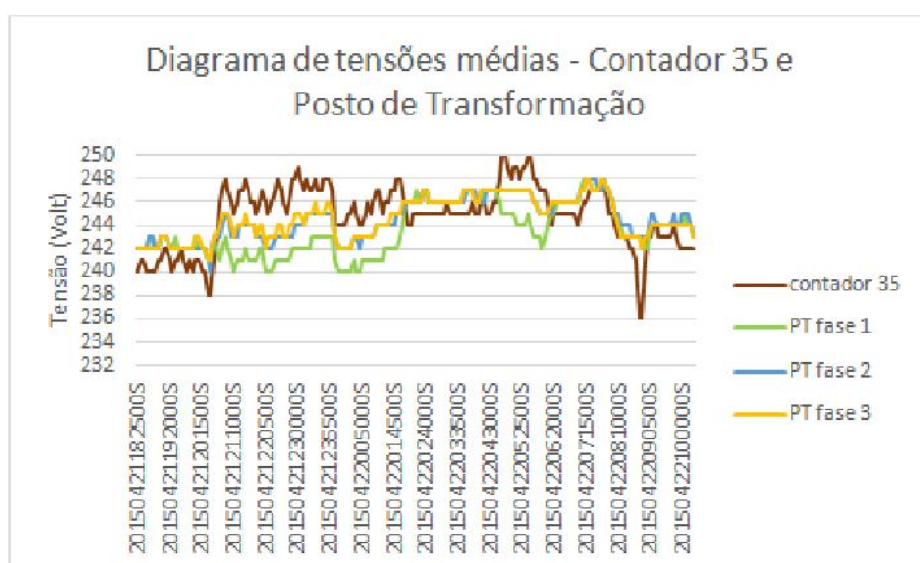


Figura 4.33: Dados da tensão média do contador 35 e das três fases do contador no Posto de Transformação

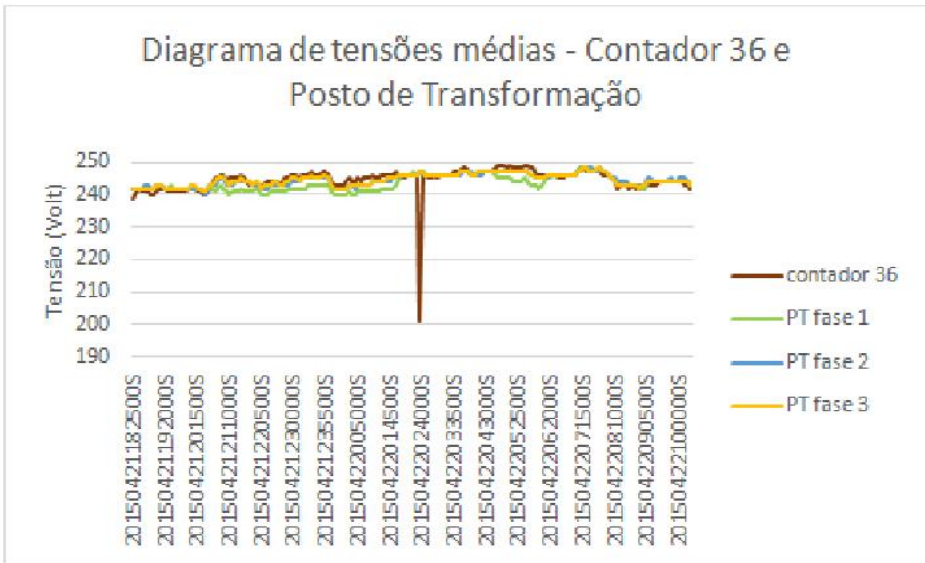


Figura 4.34: Dados da tensão média do contador 36 e das três fases do contador no Posto de Transformação

Anexo B

Neste anexo serão representadas as figuras referentes à secção 3.3.

Dos 36 contadores trabalhados foram escolhidos apenas 10 contadores para serem representados neste anexo que fornecem uma ideia representativa das séries obtidas para todos os contadores.

As figuras seguintes apresentam as séries das diferenças entre fase e contador.

