

4 Análise dos tributos das concessionárias selecionadas

Neste capítulo serão abordados os subsídios teóricos dos modelos estatísticos através da análise das séries temporais correspondentes aos tributos e encargos incidentes sobre a tarifa de energia elétrica. Esses tributos representam uma parcela importante da conta de luz (aproximadamente 33.45%). Um estudo de caso foi realizado com uma amostra de 6 (seis) concessionárias brasileiras de energia elétrica.

4.1. Séries temporais

Para fazer a análise das séries dos tributos das concessionárias, será visto a parte conceitual dedicada ao estudo de séries temporais e sua aplicabilidade na previsão de valores futuros dessas séries. São introduzidos os conceitos de série temporal e previsão, além da classificação e descrição de alguns métodos de previsão. Também são feitas determinadas considerações a respeito desses métodos e sobre as medidas de desempenho usualmente adotadas na avaliação de tais procedimentos.

4.1.1. Definição

Uma definição clara de (Souza, 1989) afirma que a classe de fenômenos cujo processo observacional e conseqüente quantificação numérica geram uma seqüência de dados distribuídos no tempo é denominada série temporal.

Segundo (Anderson, 1971), a natureza de uma série temporal e a estrutura de seu mecanismo gerador estão relacionados com o intervalo de ocorrência das observações no tempo³⁰. Caso o levantamento das observações da série possa ser feito a qualquer momento do tempo, a série temporal é dita contínua, sendo

³⁰ ANDERSON, T. W., 1971. *The Statistical Analysis of Time Series*. New York: John Wiley & Sons

denotada por $x(t)$. Entretanto, de acordo com (Nelson, 1973; Granger e Newbold, 1977), na maioria das séries, as observações são tomadas em intervalos de tempo discretos e eqüidistantes.

Por sua parte (Souza, 1989) afirma que uma série temporal *discreta* pode ser representada por $X^T = \{x_1, x_2, \dots, x_T\}$, sendo que cada observação discreta x_t está associada a um instante de tempo distinto, existindo uma relação de dependência serial entre essas observações.

(Granger e Newbold, 1977) coloca como objetivo inicial da análise de séries temporais a realização de inferências sobre as características básicas do mecanismo gerador do processo estocástico das observações da série. Assim, com forme (Barbancho, 1970), através da abstração de regularidades contidas nos fenômenos observáveis de uma série temporal existe a possibilidade de se construir um modelo matemático como uma representação simplificada da realidade.