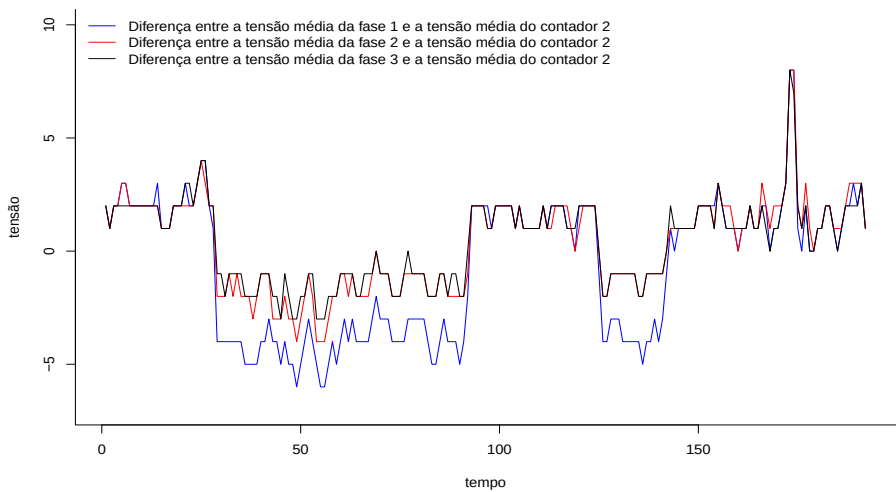


Figura 4.35: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador

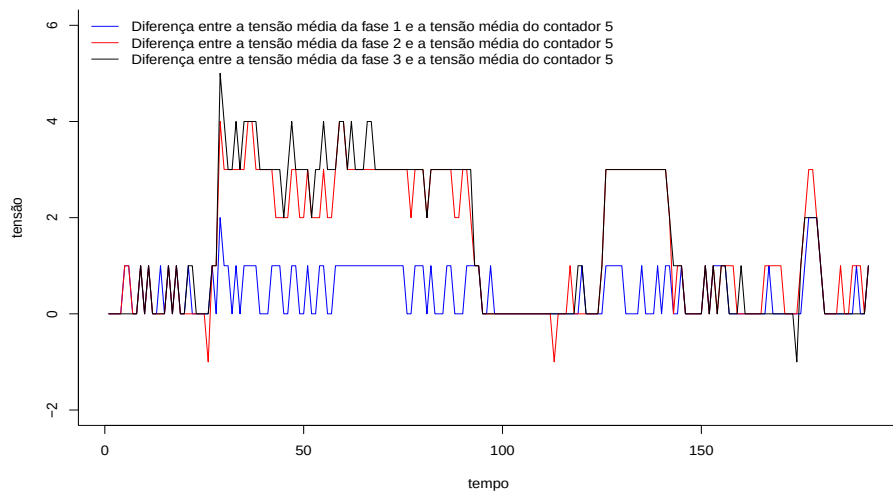


Figura 4.36: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 5

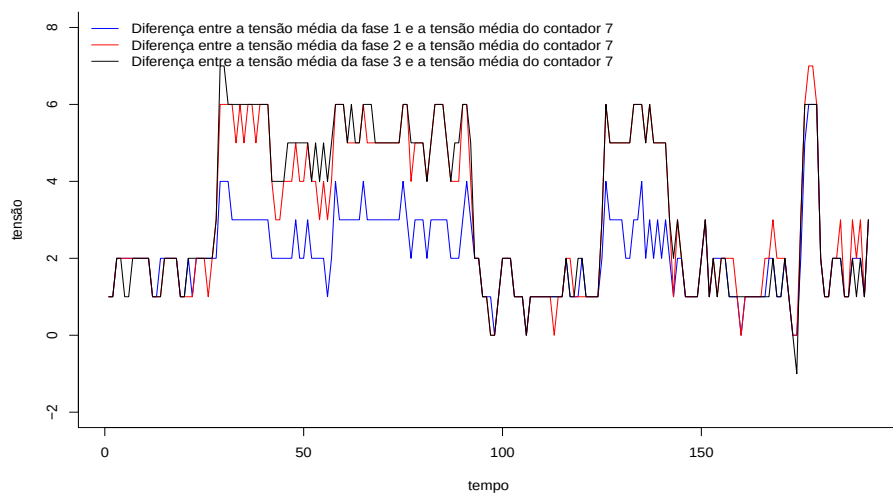


Figura 4.37: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 7

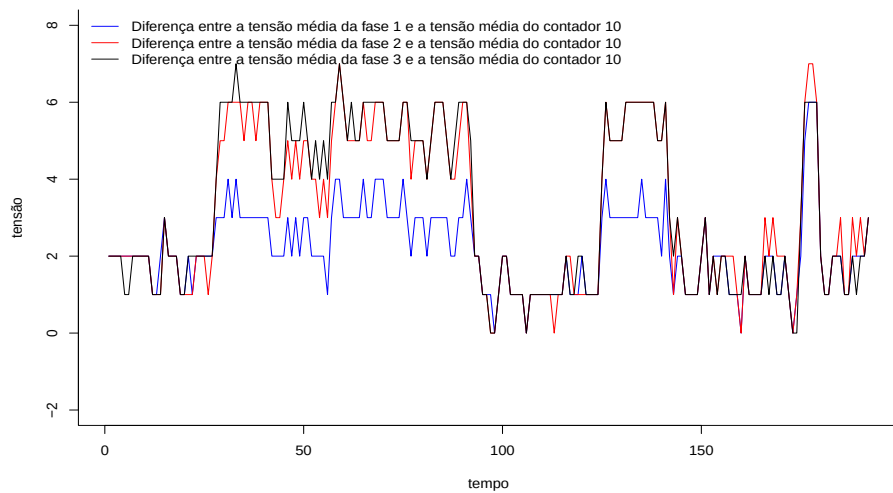


Figura 4.38: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 10

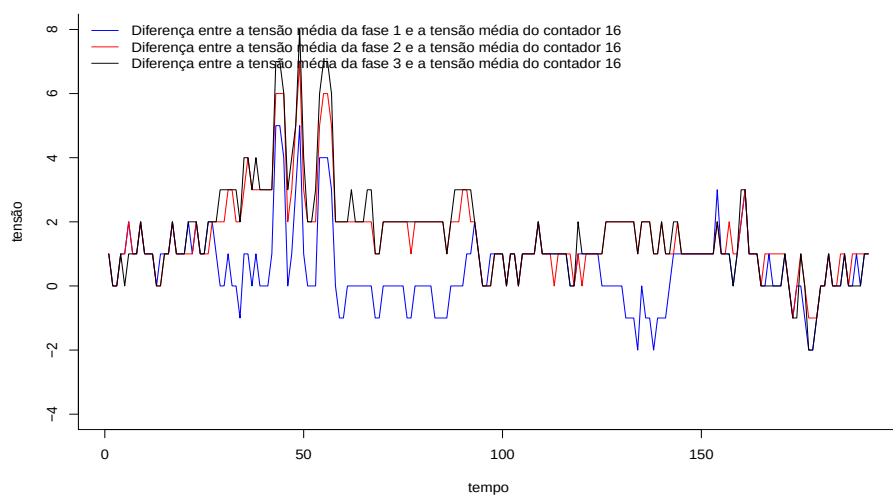


Figura 4.39: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 16

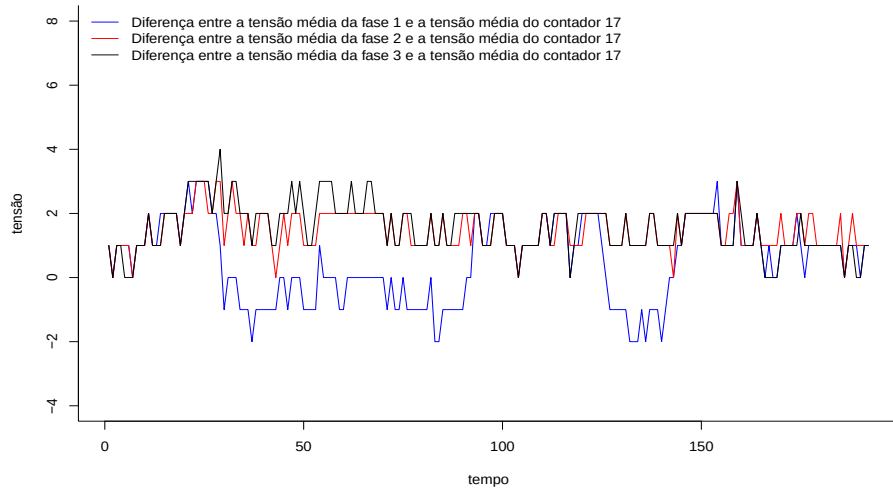


Figura 4.40: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 17

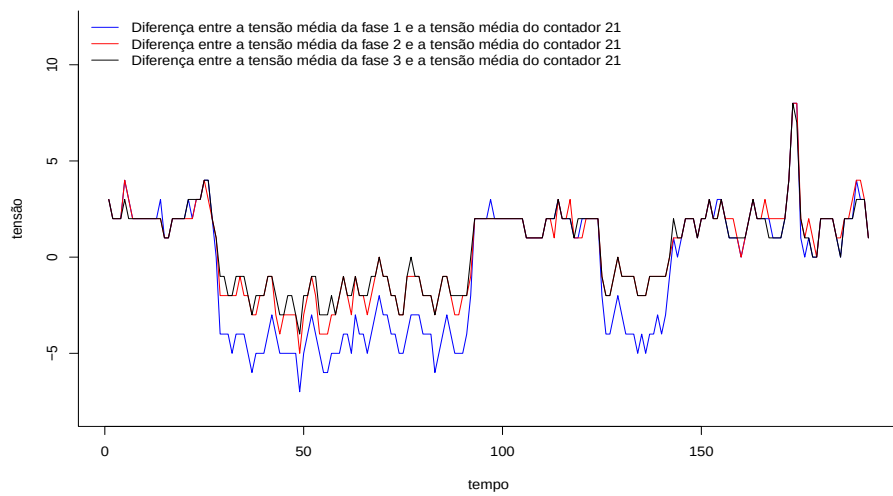


Figura 4.41: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 21

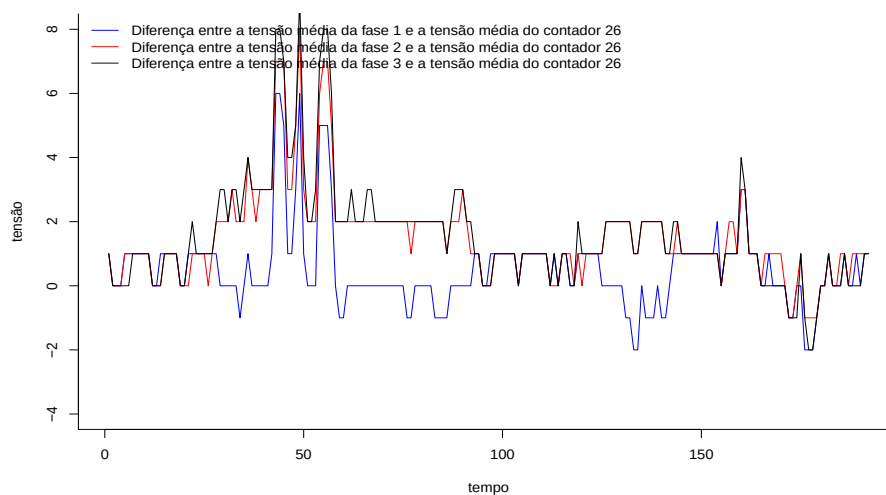


Figura 4.42: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 26

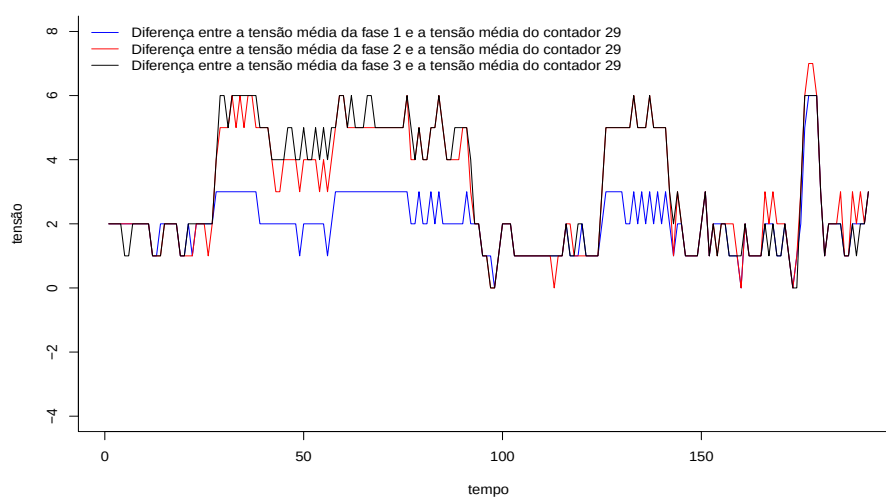


Figura 4.43: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 29

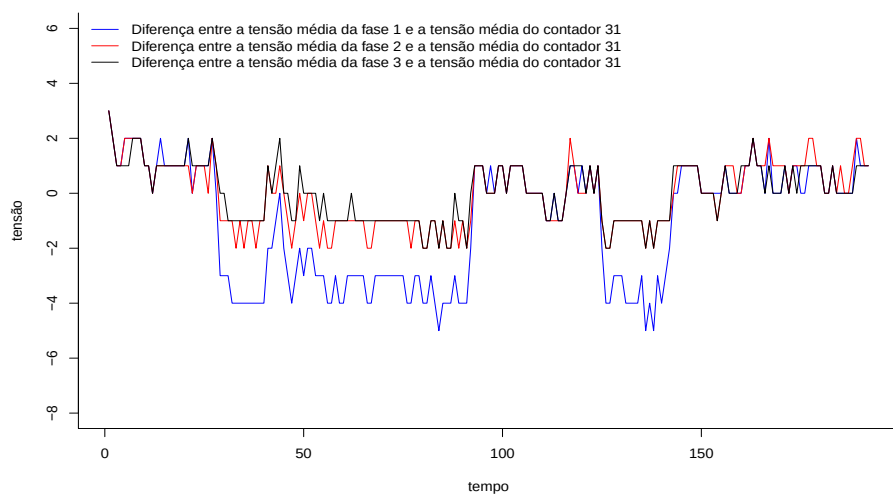


Figura 4.44: Série que contém a diferença entre a tensão média das três fases no Posto de Transformação e a tensão média do contador 31

As seguintes figuras são referentes à série das derivadas geradas para os 10 contadores.

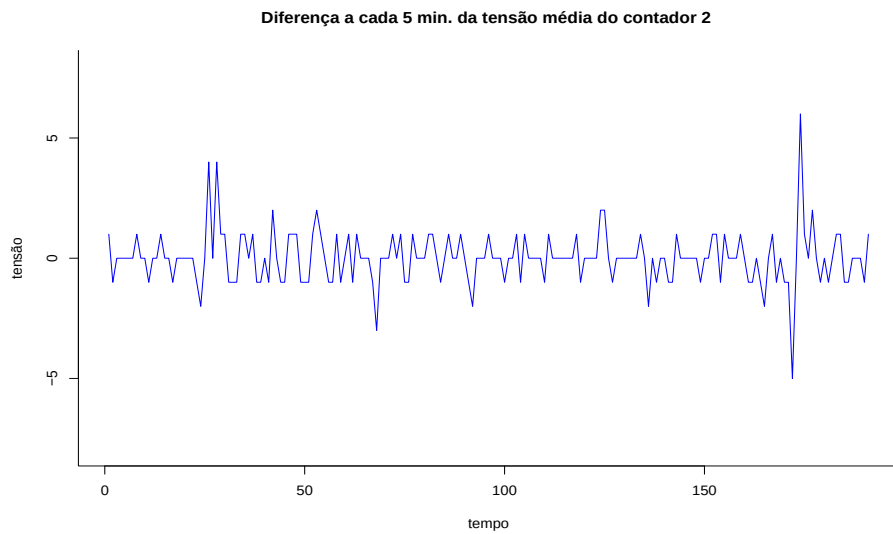


Figura 4.45: Série que contém as derivadas da tensão média do contador 2

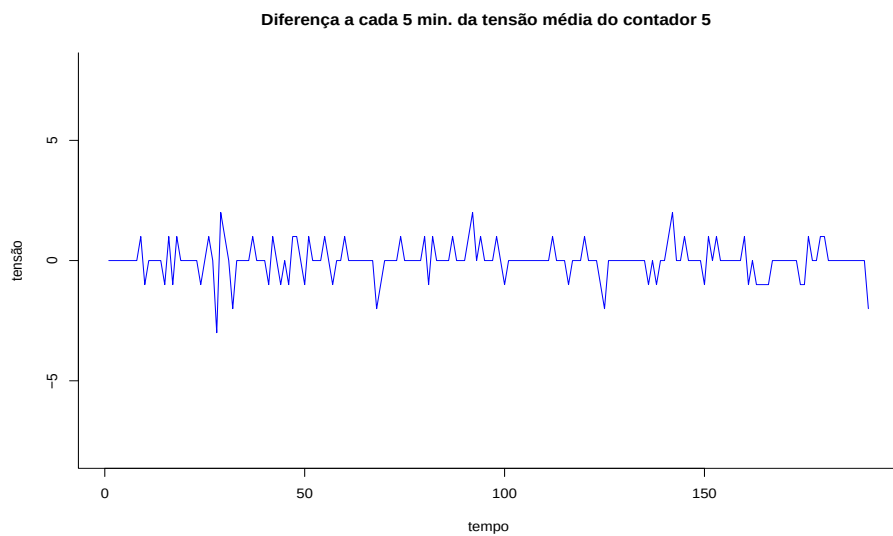


Figura 4.46: Série que contém as derivadas da tensão média do contador 5

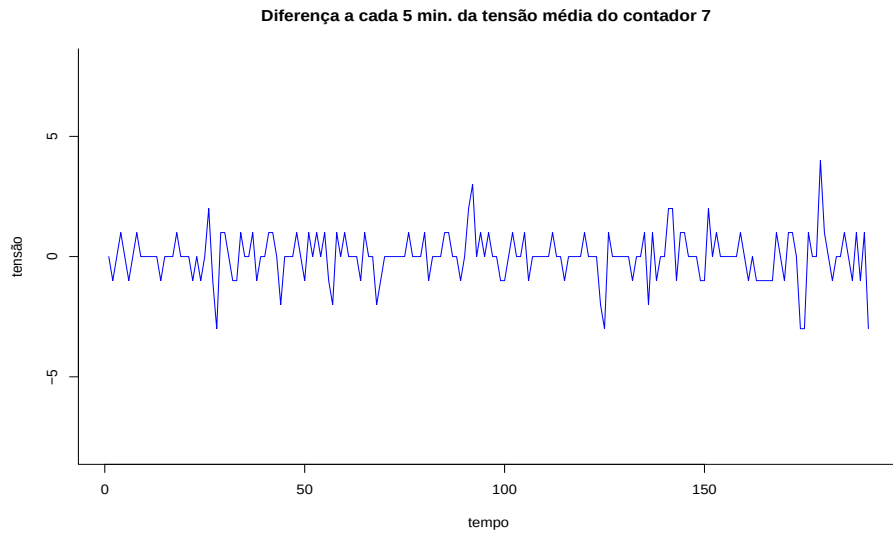


Figura 4.47: Série que contém as derivadas da tensão média do contador 7

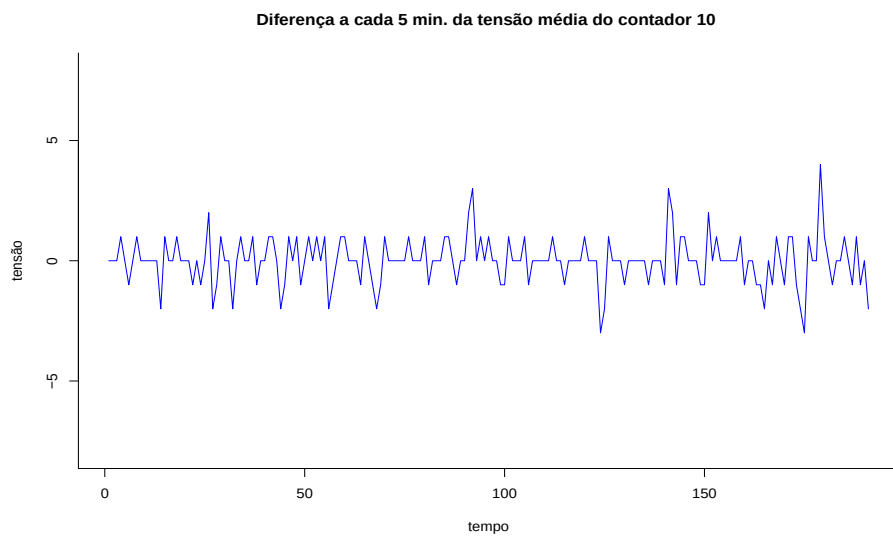


Figura 4.48: Série que contém as derivadas da tensão média do contador 10

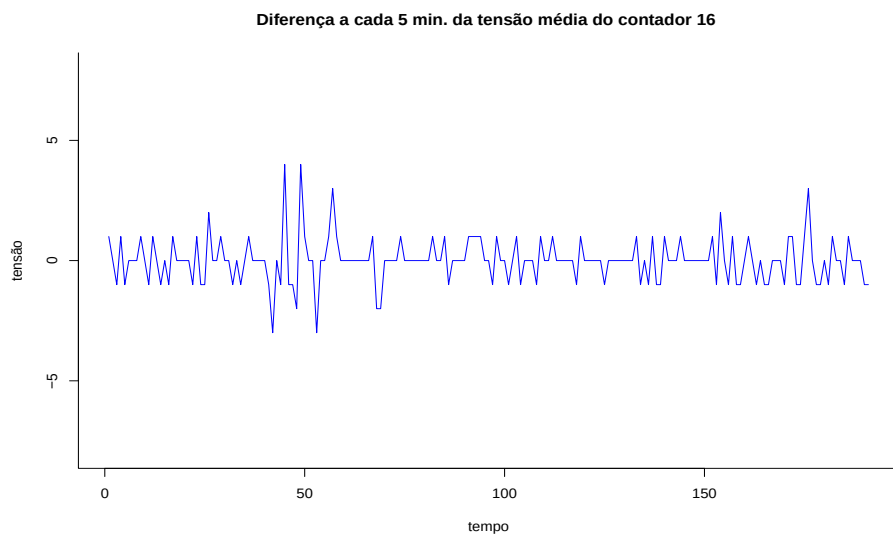


Figura 4.49: Série que contém as derivadas da tensão média do contador 16

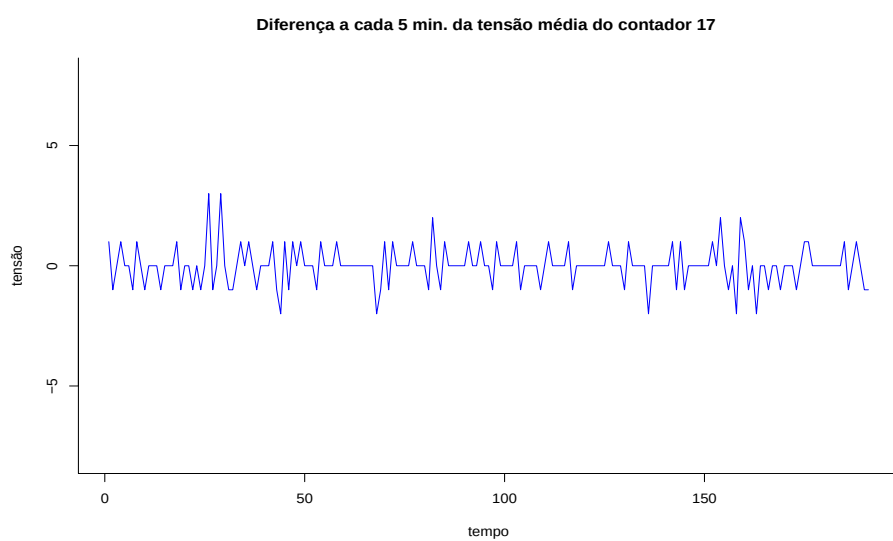


Figura 4.50: Série que contém as derivadas da tensão média do contador 17

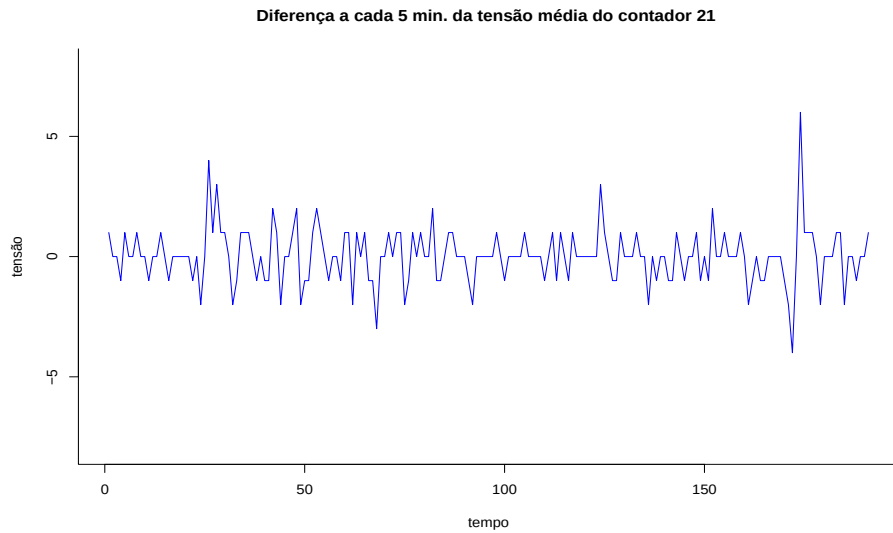


Figura 4.51: Série que contém as derivadas da tensão média do contador 21

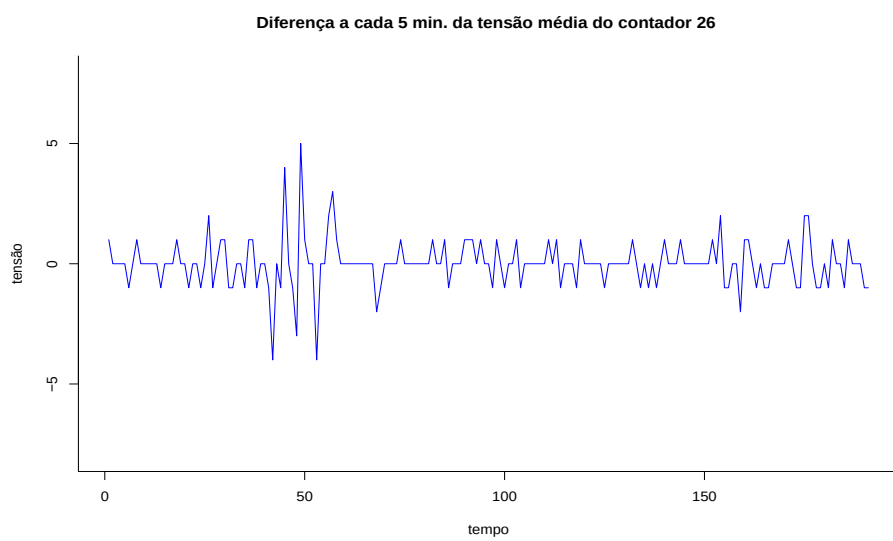


Figura 4.52: Série que contém as derivadas da tensão média do contador 26

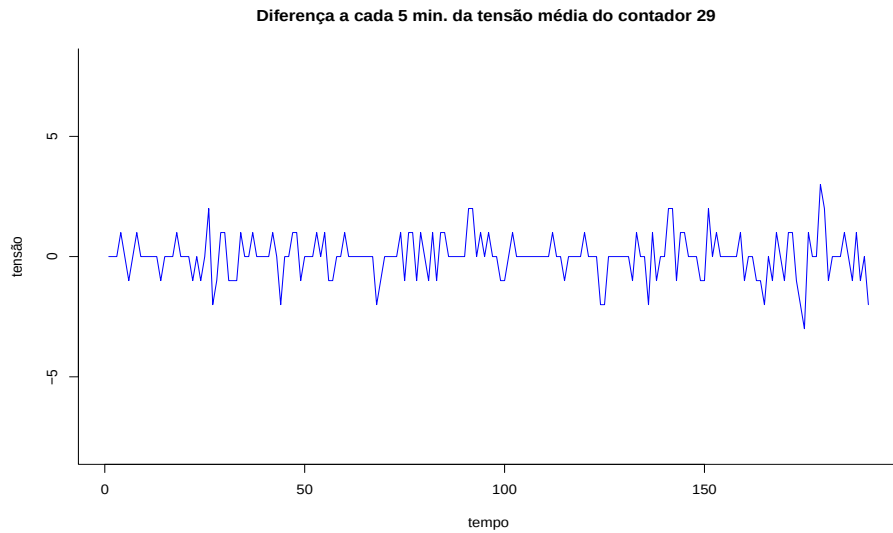


Figura 4.53: Série que contém as derivadas da tensão média do contador 29

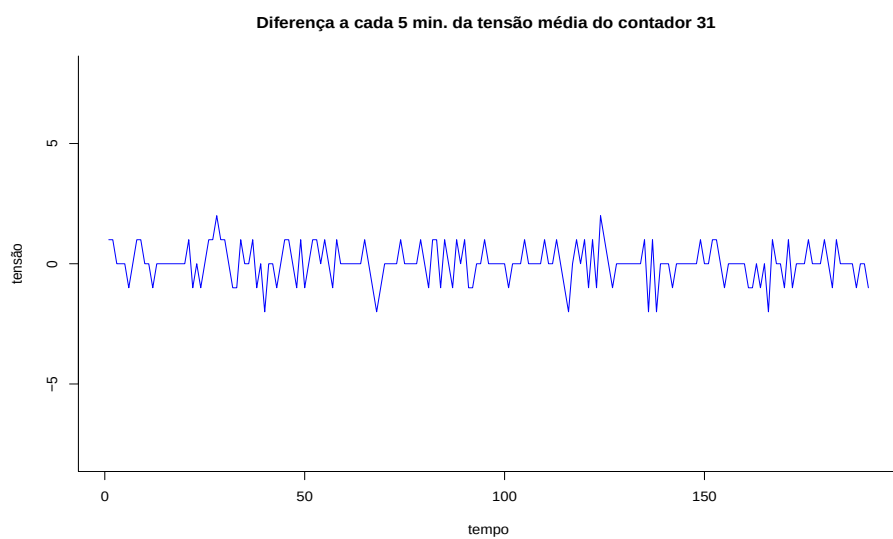


Figura 4.54: Série que contém as derivadas da tensão média do contador 31

As figuras seguintes são referentes aos gráficos das diferenças das derivadas, para os 10 contadores.

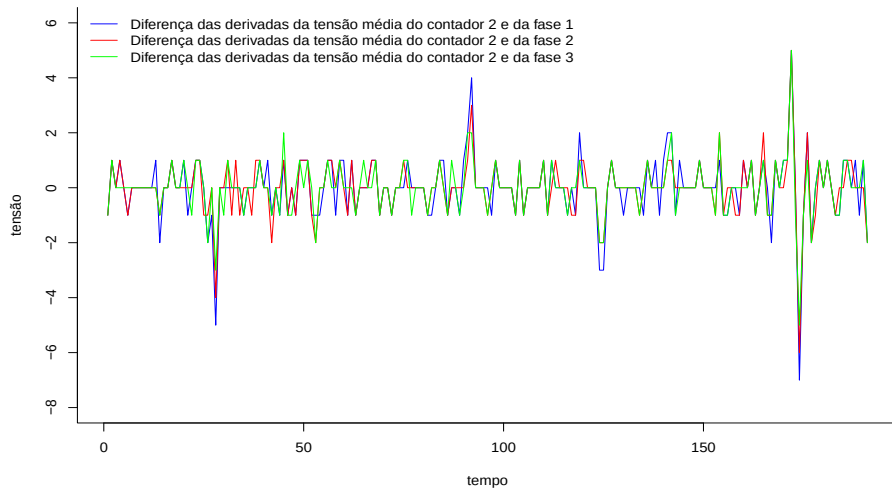


Figura 4.55: Série que contém as diferenças entre as derivadas da tensão média do contador 2 e as derivadas da tensão média das 3 fases no Posto de Transformação

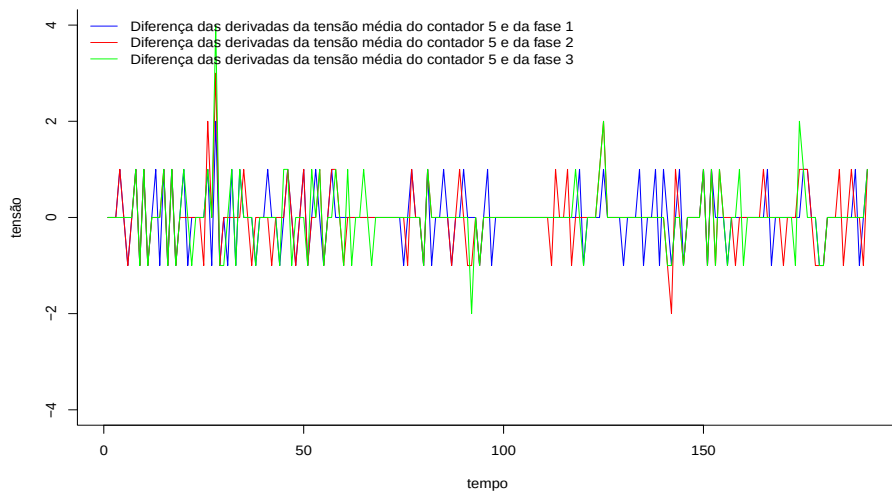


Figura 4.56: Série que contém as diferenças entre as derivadas da tensão média do contador 5 e as derivadas da tensão média das 3 fases no Posto de Transformação

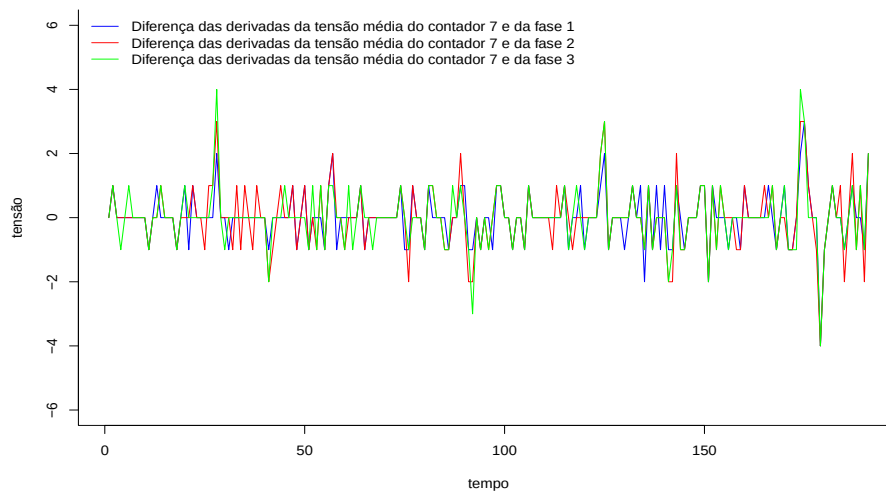


Figura 4.57: Série que contém as diferenças entre as derivadas da tensão média do contador 7 e as derivadas da tensão média das 3 fases no Posto de Transformação

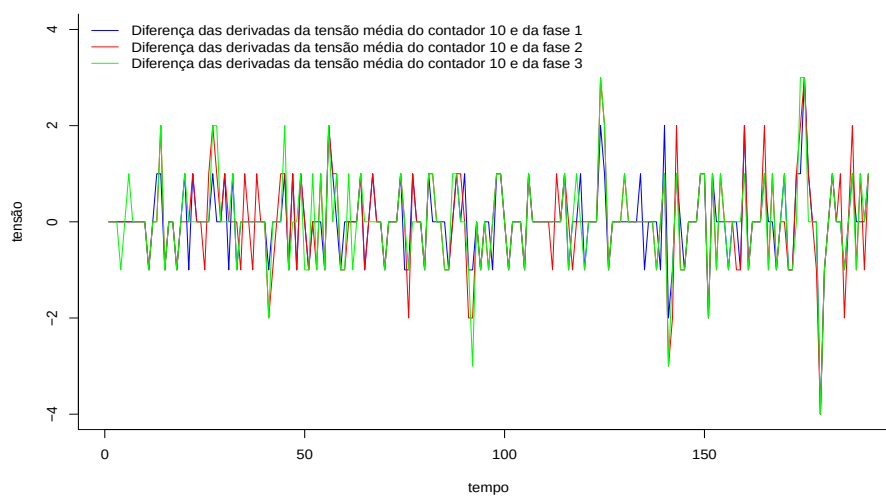


Figura 4.58: Série que contém as diferenças entre as derivadas da tensão média do contador 10 e as derivadas da tensão média das 3 fases no Posto de Transformação

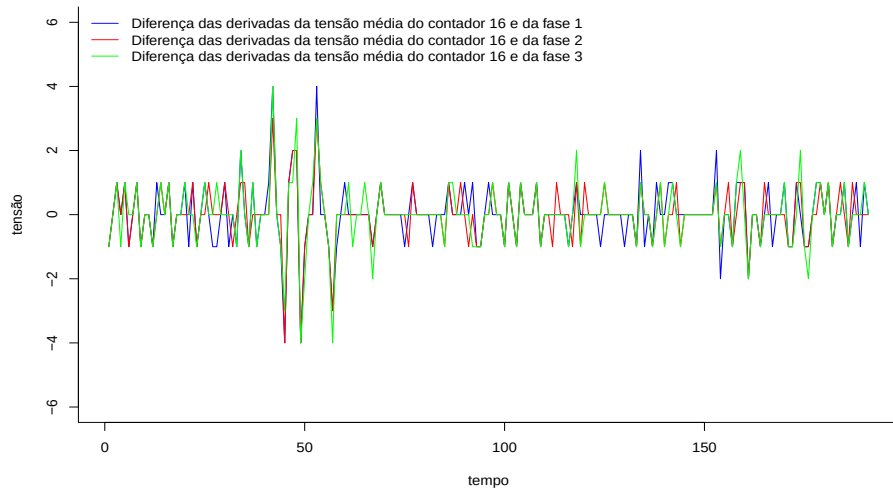


Figura 4.59: Série que contém as diferenças entre as derivadas da tensão média do contador 16 e as derivadas da tensão média das 3 fases no Posto de Transformação

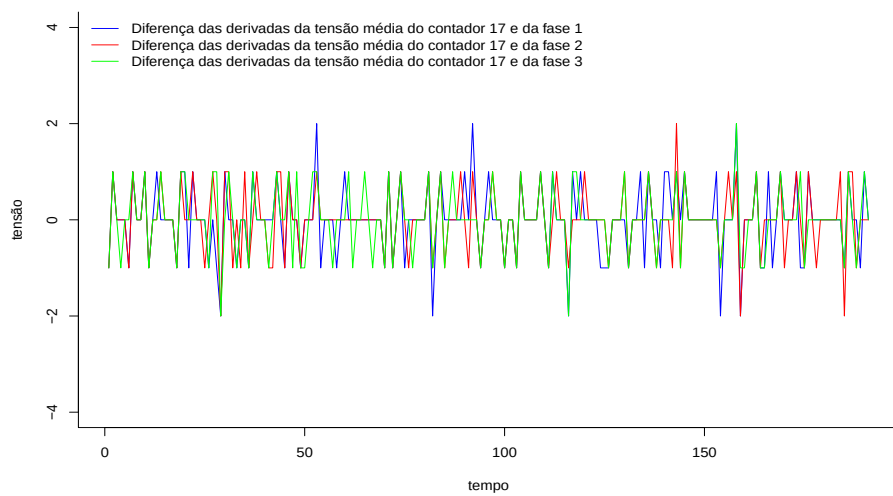


Figura 4.60: Série que contém as diferenças entre as derivadas da tensão média do contador 17 e as derivadas da tensão média das 3 fases no Posto de Transformação