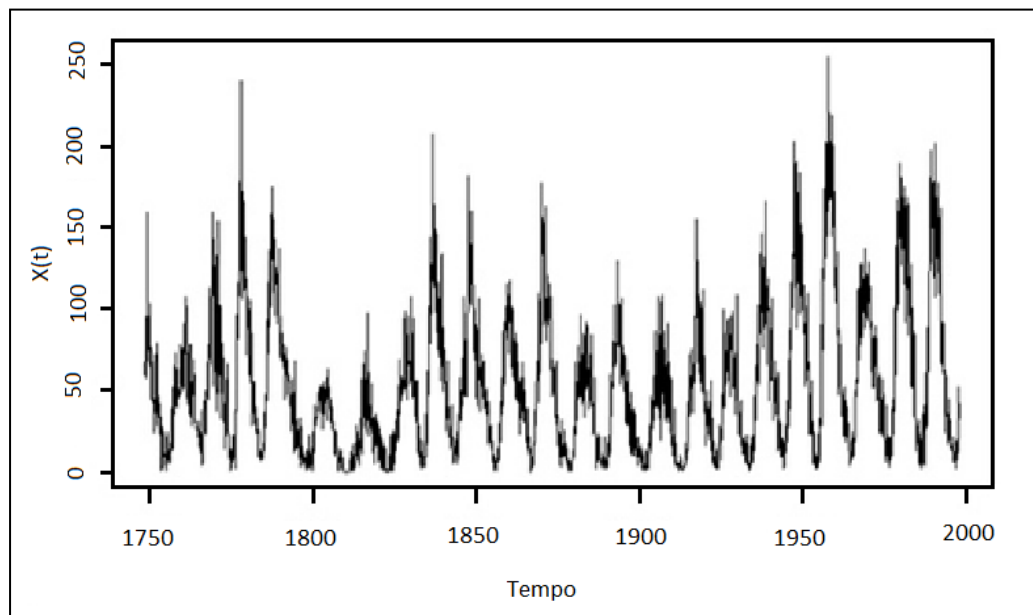


Figura 16 - Série temporal com dados anuais.

Em quanto a sua aplicação segundo (Nelson, 1973), após a formulação do modelo matemático, obtido pela seleção entre as alternativas de classes de modelos identificadas como apropriadas para essa representação e subsequente estimação de seus parâmetros, é possível utilizá-lo para testar alguma hipótese ou

teoria a respeito do mecanismo gerador do processo estocástico e realizar a previsão de valores futuros da série temporal.

4.1.2. Componentes de uma série temporal

Uma série temporal é constituída por um conjunto de componentes não-observáveis (Souza, 1989). Dessa forma, pela identificação das componentes individuais presentes no padrão básico da série histórica de dados (tendência, ciclo, sazonalidade e aleatoriedade), a extrapolação para o futuro pode ser realizada (Oconnor e Lawrence, 1992).

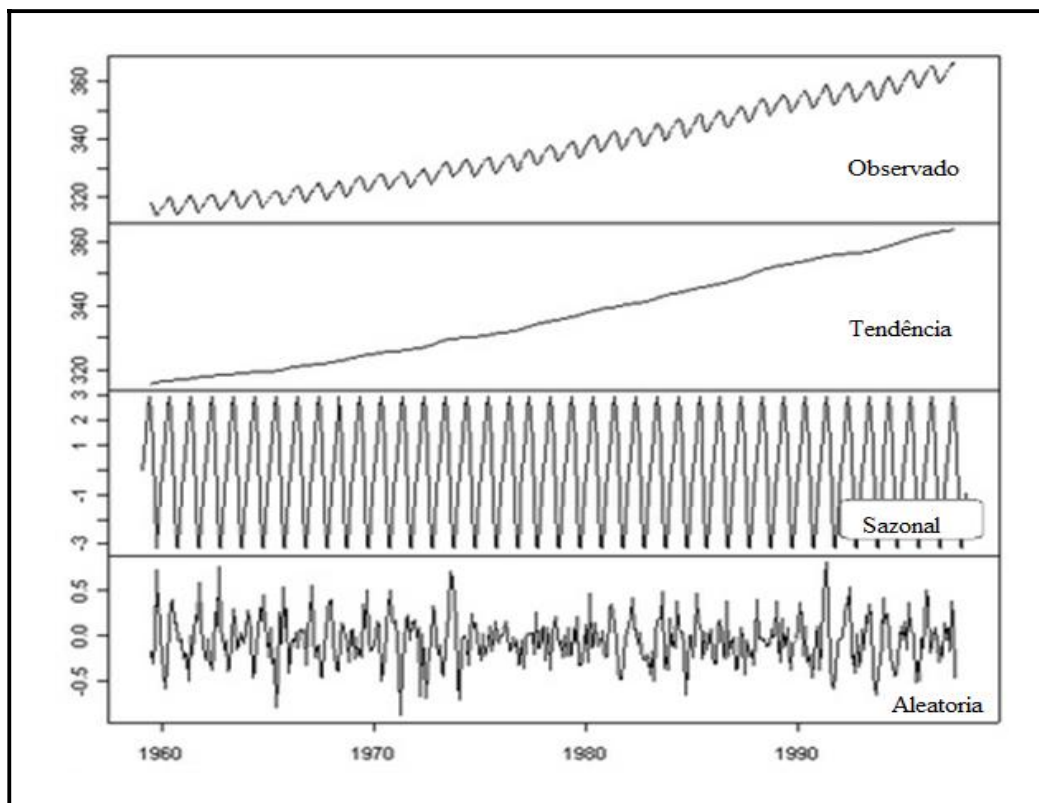


Figura 17 - Componentes gráficos de uma série temporal

(Casals, Jerez *et al.*, 2009) expressa o relacionamento entre as componentes não-observáveis da série temporal através da equação (4.1).