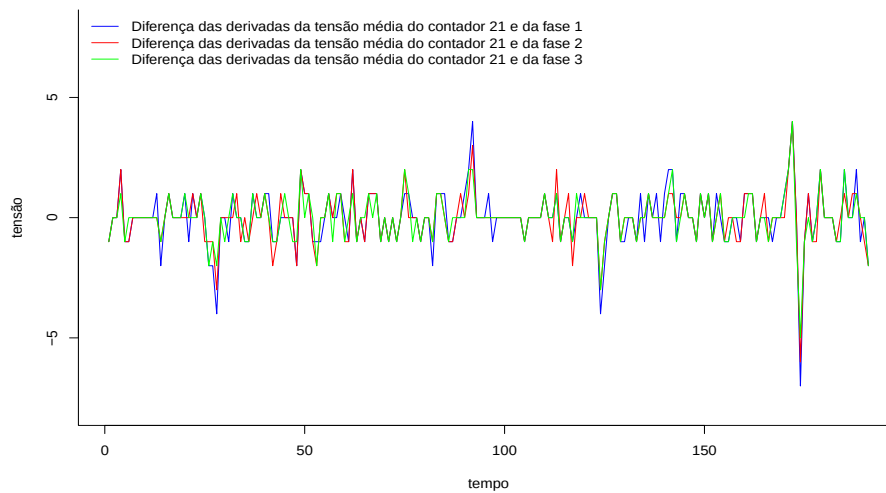


Figura 4.61: Série que contém as diferenças entre as derivadas da tensão média do contador 21 e as derivadas da tensão média das 3 fases no Posto de Transformação

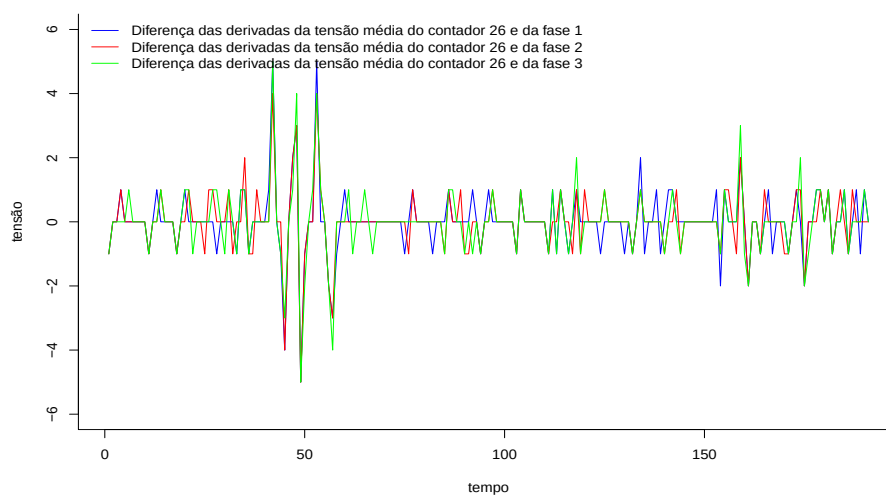


Figura 4.62: Série que contém as diferenças entre as derivadas da tensão média do contador 26 e as derivadas da tensão média das 3 fases no Posto de Transformação

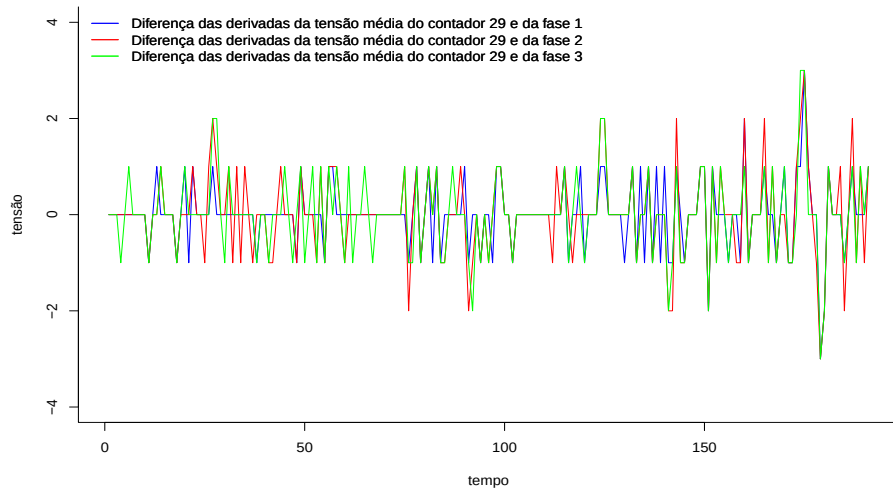


Figura 4.63: Série que contém as diferenças entre as derivadas da tensão média do contador 29 e as derivadas da tensão média das 3 fases no Posto de Transformação

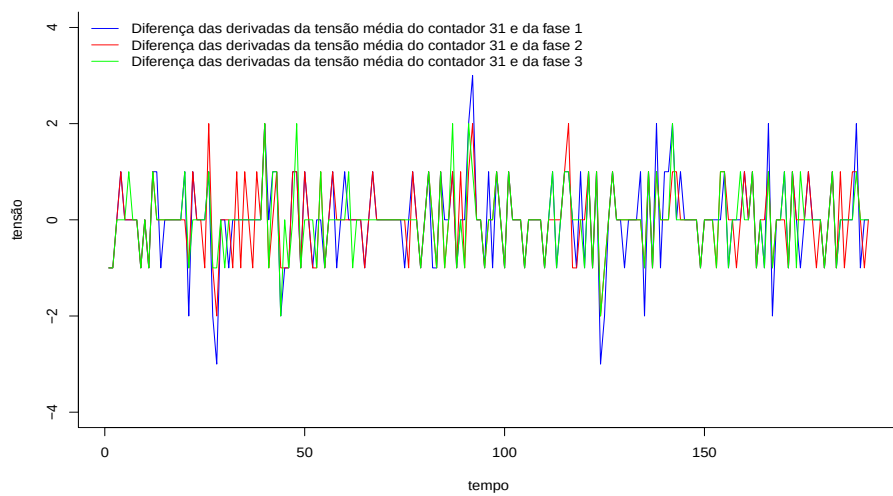


Figura 4.64: Série que contém as diferenças entre as derivadas da tensão média do contador 31 e as derivadas da tensão média das 3 fases no Posto de Transformação

Anexo C

Neste anexo serão apresentadas as figuras decorrentes da análise feita para a fase 2 e 3 do Posto de Transformação. Bem como os resultados dos testes de Ljung Box e Shapiro - Wilk para todos os contadores que não foram trabalhados na secção 3.4.2. E ainda as linhas de tendência para a fase 1.

Assim de seguida são apresentadas as figuras correspondentes às linhas de tendência.



Figura 4.65: Tendência exponencial da série de tensão média da fase 1 do Posto de Transformação

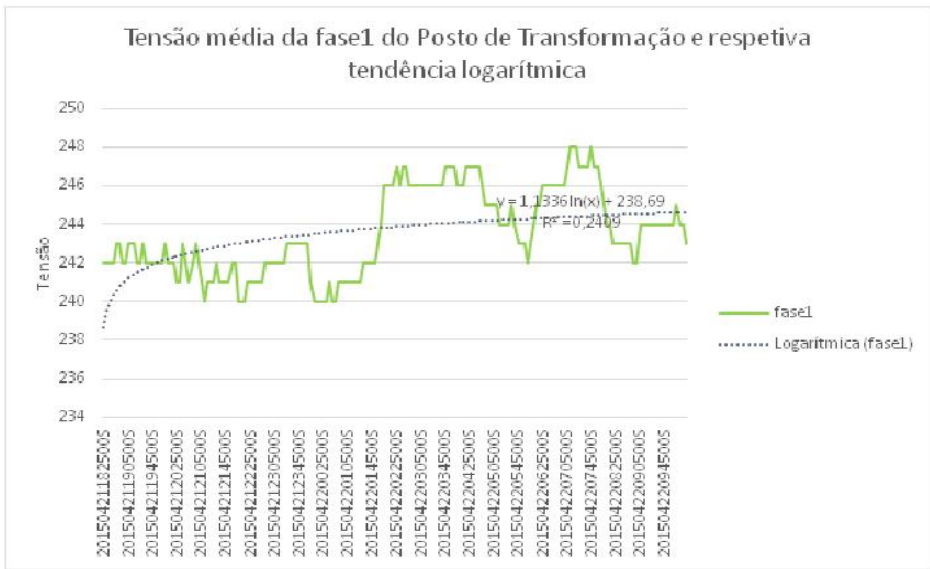


Figura 4.66: Tendência logarítmica da série de tensão média da fase 1 do Posto de Transformação

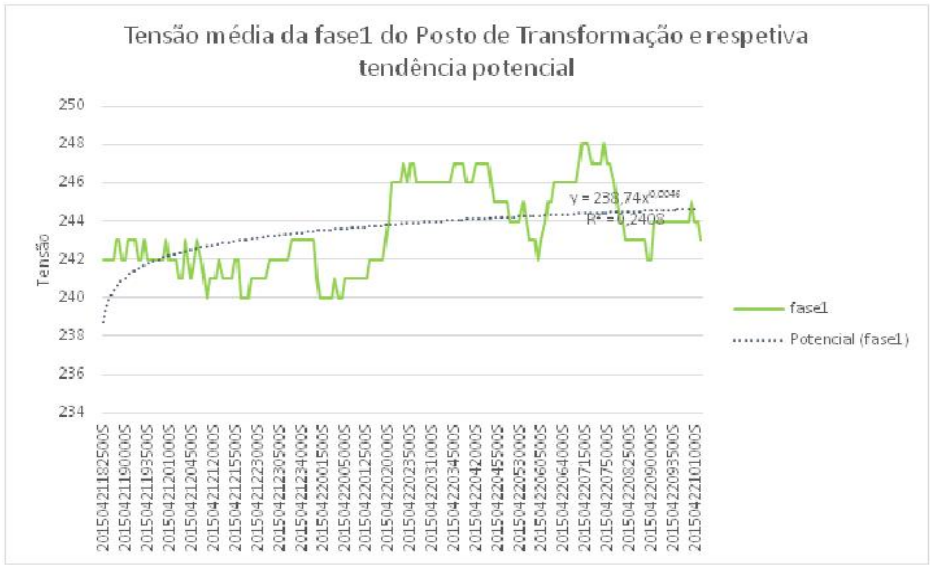


Figura 4.67: Tendência potencial da série de tensão média da fase 1 do Posto de Transformação

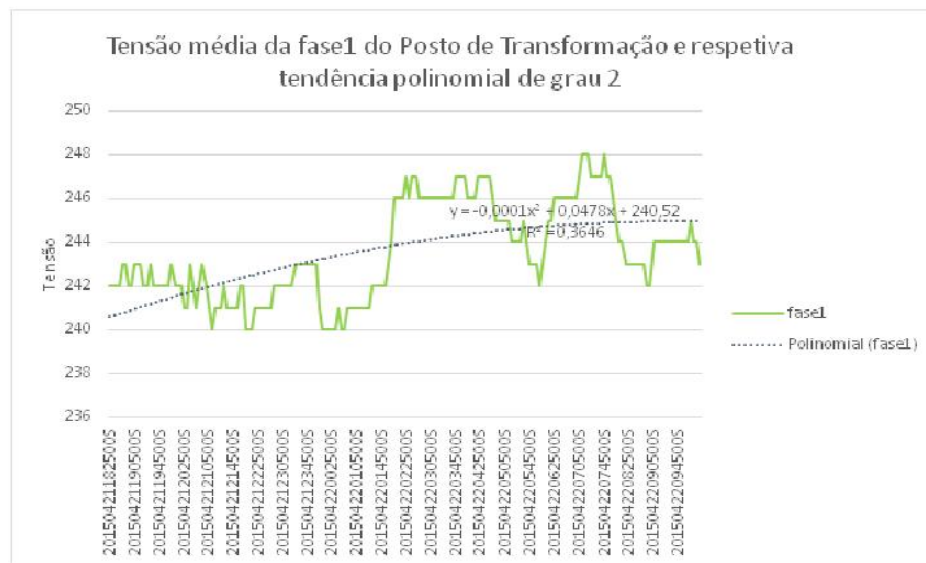


Figura 4.68: Tendência polinomial de grau 2 da série de tensão média da fase 1 do Posto de Transformação

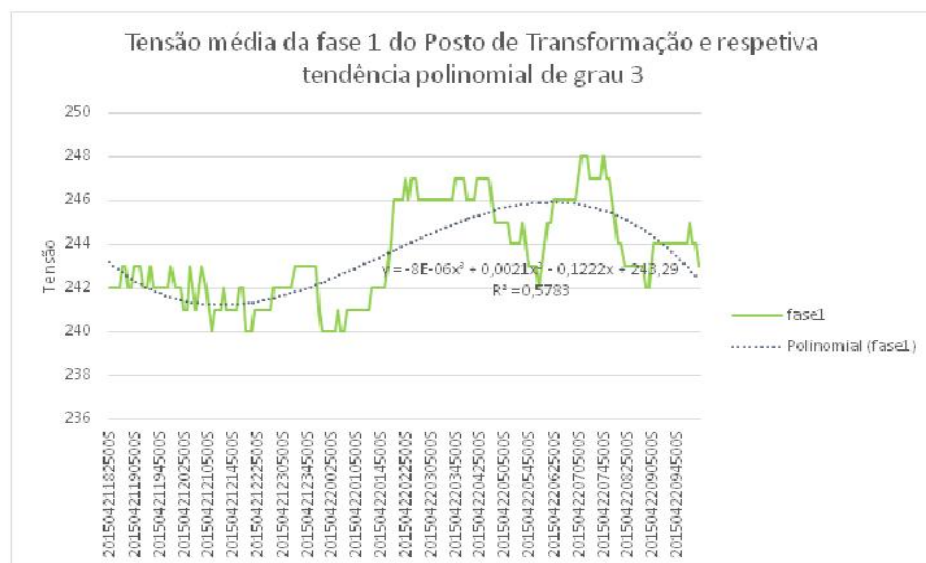


Figura 4.69: Tendência polinomial de grau 3 da série de tensão média da fase 1 do Posto de Transformação

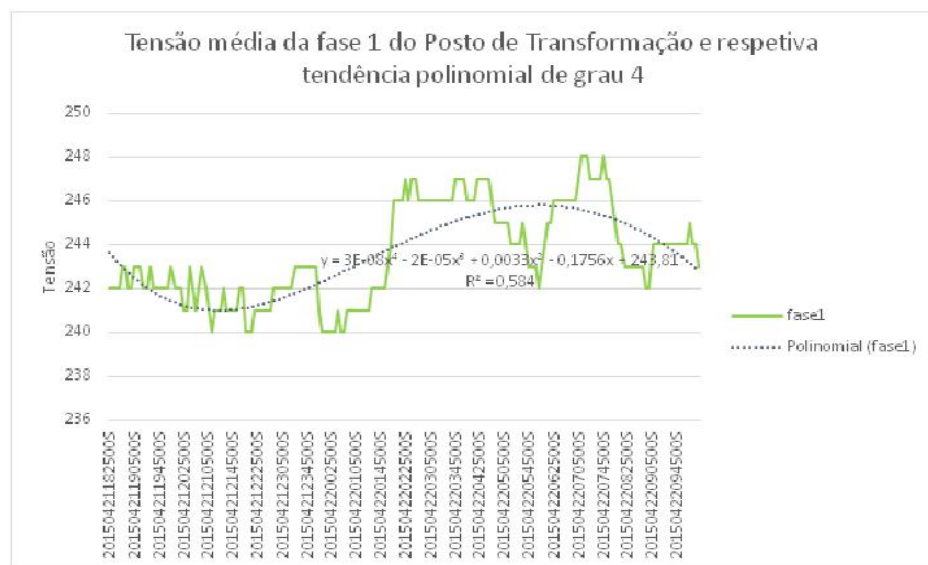


Figura 4.70: Tendência polinomial de grau 4 da série de tensão média da fase 1 do Posto de Transformação



Figura 4.71: Tendência polinomial de grau 5 da série de tensão média da fase 1 do Posto de Transformação



Figura 4.72: Tendência linear da série de tensão média da fase 1 do Posto de Transformação

Para o modelo ARIMA(10,1,7) da fase 2, a análise dos resíduos:

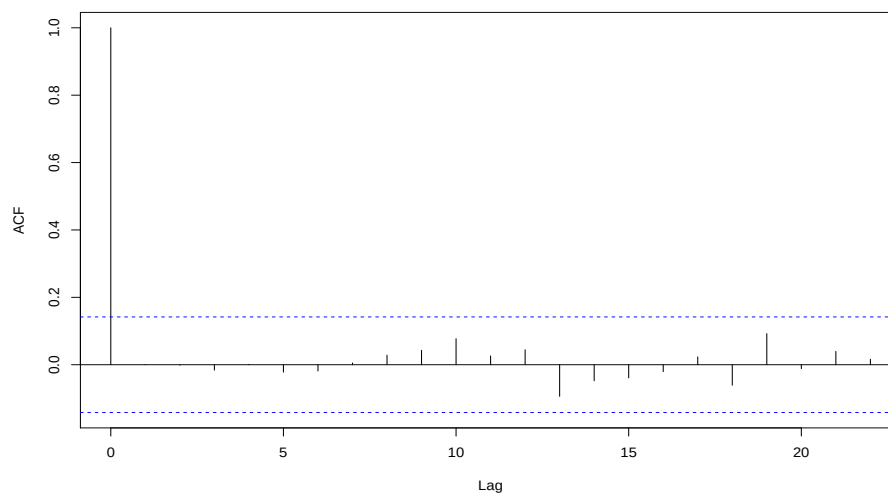


Figura 4.73: Função de autocorrelação dos resíduos do modelo para a fase 2 do Posto de Transformação

Teste para a Independência dos Resíduos	
Teste de Ljung - Box	valor p = 0.9988

Tabela 4.1: Aplicação do teste de Ljung - Box para testar a correlação dos resíduos do modelo para a fase 2

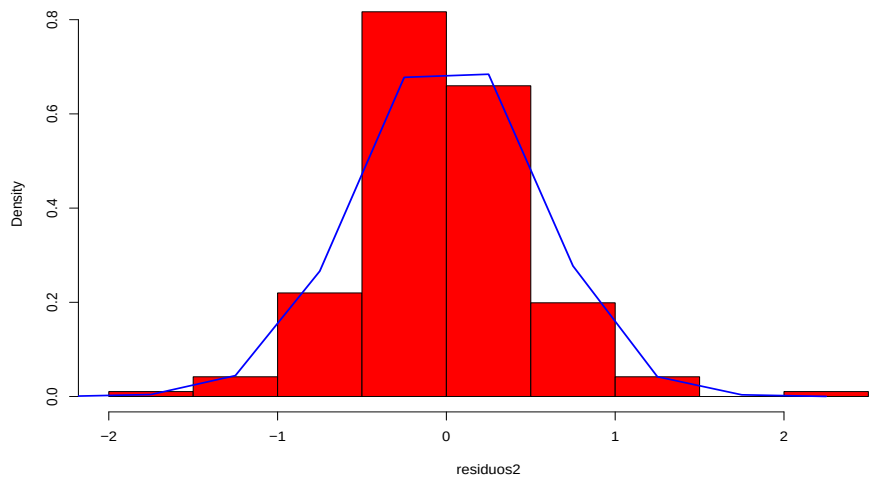


Figura 4.74: Histograma dos resíduos do modelo para a fase 2 do Posto de Transformação

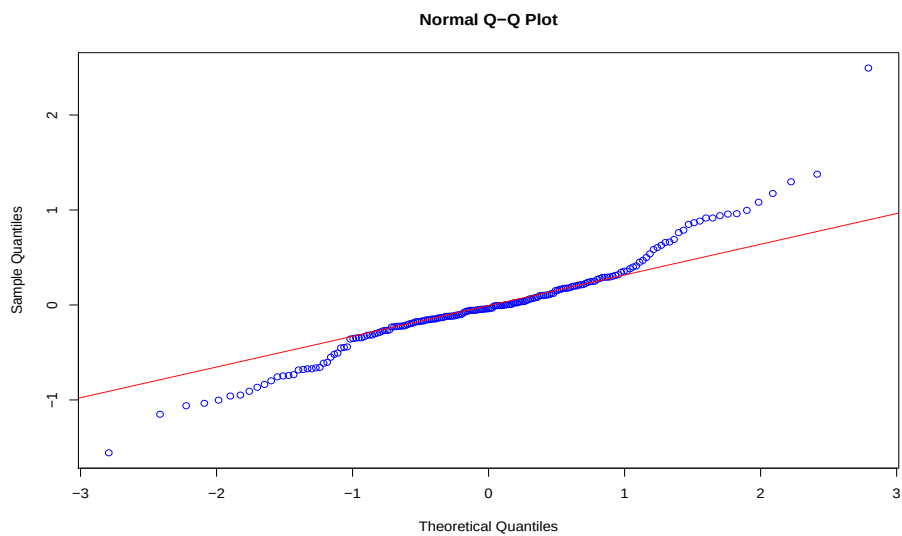


Figura 4.75: Q - Q plot dos resíduos do modelo da fase 2

Teste para a Normalidade dos Resíduos	
Teste de Shapiro - Wilk	Valor $p = 2.111e^{-6}$

Tabela 4.2: Aplicação do teste de Shapiro para a normalidade dos resíduos do modelo para a fase 2

Para o modelo ARIMA(8,1,4) fase 3,a análise dos resíduos :

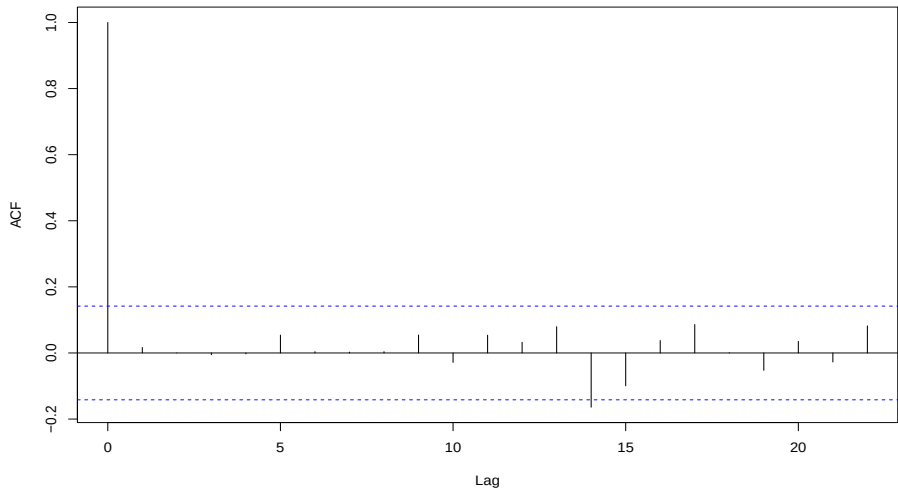


Figura 4.76: Função de autocorrelação dos resíduos do modelo para a fase 3 do Posto de Transformação

Teste para a Independência dos Resíduos	
Teste de Ljung - Box	valor p = 0.9988

Tabela 4.3: Aplicação do teste de Ljung - Box para testar a correlação dos resíduos do modelo para a fase 3

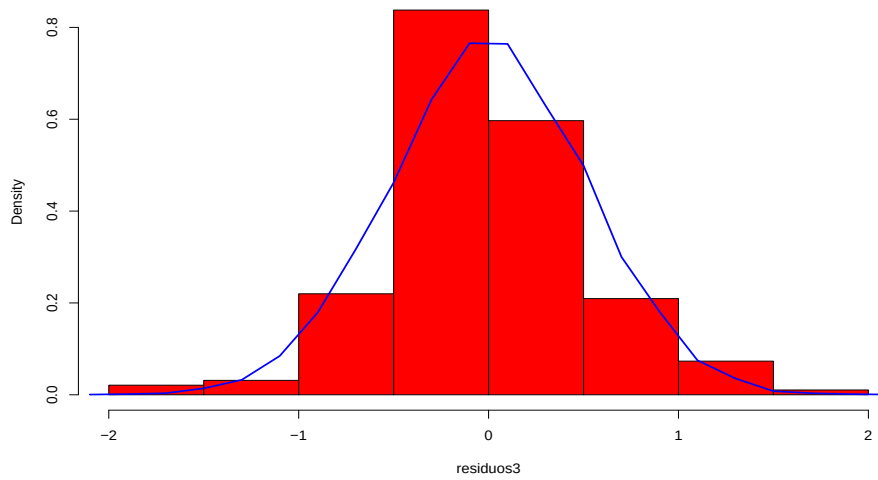


Figura 4.77: Histograma dos resíduos do modelo para a fase 3 do Posto de Transformação

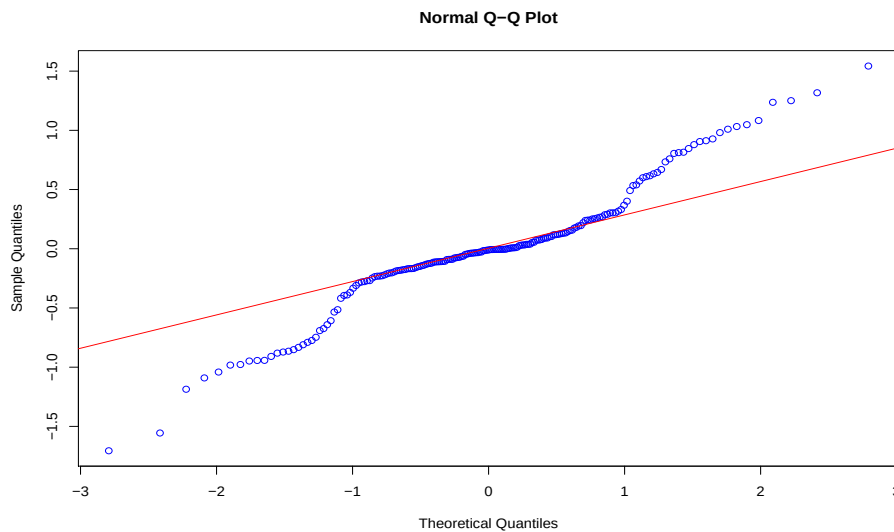


Figura 4.78: Q - Q plot dos resíduos do modelo da fase 3

Teste de Ljung - Box e de Shapiro - Wilk para a rede 1:

Teste para a Normalidade dos Resíduos	
Teste de Shapiro - Wilk	Valor p = $2.111e^{-6}$

Tabela 4.4: Aplicação do teste de Shapiro para a normalidade dos resíduos do modelo para a fase 3

Teste para a Independência e normalidade dos Resíduos para a rede 1			
Contador	Modelo Ótimo	Teste de Ljung - Box	Teste de Shapiro - Wilk
2	Fase 2	valor p =1	valor p =1.324e-09
3	Fase 1	valor p =0.8211	valor p =0.002657
4	Fase 3	valor p =0.9241	valor p =2.203e-05
5	Fase 1	valor p =0.6954	valor p =1.134e-10
6	Fase 2	valor p =0.6954	valor p =7.044e-05
7	Fase 1	valor p =0.6538	valor p =2.995e-06
8	Fase 2	valor p =0.9976	valor p =0.3219
9	Fase 2	valor p =0.9975	valor p =0.427
10	Fase 2	valor p =0.9249	valor p =9.381e-06
11	Fase 3	valor p =0.9894	valor p =0.0001591
12	Fase 1	valor p =0.9887	valor p =3.03e-10
13	Fase 1	valor p =0.9954	valor p =1.057e-06
14	Fase 1	valor p =0.9867	valor p =9.891e-08
15	Fase 1	valor p =0.9035	valor p =2.571e-07
16	Fase 1	valor p =0.9491	valor p =9.373e-07
17	Fase 3	valor p =0.6013	valor p =0.0004684

Tabela 4.5: Aplicação do teste de Ljung - Box aos resíduos do modelo ótimo para a rede 1

Teste para a Independência e normalidade dos Resíduos para a rede 1			
Contador	Modelo Ótimo	Teste de Ljung - Box	Teste de Shapiro - Wilk
18	Fase 1	valor p =0.2146	valor p =0.009116
19	Fase 1	valor p =0.9741	valor p =4.923e-11
20	Fase 2	valor p =0.9191	valor p =0.09939
21	Fase 1	valor p =0.911	valor p =7.213e-07
22	Fase 2	valor p =0.9487	valor p =0.2184
23	Fase 2	valor p =0.8923	valor p =0.00181
24	Fase 3	valor p =0.9838	valor p =0.009794
25	Fase 1	valor p =0.8561	valor p =1.248e-07
26	Fase 1	valor p =0.9934	valor p =2.18e-09
27	Fase 2	valor p =0.9953	valor p =0.0003677
28	Fase 2	valor p =0.9866	valor p =0.03596
29	Fase 1	valor p = 0.1317	valor p =1.216e-06
30	Fase 2	valor p =0.9997	valor p =0.001086
31	Fase 2	valor p =0.758	valor p =0.199
32	Fase 1	valor p =0.7216	valor p =7.553e-08
33	Fase 3	valor p =0.6966	valor p =0.4828
34	Fase 1	valor p = 0.9137	valor p =3.163e-07
35	Fase 2	valor p =1	valor p =7.711e-07
36	Fase 1	valor p =1	valor p < 2.2e - 16

Tabela 4.6: Aplicação do teste de Ljung - Box aos resíduos do modelo ótimo para a rede 1

Anexo D

Neste anexo são apresentadas tabelas com o valor das medidas de dispersão para as séries das diferenças de cada contador e as medidas de dispersão para as séries das diferenças das derivadas para cada contador. São também apresentadas as tabelas com os valores do teste de hipótese das mesmas séries para cada contador, o valor do coeficiente de correlação de Spearman para as séries das tensões e as séries das derivadas. E ainda os resultados dos critérios para a rede 1, a 1ª metade da rede 1 e a 2º metade, bem como a sua classificação decorrente de cada critério e o respetivo sistema de votos.