$$x_t = f(S_t, T_t, C_t, E_t) \quad (4.1)$$

Onde S_t corresponde à componente sazonal para o período t ; T_t é a componente de tendência no período t ; C_t é a componente de ciclo no período t e E_t é a componente aleatória no período t .

Para (Morettin e Tolo, 2006), a componente sazonal representa as flutuações da série de acordo com algum fator de sazonalidade. O ciclo apresenta um comportamento similar à componente sazonal, embora tenha normalmente comprimento maior que aquela. Justamente pelo fato de não apresentar duração uniforme, a identificação da componente ciclo é mais problemática. A tendência representa o aumento ou declínio gradual nos valores das observações de uma série temporal. Com a remoção das componentes de sazonalidade, ciclo e tendência, a componente aleatória fica determinada (Oconnor e Lawrence, 1992).

(Marcek, 2001) coloca que vários procedimentos para a decomposição de séries temporais foram desenvolvidos, cada qual tentando isolar as componentes não-observáveis da série o mais acuradamente possível. O objetivo desses procedimentos consiste em remover cada uma das componentes, permitindo que o comportamento da série temporal seja melhor compreendido e, conseqüentemente, prognosticar valores futuros mais apropriados.

4.1.3. Estacionaridade de uma série

A maioria das séries temporais têm uma tendência. Seus valores médios tem uma tendência que variam ao longo do tempo. Eles são séries ou variáveis não – estacionárias.

(Martins, 2009; Butler e Bloom, 2011), deixam a entrever, quando as séries não têm um curso definido, se conhecem como andar aleatório³¹ (random walk) é também uma propriedade de muitas variáveis não estacionárias.

³¹ Segundo Damodar Gujarati (p. 702) em Econometria, uma serie temporal que têm uma raiz unitária se conhece como passeio aleatória. Uma passeio aleatória é um exemplo de uma serie não-estacionaria.

4.1.4. Teste de estacionariedade

4.1.4.1. Testes informais: análise de dados

A estacionariedade de uma variável pode ser examinada observando a sua representação gráfica. Na figura 18, as séries das alíquotas do PIS/COFINS são traçadas para cada uma das seis distribuidoras (LIGHT, ELETROPAULO, CEMIG, AMPLA, ELEKTRO e CPFL) que se amostram no Anexo 3, os que antes de fazer algum teste analítico, podem-se observar quais deles têm estacionariedade. Neste caso pode-se observar que são séries que aparentemente possuem estacionariedade, com exceção da CEMIG e ELETROPAULO, que mais na frente poderão-se fazer os testes analíticos



Figura 18 - Dados originais das alíquotas PIS/COFINS de seis concessionárias

4.1.4.2. Testes Formais: O teste de Dickey-Fuller

Na prática o teste mais utilizado na identificação de não-estacionariedade é o de Dickey-Fuller (Augmented Dickey-Fuller Unit Root Test), que consiste na verificação da existência ou não de uma raiz unitária na série, por meio de um teste de hipóteses. A hipótese nula considerada é a de que existe uma raiz unitária. Portanto, é necessário que a hipótese nula seja rejeitada para que a série seja considerada estacionária (Aneel, 2008).