Medidas de dispersão calculadas para cada uma das séries de diferenças de cada contador.

Medidas de Dispersão para as Três Séries de Diferenças dos Contadores			
Contador	Diferença	Desvio Padrão	Variância
Contador 1	Diferença Fase 1	0.5693602	0.324171
	Diferença Fase 2	1.46154	2.136098
	Diferença Fase 3	1.596508	2.548838
Contador 2	Diferença Fase 1	3.005227	9.031386
	Diferença Fase 2	2.031144	4.125545
	Diferença Fase 2	1.833254	3.36082
Contador 3	Diferença Fase 1	2.695073	7.263416
	Diferença Fase 2	1.728008	2.986011
	Diferença Fase 3	1.511076	2.28335
Contador 4	Diferença Fase 1	0.9378017	0.8794721
	Diferença Fase 2	0.5211234	0.2715696
	Diferença Fase 3	0.6719698	0.4515434
Contador 5	Diferença Fase 1	0.5365156	0.287849
	Diferença Fase 2	1.366186	1.866465
	Diferença Fase 3	1.507317	2.272006

Tabela 4.7: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças dos 36 contadores

Medidas de Dispersão para as Três Séries de Diferenças dos Contadores			
Contador	Diferença	Desvio Padrão	Variância
Contador 6	Diferença Fase 1	2.673965	7.150087
	Diferença Fase 2	1.695188	2.873664
	Diferença Fase 3	1.488384	2.215287
Contador 7	Diferença Fase 1	1.059271	1.122055
	Diferença Fase 2	1.965445	3.862974
	Diferença Fase 3	2.065194	4.265025
Contador 8	Diferença Fase 1	1.31178	1.720768
	Diferença Fase 2	0.6616645	0.4378
	Diferença Fase 3	0.8429042	0.7104876
Contador 9	Diferença Fase 1	1.332969	1.776805
	Diferença Fase 2	0.6927699	0.4799302
	Diferença Fase 3	0.8296822	0.6883726
Contador 10	Diferença Fase 1	1.084105	1.175284
	Diferença Fase 2	2.011251	4.04513
	Diferença Fase 3	2.107696	4.442381
Contador 11	Diferença Fase 1	1.233784	1.522224
	Diferença Fase 2	0.6995854	0.4894197
	Diferença Fase 3	0.8800186	0.7744328
Contador 12	Diferença Fase 1	1.451965	2.108202
	Diferença Fase 2	1.650194	2.72314
	Diferença Fase 3	1.905997	3.632826
Contador 13	Diferença Fase 1	3.276906	10.73811
	Diferença Fase 2	2.298981	5.285313
	Diferença Fase 3	2.086631	4.35403
Contador 14	Diferença Fase 1	1.425124	2.030977
	Diferença Fase 2	1.753075	3.073271
	Diferença Fase 3	2.007614	4.030514

Tabela 4.8: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças dos 36 contadores

Medidas de Dispersão para as Três Séries de Diferenças dos Contadores			
Contador	Diferença	Desvio Padrão	Variância
Contador 15	Diferença Fase 1	1.521085	2.3137
	Diferença Fase 2	1.181212	1.395261
	Diferença Fase 3	1.328296	1.764371
Contador 16	Diferença Fase 1	1.182596	1.398533
	Diferença Fase 2	1.281114	1.641252
	Diferença Fase 3	1.543437	2.382199
Contador 17	Diferença Fase	1.244874	1.549711
	Diferença Fase 2	0.6220918	0.3869983
	Diferença Fase 3	0.7842666	0.6150742
Contador 18	Diferença Fase 1	2.066283	4.269524
	Diferença Fase 2	1.129668	1.276151
	Diferença Fase 3	0.9930908	0.9862293
Contador 19	Diferença Fase 1	2.165895	4.691099
	Diferença Fase 2	1.323278	1.751063
	Diferença Fase 3	1.250905	1.564763
Contador 20	Diferença Fase 3	1.945147	3.783595
	Diferença Fase 3	1.089262	1.186491
	Diferença Fase 3	1.011375	1.022878
Contador 21	Diferença Fase 1	3.241582	10.50785
	Diferença Fase 2	2.256118	5.090069
	Diferença Fase 3	2.053116	4.215287
Contador 22	Diferença Fase 1	1.233221	1.520833
	Diferença Fase 2	0.7098587	0.5038994
	Diferença Fase 3	0.8956063	0.8021106
Contador 23	Diferença Fase 1	2.165895	4.691099
	Diferença Fase 2	1.323278	1.751063
	Diferença Fase 3	1.250905	1.564763

Tabela 4.9: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças dos 36 contadores

Medidas de Dispersão para as Três Séries de Diferenças dos Contadores			
Contador	Diferença	Desvio Padrão	Variância
Contador 24	Diferença Fase 1	1.241101	1.54033
	Diferença Fase 2	0.7137854	0.5094895
	Diferença Fase 3	0.9056734	0.8202443
Contador 25	Diferença Fase 1	1.264661	1.599367
	Diferença Fase 2	1.432824	2.052983
	Diferença Fase 3	1.698467	2.884789
Contador 26	Diferença Fase 1	1.278045	1.633399
	Diferença Fase 2	1.457953	2.125627
	Diferença Fase 3	1.704813	2.906386
Contador 27	Diferença Fase 1	2.005229	4.020942
	Diferença Fase 2	1.255257	1.575671
	Diferença Fase 3	1.172034	1.373664
Contador 28	Diferença Fase 1	1.285231	1.13368
	Diferença Fase 2	0.3285068	0.5731551
	Diferença Fase 3	0.5691263	0.7544046
Contador 29	Diferença Fase 1	0.9346999	0.8736638
	Diferença Fase 2	1.795835	3.225022
	Diferença Fase 3	1.890548	3.574171
Contador 30	Diferença Fase 1	2.280641	5.201325
	Diferença Fase 2	1.314439	1.727749
	Diferença Fase 3	1.106807	1.225022
Contador 31	Diferença Fase 1	2.14005	4.579816
	Diferença Fase 2	1.183391	1.400414
	Diferença Fase 3	1.038315	1.078098
Contador 32	Diferença Fase 1	0.9976931	0.9953916
	Diferença Fase 2	1.789263	3.201462
	Diferença Fase 3	1.871528	3.502618

Tabela 4.10: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças dos 36 contadores

Medidas de Dispersão para as Três Séries de Diferenças dos Contadores			
Contador	Diferença	Desvio Padrão	Variância
Contador 33	Diferença Fase 1	1.324853	1.755236
	Diferença Fase 2	0.7372942	0.5436027
	Diferença Fase 3	0.8832199	0.7800774
Contador 34	Diferença Fase 1	1.058447	1.12031
	Diferença Fase 2	1.830329	3.350104
	Diferença Fase 3	1.914733	3.666203
Contador 35	Diferença Fase 1	3.224308	10.39616
	Diferença Fase 2	2.231227	4.978376
	Diferença Fase 3	2.029176	4.117556
Contador 36	Diferença Fase 1	4.015676	16.12565
	Diferença Fase 2	3.53661	12.50761
	Diferença Fase 3	3.483972	12.13806

Tabela 4.11: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças dos 36 contadores

Os valores obtidos do teste de Brown - Forsythe para os 36 contadores encontra-se na tabela seguinte:

Valor do Teste de Hipótese de Brown-Forsythe				
Contador	Todas as fases	Fase 1 e 2	Fase 1 e 3	Fase 2 e 3
Contador 1	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.1022$
Contador 2	valor $p = 4.436e^{-11}$	valor $p = 7.726e^{-07}$	valor $p = 4.072e^{-09}$	valor $p = 0.229$
Contador 3	valor $p = 9.82e^{-13}$	valor $p = 1.941e^{-07}$	valor $p = 2.679e^{-10}$	valor $p = 0.1252$
Contador 4	valor $p = 2.252e^{-15}$	valor $p = 3.673e^{-15}$	valor $p = 2.084e^{-07}$	valor $p = 0.003521$
Contador 5	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.06659$
Contador 6	valor $p = 4.777e^{-15}$	valor $p = 5.528e^{-10}$	valor $p = 2.474e^{-13}$	valor $p = 0.3492$
Contador 7	valor $p < 2.2e^{-16}$	valor $p = 2.227e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.4781$
Contador 8	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 2.932e^{-10}$	valor $p = 0.002652$
Contador 9	valor $p = 5.495e^{-12}$	valor $p = 3.517e^{-10}$	valor $p = 7.614e^{-08}$	valor $p = 0.1967$
Contador 10	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.4221$
Contador 11	valor $p = 1.365e^{-12}$	valor $p = 2.157e^{-12}$	valor $p = 6.786e^{-06}$	valor $p = 0.004813$
Contador 12	valor $p = 0.0362$	valor $p = 0.176$	valor $p = 0.01109$	valor $p = 0.2109$
Contador 13	valor $p = 1.786e^{-09}$	valor $p = 5.385e^{-06}$	valor $p = 5.246e^{-08}$	valor $p = 0.2743$
Contador 14	valor $p = 2.92e^{-06}$	valor $p = 0.0001568$	valor $p = 1.91e^{-06}$	valor $p = 0.202$
Contador 15	valor $p = 0.2256$	-	-	-
Contador 16	valor $p = 0.1225$	-	-	-
Contador 17	valor $p = 7.034e^{-12}$	valor $p = 7.809e^{-11}$	valor $p = 1.463e^{-06}$	valor $p = 0.03811$
Contador 18	valor $p < 2.2e^{-16}$	valor $p = 6.706e^{-12}$	valor $p = 1.067e^{-15}$	valor $p = 0.02979$
Contador 19	valor $p = 2.231e^{-11}$	valor $p = 5.664e^{-07}$	valor $p = 1.114e^{-09}$	valor $p = 0.1737$
Contador 20	valor $p < 2.2e^{-16}$	valor $p = 1.587e^{-10}$	valor $p = 2.558e^{-13}$	valor $p = 0.1038$
Contador 21	valor $p = 1.069e^{-10}$	valor $p = 1.19e^{-06}$	valor $p = 7.024e^{-09}$	valor $p = 0.2428$
Contador 22	valor $p = 4.712e^{-11}$	valor $p = 6.107e^{-11}$	valor $p = 2.576e^{-05}$	valor $p = 0.008767$
Contador 23	valor $p = 2.467e^{-14}$	valor $p = 2.126e^{-09}$	valor $p = 2.176e^{-11}$	valor $p = 0.3007$

Tabela 4.12: Valor do teste de hipótese de Brown-Forsythe para as séries das diferenças dos 36 Contadores

Valor do Teste de Hipótese de Brown-Forsythe				
Contador	Todas as fases	Fase 1 e 2	Fase 1 e 3	Fase 2 e 3
Contador 24	valor $p = 8.41e^{-10}$	valor $p = 2.58e^{-10}$	valor $p = 0.0002682$	valor $p = 0.004729$
Contador 25	valor $p = 0.01019$	valor $p = 0.1067$	valor $p = 0.003047$	valor $p = 0.1425$
Contador 26	valor $p = 0.0791$	-	-	-
Contador 27	valor $p = 2.509e^{-14}$	valor $p = 2.004e^{-08}$	valor $p = 1.53e^{-11}$	valor $p = 0.07114$
Contador 28	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 4.943e^{-10}$	valor $p = 0.0004327$
Contador 29	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.3645$
Contador 30	valor $p < 2.2e^{-16}$	valor $p = 1.676e^{-13}$	valor $p < 2.2e^{-16}$	valor $p = 0.09207$
Contador 31	valor $p = 2.1e^{-12}$	valor $p < 2.2e^{-16}$	valor $p = 0.01958$	valor $p = 2.717e^{-09}$
Contador 32	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.3403$
Contador 33	valor $p = 3.583e^{-16}$	valor $p = 9.771e^{-15}$	valor $p = 9.495e^{-09}$	valor $p = 0.02151$
Contador 34	valor $p < 2.2e^{-16}$	valor $p = 2.556e^{-16}$	valor $p < 2.2e^{-16}$	valor $p = 0.413$
Contador 35	valor $p = 2.184e^{-09}$	valor $p = 1.168e^{-06}$	valor $p = 1.151e^{-08}$	valor $p = 0.4684$
Contador 36	valor $p = 0.006096$	valor $p = 0.013$	valor $p = 0.003489$	valor $p = 0.6457$

Tabela 4.13: Valor do teste de hipótese de Brown-Forsythe para as séries das diferenças dos 36 contadores

As medidas de dispersão das séries das derivadas dos 36 contadores encontram-se a seguir representadas.