O teste começa por supor que a série X_t segue um processo autoregressivo de primeira ordem $AR(1)$, da forma:

$$X_t = rX_{t-1} + e_t \quad (4.2)$$

E se testa para o caso de que r seja igual a 1: ($r = 1$), (*unity*, daqui a expressão raiz unitária). (Dickey e Fuller, 1979) consideram três diferentes equações de regressão que se podem utilizar para testar a presença de uma raiz unitária:

$$X_t = rX_{t-1} + e_t \quad (4.3)$$

$$X_t = a_0 + rX_{t-1} + e_t \quad (4.4)$$

$$X_t = a_0 + rX_{t-1} + a_1 + e_t \quad (4.5)$$

A diferença entre as três regressões se deve a presença de elementos determinísticos a_0 e a_1 . O primeiro é um modelo de andar aleatório, o segundo adiciona um intercepto e o terceiro incluem ambos, ou seja, um intercepto e uma linha de tendência.

O parâmetro de interesse em todas as equações de regressão é r : se $r = 1$, a série contém uma raiz unitária. Uma versão simples do teste consiste em estimar uma (ou más) das equações antes indicadas mediante o método dos Mínimos Quadrados Ordinários (MCO) com o fim de obter o valor estimado de r e seu correspondente erro padrão associado. Comparando o estatístico resultante (t^* tau) com o valor apropriado reportado nas tabelas Dickey-Fuller, o pesquisador pode determinar se aceita ou não a **hipótese nula $r = 1$** . Porém há um problema. Embora possa parecer obvio que poda-se testar isso usando o estatístico $(\hat{r} - 1)/s_r$, onde \hat{r} é a estimativa de r e s_r é uma estimativa do seu desvio padrão. Mas há um problema. Neste teste a hipótese nula é $H_0 : r = 1$, onde se diz

que X têm uma raiz unitária. A hipótese alternativa é $H_1 : r < 1$. Se a hipótese alternativa é correta, então X é estacionária. Caso contrário, se a hipótese nula é correta, então a variável é não estacionária.

Suponha que reescrevemos o modelo de X_t da maneira seguinte:

$$DX_t = a_0 + gX_{t-1} + e_t \quad (4.6)$$

Onde $g = (r-1)$. Neste caso a hipótese nula é $H_0 : g = 0$, em cujo caso $r = 1$ e a variável é não estacionária. A hipótese alternativa é $H_1 : g < 0$ significando que $r < 1$ e que X é estacionária. Assim, tem-se a seguinte estatística:

$$\hat{g} / s_g \quad (4.7)$$

Onde \hat{g} é estimativa de g na regressão (4.6) e s_g é uma estimativa de seu desvio padrão. Isso é fácil de calcular, pois é justamente o estatístico t de g na regressão. Rejeita-se a hipótese nula e concluímos que X é estacionária se a relação de t no g é suficientemente negativa. O problema é que a estatística não segue a distribuição do t quando seja não estacionário se a hipótese nula é correta. (Dickey e Fuller, 1979) resolveram este problema gerando números aleatórios (mediante um experimento de Monte Carlo) dando lugar a valores críticos corretos. Os resultados obtidos do teste Dickey-Fuller são amostrados na seguinte tabela:

Tabela 6 - Testes de Raízes Unitárias (ADF)

Variável	Característica	Razão t	Valor crítico		
			1%	5%	10%
Light	Intercepto	-4.193(*)	-3.670	-2.964	-2.621
Light	Intercepto e tendência.	-4.237(**)	-4.297	-3.568	-3.218
Eletropaulo	Intercepto	-1.392	-3.670	-2.964	-2.621
Eletropaulo	Intercepto e tendência.	-2.989	-4.297	-3.568	-3.218
DEletropaulo	Intercepto	-5.842(*)	-3.679	-2.968	-2.623
DEletropaulo	Intercepto e tendência.	-5.7374(*)	-4.310	-3.574	-3.222
Cemig	Intercepto	-1.571	-3.788	-3.012	-2.646
Cemig	Intercepto e tendência.	-2.704	-4.468	-3.645	-3.261
DCemig	Intercepto	-5.827(*)	-3.808	-3.021	-2.650
DCemig	Intercepto e tendência.	-5.627(*)	-4.498	-3.658	-3.269
Ampla	Intercepto	-5.252(*)	-3.670	-2.964	-2.621
Ampla	Intercepto e tendência.	-5.278(*)	-4.297	-3.568	-3.218
Elektro	Intercepto	-3.401(**)	-3.670	-2.964	-2.621
Elektro	Intercepto e tendência.	-3.278(***)	-4.297	-3.568	-3.218
Cpfl	Intercepto	-3.756(*)	-3.670	-2.964	-2.621
Cpfl	Intercepto e tendência.	-3.693(**)	-4.297	-3.568	-3.218