Medidas de Dispersão para as Três Séries de Diferenças das Derivadas dos Contadores			
Contador	Diferença	Desvio Padrão	Variância
Contador 1	Diferença Fase 1	0.598243	0.3578947
	Diferença Fase 2	0.7181848	0.5157895
	Diferença Fase 3	0.672779	0.4526316
Contador 2	Diferença Fase 1	1.167532	1.36313
	Diferença Fase 2	1.018241	1.036815
	Diferença Fase 3	0.9542006	0.9104988
Contador 3	Diferença Fase 1	1.038711	1.07892
	Diferença Fase 2	0.9262112	0.8578672
	Diferença Fase 3	0.8175534	0.6683935
Contador 4	Diferença Fase 1	0.6324555	0.4
	Diferença Fase 2	0.5428967	0.2947368
	Diferença Fase 3	0.598243	0.3578947
Contador 5	Diferença Fase 1	0.6112746	0.3736567
	Diferença Fase 2	0.6529075	0.4262882
	Diferença Fase 3	0.6529075	0.4262882
Contador 6	Diferença Fase 1	1.038711	1.07892
	Diferença Fase 2	0.9031954	0.8157619
	Diferença Fase 3	0.8366436	0.6999724
Contador 7	Diferença Fase 1	0.8077879	0.6525214
	Diferença Fase 2	0.9840279	0.9683108
	Diferença Fase 3	0.9346518	0.873574

Tabela 4.14: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças das derivadas dos 36 contadores

Medidas de Dispersão para as Três Séries de Diferenças das Derivadas dos Contadores

Contador	Diferença	Desvio Padrão	Variância
Contador 8	Diferença Fase 1	0.7254763	0.5263158
	Diferença Fase 2	0.680557	0.4631579
	Diferença Fase 3	0.7745967	0.6
Contador 9	Diferença Fase 1	0.8078562	0.6526316
	Diferença Fase 2	0.7326951	0.5368421
	Diferença Fase 3	0.7677719	0.5894737
Contador 10	Diferença Fase 1	0.8110901	0.6578672
	Diferença Fase 2	0.99206	0.984183
	Diferença Fase 3	0.9375072	0.8789198
Contador 11	Diferença Fase 1	0.7070873	0.4999724
	Diferença Fase 2	0.6688355	0.4473409
	Diferença Fase 3	0.6766588	0.4578672
Contador 12	Diferença Fase 1	1.130893	1.27892
	Diferença Fase 2	1.048796	1.099972
	Diferença Fase 3	1.12623	1.268393
Contador 13	Diferença Fase 1	1.27526	1.626288
	Diferença Fase 2	1.149359	1.321025
	Diferença Fase 3	1.018241	1.036815
Contador 14	Diferença Fase 1	1.100239	1.210526
	Diferença Fase 2	1.020836	1.042105
	Diferença Fase 3	1.109765	1.231579
Contador 15	Diferença Fase 1	0.8735733	0.7631303
	Diferença Fase 2	0.8110901	0.6578672
	Diferença Fase 3	0.8429109	0.7104988
Contador 16	Diferença Fase 1	0.9733285	0.9473684
	Diferença Fase 2	0.8705715	0.7578947
	Diferença Fase 3	0.9569689	0.9157895
Contador 17	Diferença Fase 1	0.7326951	0.5368421
	Diferença Fase 2	0.680557	0.4631579
	Diferença Fase 3	0.66491	0.4421053

Tabela 4.15: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças das derivadas dos 36 Contadores

Medidas de Dispersão para as Três Séries de Diferenças das Derivadas dos Contadores

Contador	Diferença	Desvio Padrão	Variância
Contador 18	Diferença Fase 1	0.8855411	0.784183
	Diferença Fase 2	0.7504201	0.5631303
	Diferença Fase 3	0.7433734	0.552604
Contador 19	Diferença Fase 1	1.3416	1.79989
	Diferença Fase 2	1.231129	1.515679
	Diferença Fase 3	1.133181	1.2841
Contador 20	Diferença Fase 1	0.9088827	0.8260678
	Diferença Fase 2	0.771031	0.5944888
	Diferença Fase 3	0.7641744	0.5839625
Contador 21	Diferença Fase 1	1.183169	1.39989
	Diferença Fase 2	1.041201	1.0841
	Diferença Fase 3	0.945847	0.8946266
Contador 22	Diferença Fase 1	0.7608859	0.5789474
	Diferença Fase 2	0.7608859	0.5789474
	Diferença Fase 3	0.753937	0.5684211
Contador 23	Diferença Fase 1	0.9782141	0.9569027
	Diferença Fase 2	0.8006955	0.6411133
	Diferença Fase 3	0.8265703	0.6832185
Contador 24	Diferença Fase 1	0.7469235	0.5578947
	Diferença Fase 2	0.6488857	0.4210526
	Diferença Fase 3	0.7469235	0.5578947
Contador 25	Diferença Fase 1	1.046297	1.094737
	Diferença Fase 2	0.9459053	0.8947368
	Diferença Fase 3	1.025978	1.052632
Contador 26	Diferença Fase 1	1.005249	1.010526
	Diferença Fase 2	0.9459053	0.8947368
	Diferença Fase 3	1.020836	1.042105
Contador 27	Diferença Fase 1	0.9838598	0.9679802
	Diferença Fase 2	0.8941807	0.7995591
	Diferença Fase 3	0.8395369	0.7048223

Tabela 4.16: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças das derivadas dos 36 Contadores

Medidas de Dispersão para as Três Séries de Diferenças das Derivadas dos Contadores			
Contador	Diferença	Desvio Padrão	Variância
Contador 28	Diferença Fase 1	0.6529075	0.4262882
	Diferença Fase 2	0.619825	0.384183
	Diferença Fase 3	0.6282589	0.3947093
Contador 29	Diferença Fase 1	0.7070873	0.4999724
	Diferença Fase 2	0.8855411	0.784183
	Diferença Fase 3	0.8366436	0.6999724
Contador 30	Diferença Fase 1	0.849132	0.7210251
	Diferença Fase 2	0.7504201	0.5631303
	Diferença Fase 3	0.6843928	0.4683935
Contador 31	Diferença Fase 1	0.9002311	0.8104161
	Diferença Fase 2	0.7608135	0.5788371
	Diferença Fase 3	0.7181081	0.5156793
Contador 32	Diferença Fase 1	0.8110901	0.6578672
	Diferença Fase 2	0.9431045	0.8894461
	Diferença Fase 3	0.9205112	0.8473409
Contador 33	Diferença Fase 1	0.8303289	0.6894461
	Diferença Fase 2	0.8110901	0.6578672
	Diferença Fase 3	0.8175534	0.6683935
Contador 34	Diferença Fase 1	0.784705	0.6157619
	Diferença Fase 2	0.9318763	0.8683935
	Diferença Fase 3	0.8914647	0.7947093
Contador 35	Diferença Fase 1	1.149359	1.321025
	Diferença Fase 2	1.033631	1.068393
	Diferença Fase 3	0.9542006	0.9104988
Contador 36	Diferença Fase 1	4.74396	22.50515
	Diferença Fase 2	4.719489	22.27357
	Diferença Fase 3	4.719489	22.27357

Tabela 4.17: Medidas de dispersão - desvio padrão e variância - para as três séries de diferenças das derivadas dos 36 Contadores

Os valores obtidos do teste de Brown - Forsythe entre as séries das derivadas para os 36 contadores encontra-se na seguinte tabela:

Valor do Teste de Hipótese de Brown-Forsythe				
Contador	Todas as fases	Fase 1 e 2	Fase 1 e 3	Fase 2 e 3
Contador 1	valor p = 0.2833	-	-	-
Contador 2	valor p = 0.2098	-	-	-
Contador 3	valor p = 0.1266	-	-	-
Contador 4	valor p = 0.09707	-	-	-
Contador 5	valor p = 0.8021	-	-	-
Contador 6	valor p = 0.2253	-	-	-
Contador 7	valor p = 0.184	-	-	-
Contador 8	valor p = 0.4262	-	-	-
Contador 9	valor p = 0.4267	-	-	-
Contador 10	valor p = 0.09754	-	-	-
Contador 11	valor p = 0.7244	-	-	-
Contador 12	valor p = 0.7456	-	-	-
Contador 13	valor p = 0.1605	-	-	-
Contador 14	valor p = 0.7045	-	-	-
Contador 15	valor p = 0.9043	-	-	-
Contador 16	valor p = 0.5584	-	-	-
Contador 17	valor p = 0.7377	-	-	-
Contador 18	valor p = 0.2519	-	-	-
Contador 19	valor p = 0.2945	-	-	-
Contador 20	valor p = 0.1641	-	-	-
Contador 21	valor p = 0.1126	-	-	-
Contador 22	valor p = 0.8169	-	-	-

Tabela 4.18: Valor do teste de hipótese de Brown-Forsythe para as séries de diferenças das derivadas dos 36 Contadores

Valor do Teste de Hipótese de Brown-Forsythe				
Contador	Todas as fases	Fase 1 e 2	Fase 1 e 3	Fase 2 e 3
Contador 23	valor p = 0.08135	-	-	-
Contador 24	valor p = 0.2034	-	-	-
Contador 25	valor p = 0.4011	-	-	-
Contador 26	valor p = 0.8928	-	-	-
Contador 27	valor p = 0.2001	-	-	-
Contador 28	valor p = 0.8972	-	-	-
Contador 29	valor p = 0.04725	valor p = 0.01604	valor p = 0.06982	valor p = 0.5418
Contador 30	valor p = 0.1605	-	-	-
Contador 31	valor p = 0.06181	-	-	-
Contador 32	valor p = 0.2423	-	-	-
Contador 33	valor p = 0.8277	-	-	-
Contador 34	valor p = 0.2135	-	-	-
Contador 35	valor p = 0.2523	-	-	-
Contador 36	valor p = 0.973	-	-	-

Tabela 4.19: Valor do teste de hipótese de Brown-Forsythe para as séries de diferenças das derivadas dos 36 Contadores

O coeficiente de correlação de Spearman entre as séries de tensão média das fases do Posto de Transformação e a tensão média dos contadores, para os 36 contadores encontra-se na seguinte tabela:

Coeficiente de Correlação de Spearman para as Séries de Tensão

Contador	Fase 1	Fase 2	Fase 3
Contador 1	0.9658074	0.763069	0.6990243
Contador 2	0.1734574	0.5968477	0.6915795
Contador 3	0.2587017	0.6542079	0.7477487
Contador 4	0.8906105	0.9641781	0.9332925
Contador 5	0.9698603	0.8055055	0.745909
Contador 6	0.2738124	0.6731214	0.7620442
Contador 7	0.9279967	0.7043779	0.6564593
Contador 8	0.8312952	0.9430018	0.9008848
Contador 9	0.8224295	0.9395011	0.9092664
Contador 10	0.9275524	0.6968279	0.648839
Contador 11	0.8495979	0.937144	0.8960356
Contador 12	0.9225147	0.8627365	0.7780241
Contador 13	0.069239	0.5125564	0.6170697
Contador 14	0.9351943	0.8338727	0.7459775
Contador 15	0.7701024	0.8137705	0.7605207
Contador 16	0.9117893	0.883643	0.8053633
Contador 17	0.8444933	0.9509836	0.9162201
Contador 18	0.5825413	0.8745808	0.9109364
Contador 19	0.1768492	0.5932041	0.6898423
Contador 20	0.656999	0.9065614	0.9255961
Contador 21	0.07268196	0.5177674	0.6193504
Contador 22	0.8494686	0.9324393	0.8894573

Tabela 4.20: Coeficiente de correlação de Spearman entre os valores das séries de tensão média das três fases do posto de transformação e os valores de tensão média dos contadores

Coeficiente de Correlação de Spearman para as Séries de Tensão

Contador	Fase 1	Fase 2	Fase 3
Contador 23	0.6392789	0.902196	0.920585
Contador 24	0.8435896	0.9342733	0.8865292
Contador 25	0.9234748	0.8719442	0.7864944
Contador 26	0.9275455	0.8728097	0.7954874
Contador 27	0.6767355	0.9052762	0.9318891
Contador 28	0.864909	0.9507866	0.9144667
Contador 29	0.9354229	0.7357393	0.6895225
Contador 30	0.4376659	0.7813216	0.8544522
Contador 31	0.5395244	0.8502396	0.8937871
Contador 32	0.9313241	0.7469827	0.7029419
Contador 33	0.821554	0.9267708	0.8916875
Contador 34	0.9284302	0.7472693	0.7035873
Contador 35	0.07074947	0.5193251	0.6206496
Contador 36	0.5377639	0.8411135	0.8846622

Tabela 4.21: Coeficiente de correlação de Spearman entre os valores das séries de tensão média das três fases do posto de transformação e os valores de tensão média dos contadores

O coeficiente de correlação de Spearman entre as séries das derivadas do Posto de Transformação e dos contadores, para os 36 contadores:

Coeficiente de Correlação de Spearman para as Séries das Derivadas

Contador	Fase 1	Fase 2	Fase 3
Contador 1	0.5753444	0.3735381	0.4851053
Contador 2	0.2105251	0.3896985	0.444091
Contador 3	0.21556	0.3082888	0.4709407
Contador 4	0.428992	0.5316705	0.4142452
Contador 5	0.5529131	0.4710801	0.5053344
Contador 6	0.2245401	0.4057513	0.4555879
Contador 7	0.5724154	0.2697296	0.4227629
Contador 8	0.428125	0.4666858	0.2732781
Contador 9	0.3146826	0.4410548	0.3479305
Contador 10	0.5438097	0.2147425	0.3631805
Contador 11	0.4638469	0.4163991	0.4523589
Contador 12	0.3802189	0.4189313	0.258077
Contador 13	0.1691328	0.3246793	0.5022413
Contador 14	0.3604638	0.4076109	0.2298108
Contador 15	0.4931593	0.4679896	0.4800659
Contador 16	0.3052629	0.4262138	0.2677814
Contador 17	0.4520803	0.4433206	0.4790014
Contador 18	0.19953	0.3430711	0.3680839
Contador 19	0.2321278	0.3306284	0.4872994
Contador 20	0.1739419	0.3321549	0.3971313
Contador 21	0.2116206	0.3645516	0.5138005
Contador 22	0.4055803	0.3255254	0.3720856

Tabela 4.22: Coeficiente de correlação de Spearman entre os valores das séries das derivadas das três fases do posto de transformação e os valores das derivadas dos contadores

Coeficiente de Correlação de Spearman para as Séries das Derivadas

Contador	Fase 1	Fase 2	Fase 3
Contador 23	0.1228281	0.3642686	0.5138005
Contador 24	0.4276932	0.4914965	0.364909
Contador 25	0.3193148	0.4409888	0.2650331
Contador 26	0.4327652	0.4383469	0.3141427
Contador 27	0.2034871	0.2673071	0.4458358
Contador 28	0.4728919	0.4542239	0.459939
Contador 29	0.6212455	0.3189702	0.41157
Contador 30	0.2846678	0.3693719	0.4981453
Contador 31	0.1855808	0.3554326	0.4247813
Contador 32	0.5237599	0.2691844	0.3535555
Contador 33	0.3928718	0.3799698	0.3652225
Contador 34	0.5766694	0.3102918	0.41168
Contador 35	0.2354807	0.3259741	0.4411573
Contador 36	0.2286195	0.3745321	0.4076127