(*) Indica rejeição da hipótese nula de não-estacionariedade a 1% de significância.

(**) Indica rejeição da hipótese nula de não-estacionariedade a 5% de significância.

(***) Indica rejeição da hipótese nula de não-estacionariedade a 10% de significância.

Considerando as séries testadas com os dados originais, foram encontradas estacionárias somente as que correspondem às concessionárias LIGHT, AMPLA, ELEKTRO e CPFL. Enquanto que, para as outras séries (ELETROPAULO e CEMIG) foi necessário fazer o teste com a primeira diferença para as tornarem estacionárias. No caso da CEMIG, isso ocorreu por causa de ter pouca informação, pois os dados se iniciam no ano de 2005.

Quanto à sazonalidade, podemos dizer que por se tratar de dados trimestrais, se espera que a cada quatro períodos características sejam semelhantes. A função de autocorrelograma pode ajudar a ter uma visão mais clara, ao mesmo tempo em que, se pode analisar a autocorrelação possível nos dados da série. As tabelas 7 – 12, pode-se verificar tal comportamento.

4.1.5. Autocorrelação

Existem diversos métodos formais e gráficos para identificar a existência ou não de autocorrelação das séries.

4.1.5.1. Função de autocorrelação

(Souza e Camargo, 2004), define a função de autocorrelação como a autocovariância padronizada; serve para medirmos o comprimento e a memória de um processo, ou seja, a extensão para qual o valor tomado no tempo t depende daquele tomado no tempo $t-k$. A auto correlação de defasagem k é definida como:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{Cov}[Z_t, Z_{t+k}]}{\sqrt{\text{Var}(Z_t)\text{Var}(Z_{t+k})}} \quad (4.8)$$

$\text{Var}(Z_t) = \text{Var}(Z_{t+k}) = \gamma_0$ é variância do processo, $\rho_0 = 1$ e $\rho_k = \rho_{-k}$

Em outras palavras, a função de autocorrelação é simétrica sobre a origem $k = 0$

O gráfico dos coeficientes de autocorrelação ρ_k versus k é chamado de autocorrelograma.

A seguir os resultados da função de autocorrelação das séries das alíquotas PIS e COFINS das distribuidoras que formão parte deste estudo.

Tabela 7 - Autocorrelações das alíquotas da LIGHT

Autocorrelação	Correlação Parcial	AC	PAC	Q-Stat	Prob	
		1	0.198	0.198	1.3387	0.247
		2	-0.205	-0.254	2.8204	0.244
		3	-0.082	0.021	3.0665	0.381
		4	0.170	0.148	4.1582	0.385
		5	0.013	-0.091	4.1651	0.526
		6	0.101	0.212	4.5791	0.599
		7	0.234	0.191	6.9049	0.439
		8	-0.012	-0.120	6.9113	0.546
		9	-0.300	-0.183	11.096	0.269
		10	-0.070	0.012	11.338	0.332
		11	0.184	0.031	13.073	0.289
		12	-0.100	-0.235	13.606	0.327
		13	-0.206	-0.069	16.026	0.248
		14	-0.167	-0.222	17.704	0.221
		15	-0.077	-0.084	18.081	0.258
		16	-0.116	0.003	18.998	0.269