Tabela 4.23: Coeficiente de correlação de Spearman entre os valores das séries das derivadas das três fases do posto de transformação e os valores das derivadas dos contadores

Classificação da rede 1:

Resultados dos Critérios					
Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman	Mínimo do Desvio Padrão das Derivadas	Máximo do Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 2	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 3	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 4	*Fase 2	*Fase 2	*Fase 2	*Fase 2	*Fase 3
Contador 5	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 6	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 7	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 8	*Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 2
Contador 9	Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 2
Contador 10	*Fase 1	*Fase 1	*Fase 1	*Fase 1	*Fase 1
Contador 11	*Fase 2	*Fase 2	Fase 2	*Fase 1	*Fase 3
Contador 12	Fase 1	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 13	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 14	Fase 1	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 15	Fase 2	*Fase 2	Fase 2	*Fase 1	*Fase 1
Contador 16	Fase 1	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 17	*Fase 2	*Fase 2	Fase 3	*Fase 3	*Fase 3

Tabela 4.24: Classificação obtida em cada um dos critérios trabalhados para todos os contadores da rede 1

Resultados dos Critérios					
Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman	Mínimo do Desvio Padrão das Derivadas	Máximo do Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 18	*Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 19	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 20	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 21	Fase 3	*Fase 3	*Fase 3	*Fase 3	*Fase 1
Contador 22	*Fase 2	*Fase 2	Fase 3	*Fase 1	*Fase 2
Contador 23	Fase 3	*Fase 3	*Fase 2	*Fase 3	*Fase 2
Contador 24	*Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 3
Contador 25	Fase 1	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 26	Fase 1	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 27	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 28	*Fase 2	*Fase 2	Fase 2	*Fase 1	*Fase 2
Contador 29	*Fase 1	*Fase 1	*Fase 1	*Fase 1	*Fase 1
Contador 30	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 31	*Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 32	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 33	*Fase 2	*Fase 2	Fase 2	*Fase 1	*Fase 3
Contador 34	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 35	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 36	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1

Tabela 4.25: Classificação obtida em cada um dos critérios trabalhados para todos os contadores da rede 1

Classificação da 1ª metade da rede 1:

Resultados dos Critérios para a 1ª Metade					
Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman	Mínimo do Desvio Padrão das Derivadas	Máximo do Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 2	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 3	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 4	*Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 3
Contador 5	Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 6	*Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 7	Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 8	*Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 2
Contador 9	Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 3
Contador 10	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 11	*Fase 2	*Fase 2	Fase 2	*Fase 1	*Fase 1
Contador 12	*Fase 1	*Fase 1	Fase 2	*Fase 1	*Fase 1
Contador 13	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 14	*Fase 1	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 15	Fase 2	*Fase 2	Fase 2	*Fase 1	*Fase 2
Contador 16	*Fase 1	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 17	*Fase 2	*Fase 2	Fase 2	*Fase 3	*Fase 1

Tabela 4.26: Classificação obtida em cada um dos critérios trabalhados para todos os contadores da 1ª metade da rede 1

Resultados dos Critérios					
Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman	Mínimo do Desvio Padrão das Derivadas	Máximo do Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 18	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 19	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 20	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 21	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 22	*Fase 2	*Fase 2	Fase 1	*Fase 1	*Fase 1
Contador 23	Fase 3	*Fase 3	Fase 2	*Fase 2	*Fase 2
Contador 24	*Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 1
Contador 25	*Fase 1	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 26	*Fase 1	*Fase 1	Fase 2	*Fase 1	*Fase 1
Contador 27	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 28	*Fase 2	*Fase 2	Fase 2	*Fase 3	*Fase 1
Contador 29	*Fase 1	*Fase 1	*Fase 1	*Fase 1	*Fase 1
Contador 30	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 31	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 32	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 33	Fase 2	*Fase 2	Fase 1	*Fase 1	*Fase 1
Contador 34	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 35	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2
Contador 36	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 2

Tabela 4.27: Classificação obtida em cada um dos critérios trabalhados para todos os contadores da 1ª metade da rede 1

Classificação da 2ª metade da rede 1:

Resultados dos Critérios para a 2ª Metade					
Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman	Mínimo do Desvio Padrão das Derivadas	Máximo do Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 2	Fase 3	*Fase 3	Fase 3	*Fase 2	*Fase 3
Contador 3	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 4	*Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 3
Contador 5	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 6	Fase 3	*Fase 3	Fase 3	*Fase 2	*Fase 1
Contador 7	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 8	Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 1
Contador 9	Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 2
Contador 10	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 11	Fase 2	*Fase 2	Fase 3	*Fase 3	*Fase 1
Contador 12	Fase 2	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 13	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 3
Contador 14	Fase 2	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 15	Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 2
Contador 16	Fase 2	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 17	*Fase 2	*Fase 2	Fase 3	*Fase 3	*Fase 1

Tabela 4.28: Classificação obtida em cada um dos critérios trabalhados para todos os contadores da 2ª metade da rede 1

Resultados dos Critérios					
Contador	Mínimo do Desvio Padrão	Máximo do Coeficiente de Correlação de Spearman	Mínimo do Desvio Padrão das Derivadas	Máximo do Coeficiente de Correlação de Spearman das Derivadas	Modelo ARIMA
Contador 18	*Fase 3	*Fase 3	Fase 2	*Fase 2	*Fase 1
Contador 19	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 3
Contador 20	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 3
Contador 21	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 3
Contador 22	*Fase 2	*Fase 2	Fase 3	*Fase 3	*Fase 1
Contador 23	Fase 3	*Fase 3	Fase 2	*Fase 2	*Fase 2
Contador 24	*Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 1
Contador 25	Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 1
Contador 26	Fase 2	*Fase 1	Fase 2	*Fase 2	*Fase 1
Contador 27	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 28	*Fase 2	*Fase 2	Fase 2	*Fase 1	*Fase 1
Contador 29	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 30	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 3
Contador 31	Fase 3	*Fase 3	Fase 3	*Fase 2	*Fase 1
Contador 32	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 3
Contador 33	*Fase 2	*Fase 2	Fase 2	*Fase 2	*Fase 1
Contador 34	*Fase 1	*Fase 1	Fase 1	*Fase 1	*Fase 1
Contador 35	Fase 3	*Fase 3	Fase 3	*Fase 3	*Fase 1
Contador 36	Fase 3	*Fase 3	Fase 2	*Fase 2	*Fase 3

Tabela 4.29: Classificação obtida em cada um dos critérios trabalhados para todos os contadores da 2ª metade da rede 1

Sistema de Votação			
Contador	Nº de votos Rede 1	Nº de Votos 1ª Metade	Nº de Votos 2ª Metade
Contador 2	Fase 1: 0	Fase 1: 0	Fase 1: 0
	Fase 2: 1	Fase 2: 1	Fase 2: 1
	Fase 3: 4	Fase 3: 4	Fase 3: 4
Contador 3	Fase 1: 1	Fase 1: 1	Fase 1: 1
	Fase 2: 0	Fase 2: 0	Fase 2: 0
	Fase 3: 4	Fase 3: 4	Fase 3: 4
Contador 4	Fase 1: 0	Fase 1: 0	Fase 1: 0
	Fase 2: 4	Fase 2: 4	Fase 2: 0
	Fase 3: 1	Fase 3: 1	Fase 3: 5
Contador 6	Fase 1: 0	Fase 1: 0	Fase 1: 1
	Fase 2: 1	Fase 2: 1	Fase 2: 1
	Fase 3: 4	Fase 3: 4	Fase 3: 3
Contador 8	Fase 1: 0	Fase 1: 0	Fase 1: 1
	Fase 2: 5	Fase 2: 5	Fase 2: 4
	Fase 3: 0	Fase 3: 0	Fase 3: 0
Contador 9	Fase 1: 0	Fase 1: 0	Fase 1: 0
	Fase 2: 5	Fase 2: 4	Fase 2: 5
	Fase 3: 0	Fase 3: 1	Fase 3: 0
Contador 11	Fase 1: 1	Fase 1: 2	Fase 1: 1
	Fase 2: 3	Fase 2: 3	Fase 2: 2
	Fase 3: 1	Fase 3: 0	Fase 3: 2
Contador 12	Fase 1: 3	Fase 1: 4	Fase 1: 2
	Fase 2: 2	Fase 2: 1	Fase 2: 3
	Fase 3: 0	Fase 3: 0	Fase 3: 0
Contador 13	Fase 1: 1	Fase 1: 0	Fase 1: 0
	Fase 2: 0	Fase 2: 1	Fase 2: 0
	Fase 3: 4	Fase 3: 4	Fase 3: 5

Tabela 4.30: Classificação obtida do sistema de votação para a rede 1 e ambas as metades

Sistema de Votação			
Contador	Nº de votos Rede 1	Nº de Votos 1ª Metade	Nº de Votos 2ª Metade
Contador 14	Fase 1: 3	Fase 1: 3	Fase 1: 2
	Fase 2: 2	Fase 2: 2	Fase 2: 3
	Fase 3: 0	Fase 3: 0	Fase 3: 0
Contador 15	Fase 1: 2	Fase 1: 1	Fase 1: 0
	Fase 2: 3	Fase 2: 4	Fase 2: 5
	Fase 3: 0	Fase 3: 0	Fase 3: 0
Contador 16	Fase 1: 3	Fase 1: 3	Fase 1: 2
	Fase 2: 2	Fase 2: 2	Fase 2: 3
	Fase 3: 0	Fase 3: 0	Fase 3: 0
Contador 18	Fase 1: 1	Fase 1: 1	Fase 1: 1
	Fase 2: 0	Fase 2: 0	Fase 2: 2
	Fase 3: 4	Fase 3: 4	Fase 3: 2
Contador 19	Fase 1: 1	Fase 1: 0	Fase 1: 0
	Fase 2: 0	Fase 2: 1	Fase 2: 0
	Fase 3: 4	Fase 3: 4	Fase 3: 5
Contador 20	Fase 1: 0	Fase 1: 1	Fase 1: 0
	Fase 2: 1	Fase 2: 0	Fase 2: 0
	Fase 3: 4	Fase 3: 3	Fase 3: 5
Contador 21	Fase 1: 1	Fase 1: 1	Fase 1: 0
	Fase 2: 0	Fase 2: 0	Fase 2: 0
	Fase 3: 4	Fase 3: 4	Fase 3: 5
Contador 22	Fase 1: 1	Fase 1: 3	Fase 1: 1
	Fase 2: 3	Fase 2: 2	Fase 2: 2
	Fase 3: 1	Fase 3: 0	Fase 3: 2
Contador 23	Fase 1: 0	Fase 1: 0	Fase 1: 0
	Fase 2: 2	Fase 2: 3	Fase 2: 2
	Fase 3: 3	Fase 3: 2	Fase 3: 2
Contador 24	Fase 1: 0	Fase 1: 1	Fase 1: 1
	Fase 2: 4	Fase 2: 4	Fase 2: 4
	Fase 3: 1	Fase 3: 0	Fase 3: 0
Contador 25	Fase 1: 3	Fase 1: 3	Fase 1: 1
	Fase 2: 2	Fase 2: 2	Fase 2: 4
	Fase 3: 0	Fase 3: 0	Fase 3: 0

Tabela 4.31: Classificação obtida do sistema de votação para a rede 1

Sistema de Votação			
Contador	Nº de votos Rede 1	Nº de Votos 1ª Metade	Nº de Votos 2ª Metade
Contador 26	Fase 1: 3	Fase 1: 4	Fase 1: 2
	Fase 2: 2	Fase 2: 1	Fase 2: 3
	Fase 3: 0	Fase 3: 0	Fase 3: 0
Contador 27	Fase 1: 0	Fase 1: 1	Fase 1: 1
	Fase 2: 1	Fase 2: 0	Fase 2: 0
	Fase 3: 4	Fase 3: 4	Fase 3: 4
Contador 28	Fase 1: 1	Fase 1: 1	Fase 1: 2
	Fase 2: 4	Fase 2: 3	Fase 2: 2
	Fase 3: 0	Fase 3: 1	Fase 3: 0
Contador 30	Fase 1: 0	Fase 1: 1	Fase 1: 0
	Fase 2: 1	Fase 2: 0	Fase 2: 0
	Fase 3: 4	Fase 3: 4	Fase 3: 5
Contador 31	Fase 1: 0	Fase 1: 1	Fase 1: 1
	Fase 2: 1	Fase 2: 0	Fase 2: 1
	Fase 3: 4	Fase 3: 4	Fase 3: 3
Contador 32	Fase 1: 5	Fase 1: 5	Fase 1:
	Fase 2: 0	Fase 2: 0	Fase 2: 0
	Fase 3: 0	Fase 3: 0	Fase 3: 1
Contador 33	Fase 1: 1	Fase 1: 3	Fase 1: 1
	Fase 2: 3	Fase 2: 2	Fase 2: 4
	Fase 3: 1	Fase 3: 0	Fase 3: 0
Contador 34	Fase 1: 5	Fase 1: 5	Fase 1: 5
	Fase 2: 0	Fase 2: 0	Fase 2: 0
	Fase 3: 0	Fase 3: 0	Fase 3: 0
Contador 35	Fase 1: 0	Fase 1: 0	Fase 1: 1
	Fase 2: 1	Fase 2: 1	Fase 2: 0
	Fase 3: 4	Fase 3: 4	Fase 3: 4
Contador 36	Fase 1: 1	Fase 1: 0	Fase 1: 0
	Fase 2: 0	Fase 2: 0	Fase 2: 2
	Fase 3: 4	Fase 3: 4	Fase 3: 3

Tabela 4.32: Classificação obtida do sistema de votação para a rede 1

Anexo E

Comandos usados em R:

lm{stats}: esta função permite a construção de modelos lineares que se adequem aos dados.

Box.test{stats}: esta função calcula o teste de Ljung-Box ou de Box-Pierce para testar a hipótese nula de que corresponde a que as primeiras L autocorrelações sejam estatisticamente nulas numa dada série.

shapiro.test{stats}: teste para a normalidade da amostra fornecida.

var{stats}: esta função calcula a variância de um vetor, matriz ou *data frame*.

sd{stats}: esta função calcula o desvio padrão dos valores do vetor fornecido.

levene.test{car}: esta função efetua um teste para verificar a igualdade de k variâncias. É calculado o teste clássico de Levene se se usar a média, o teste de Brown-Forsythe se se usar a mediana e ainda outra variante se se usar a média truncada.

cor.test{stats}: este teste permite calcular a correlação entre dois vetores amostrais usando o coeficiente de correlação de Pearson, ou o coeficiente de correlação de Kendall ou o coeficiente de correlação de Spearman. Fornece também o valor do teste de hipótese cuja hipótese nula é o coeficiente igual a zero.