Tabela 8 - Autocorrelações das alíquotas da CEMIG

Autocorrelação	Correlação Parcial	AC	PAC	Q-Stat	Prob	
		1	0.716	0.716	12.880	0.000
		2	0.638	0.258	23.624	0.000
		3	0.488	-0.077	30.248	0.000
		4	0.380	-0.054	34.481	0.000
		5	0.217	-0.167	35.949	0.000
		6	0.112	-0.065	36.363	0.000
		7	-0.014	-0.091	36.370	0.000
		8	-0.123	-0.107	36.938	0.000
		9	-0.152	0.075	37.875	0.000
		10	-0.277	-0.198	41.242	0.000
		11	-0.323	-0.084	46.258	0.000
		12	-0.371	-0.046	53.512	0.000

Tabela 9 - Autocorrelações das alíquotas da ELEKTRO

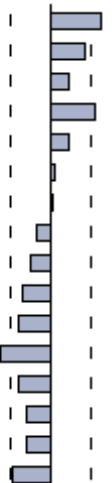

Autocorrelação	Correlação Parcial	AC	PAC	Q-Stat	Prob	
		1	0.455	0.455	7.0509	0.008
		2	0.309	0.129	10.429	0.005
		3	0.168	-0.017	11.462	0.009
		4	0.387	0.367	17.148	0.002
		5	0.156	-0.194	18.109	0.003
		6	0.039	-0.125	18.173	0.006
		7	0.013	0.100	18.180	0.011
		8	-0.131	-0.409	18.939	0.015
		9	-0.181	-0.005	20.470	0.015
		10	-0.262	-0.045	23.811	0.008
		11	-0.287	-0.365	28.027	0.003
		12	-0.443	-0.058	38.606	0.000
		13	-0.285	0.194	43.233	0.000
		14	-0.209	-0.135	45.858	0.000
		15	-0.220	0.031	48.953	0.000
		16	-0.340	-0.002	56.818	0.000

Tabela 10 - Autocorrelações das alíquotas da CPFL



Autocorrelação	Correlação Parcial	AC	PAC	Q-Stat	Prob	
		1	0.329	0.329	3.6948	0.055
		2	0.076	-0.036	3.9002	0.142
		3	0.206	0.215	5.4492	0.142
		4	0.347	0.250	10.009	0.040
		5	0.053	-0.157	10.117	0.072
		6	-0.066	-0.073	10.296	0.113
		7	-0.109	-0.200	10.805	0.147
		8	-0.200	-0.262	12.580	0.127
		9	-0.292	-0.184	16.556	0.056
		10	-0.216	-0.041	18.818	0.043
		11	-0.229	-0.049	21.492	0.029
		12	-0.273	-0.008	25.494	0.013
		13	-0.197	0.068	27.694	0.010
		14	-0.162	-0.075	29.264	0.010
		15	-0.176	-0.106	31.238	0.008
		16	-0.197	-0.212	33.887	0.006

Tabela 11 - Autocorrelações das alíquotas da ELETROPAULO

Autocorrelação	Correlação Parcial	AC	PAC	Q-Stat	Prob	
		1	0.808	0.808	22.265	0.000
		2	0.666	0.039	37.927	0.000
		3	0.611	0.180	51.574	0.000
		4	0.515	-0.101	61.612	0.000
		5	0.484	0.170	70.836	0.000
		6	0.463	0.008	79.593	0.000
		7	0.296	-0.360	83.318	0.000
		8	0.193	0.006	84.969	0.000
		9	0.100	-0.176	85.435	0.000
		10	-0.034	-0.136	85.491	0.000
		11	-0.070	0.054	85.741	0.000
		12	-0.115	-0.082	86.459	0.000
		13	-0.195	0.058	88.621	0.000
		14	-0.295	-0.300	93.851	0.000
		15	-0.429	-0.196	105.62	0.000
		16	-0.453	0.239	119.59	0.000

Tabela 12 - Autocorrelações das alíquotas da AMPLA

Autocorrelação	Correlação Parcial	AC	PAC	Q-Stat	Prob	
		1	0.051	0.051	0.0877	0.767
		2	-0.094	-0.097	0.3974	0.820
		3	0.063	0.074	0.5419	0.910
		4	0.206	0.192	2.1504	0.708
		5	-0.073	-0.087	2.3602	0.797
		6	-0.069	-0.032	2.5572	0.862
		7	0.124	0.099	3.2146	0.864
		8	-0.107	-0.171	3.7239	0.881
		9	-0.005	0.075	3.7249	0.929
		10	-0.010	-0.032	3.7293	0.959
		11	-0.046	-0.093	3.8375	0.974
		12	-0.229	-0.166	6.6624	0.879
		13	0.026	0.033	6.7012	0.917
		14	-0.034	-0.096	6.7697	0.943
		15	-0.059	0.034	6.9908	0.958
		16	-0.182	-0.169	9.2556	0.903

As tabelas acima mostram que as séries estudadas não apresentam autocorrelação significativa, mesmo nos casos da LIGHT e da AMPLA que apresentam características de ruído branco.

4.1.5.2. Teste de Durbin-Watson

Dos métodos formais para verificar a presença de autocorrelação, o mais utilizado é o critério de Durbin-Watson, para tal fim é definido o estatístico (Bhargava, Franzini *et al.*, 1982)

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=2}^n e_t^2} \quad (4.9)$$

Nesta equação, e_t é o resíduo estimado para o período t . É possível escrever a d como:

$$d = \frac{\sum_{t=2}^n e_t^2 + \sum_{t=2}^n e_{t-1}^2 - 2 \sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_t^2} = 2(1 - \rho) \quad (4.10)$$

Dado que $\sum e_t^2$ e $\sum e_{t-1}^2$ são aproximadamente iguais para amostras grandes (e dado também que $\rho = \sum e_t e_{t-1} / \sum e_t^2$).

Isso significa que ρ e d seguem uma relação inversa como no gráfico seguinte, ou seja, quando $d \rightarrow 2$, $\rho \rightarrow 0$.

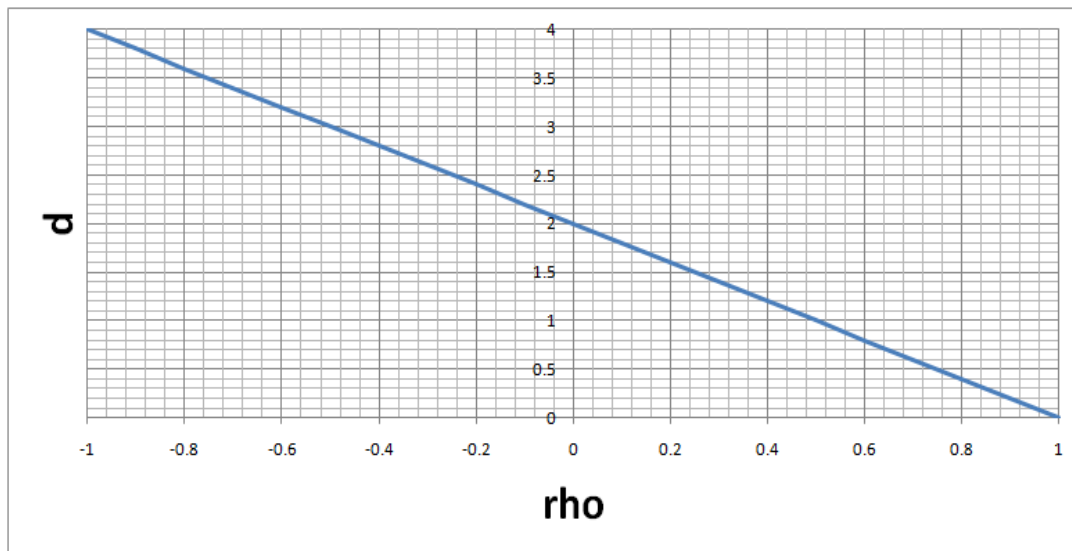


Figura 19 - Relação do estatístico Durbin-Watson e Rho

Os testes foram realizados utilizando o software Eviews, no qual a estatística d de todas as séries se aproxima do valor 2 sinalizando que a autocorrelação é muito próxima de “0”.

Tabela 13 - Teste (estatístico) de Durbin-Watson

Distribuidora	Durbin-Watson stat
LIGHT	1.696075
CEMIG	2.088297
ELEKTRO	2.242832
CPFL	1.958659
ELETROPAULO	2.063671
AMPLA	1.928446

Em relação à sazonalidade, verifica-se que as séries em estudo não apresentam essa característica já que os testes das correlações revelam características de processo de ruído branco. Então, para confirmar a ausência de sazonalidade, utiliza-se o método informal do análise gráfico (Figura 20) no qual se observa para cada trimestre que os dados estão muito dispersos em relação à média.

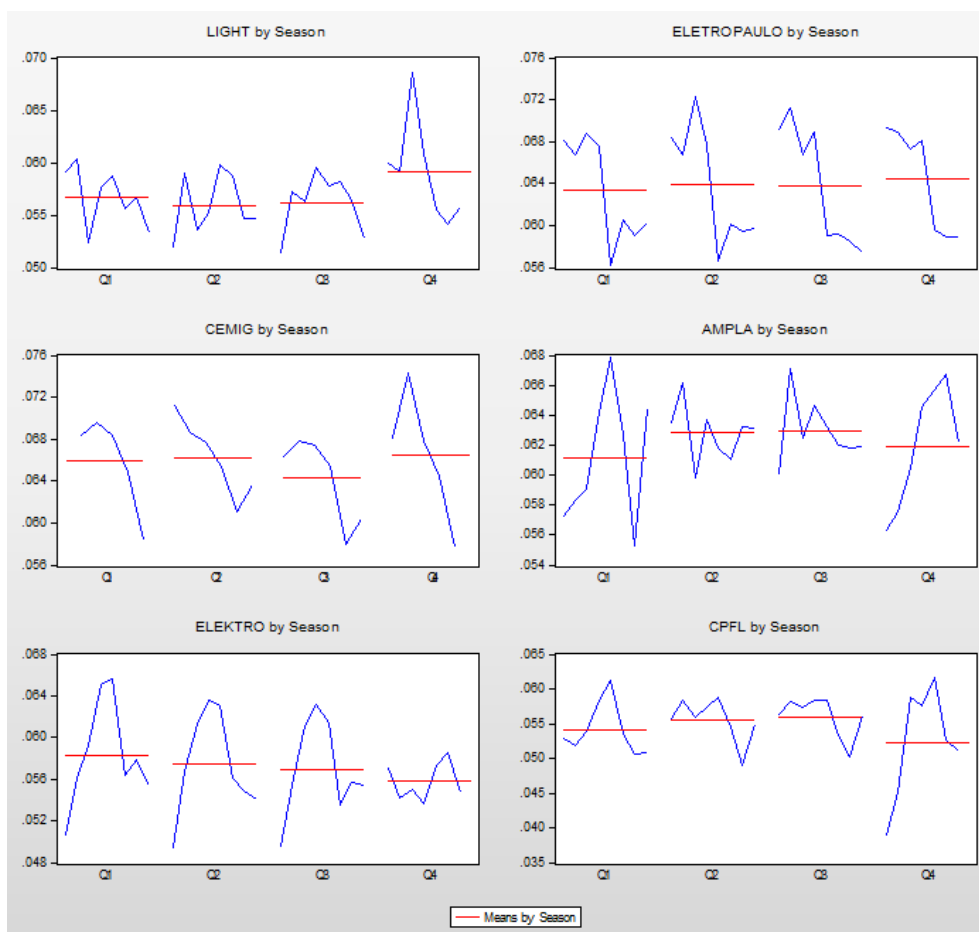


Figura 20 - Análise gráfica da sazonalidade das séries