



ConBRepro

X CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO



02 a 04
de dezembro 2020

Análise Estatística do Histórico de Valores do Preço da Liquidação das Diferenças (PLD) no Mercado Livre de Energia

Priscila Ramallo

Mestranda PPGTCA - Universidade Tecnológica Federal do Paraná

Eduarda A. Antonioli

Mestranda PPGTCA - Universidade Tecnológica Federal do Paraná

José Airton A. dos Santos

PPGTCA - Universidade Tecnológica Federal do Paraná

Carla Adriana Pizarro Schmidt

PPGTCA - Universidade Tecnológica Federal do Paraná

Resumo: Neste trabalho foram utilizados métodos estatísticos para identificar padrões e comportamentos que podem ser observados no histórico de valores médios do PLD divulgado mensalmente pela CCEE (Câmara de Comercialização de Energia Elétrica). Apesar de haver uma correlação alta entre os dados, foi possível identificar a formação de grupos, principalmente relacionadas ao ano.

Palavras-chave: PLD; mercado livre de energia; análise de variância; análise de componente principal; agrupamento.

Settlement Energy historic Price (PLD) Statistical Analyses in The Free Energy Market

Abstract: In this work, statistical methods were used to identify patterns and behaviors that can be observed in the history of average PLD values released monthly by CCEE. Although there is a high correlation between the data, it is possible to identify the formation of groups, mainly related by the year.

Keywords: PLD, free energy market, variance analysis, principal component analysis, grouping.

1. Introdução

A Lei nº 9074 de 1995 dá início ao mercado de comercialização de energia elétrica com a criação do Consumidor Livre de Energia, o qual, atendendo certos requisitos e condições, pode escolher seu fornecedor de energia elétrica (FLOREZI, 2009)

A tarifa média é formada, de modo geral, por duas parcelas: a Tarifa de Energia (TE) e a Tarifa do Uso do Sistema de Distribuição (TUSD). É obrigatória a remuneração do uso do sistema de distribuição, porém, no Ambiente de Contratação Livre (ACL) o consumidor pode

negociar, através de contratos bilaterais, de forma livre as condições, como período, preço, volume e assim por diante (COSTA; OLIVEIRA, 2019).

Para liquidar as diferenças entre a quantidade de energia contratada e a efetivamente despachada pode-se recorrer ao mercado de curto prazo, onde a energia é valorada ao Preço de Liquidação de Diferenças (PLD) (DUTRA, 2008)

O PLD é determinado e divulgado semanalmente pela Câmara de Comercialização de Energia Elétrica (CCEE), limitados por um preço máximo e mínimo, e fazendo algumas adequações, como a retirada de dados de restrições operativas. O valor divulgado é utilizado na contabilização do mercado de curto prazo (SOZZI, 2014).

Quando se analisa uma série de dados é interesse que haja uma avaliação preliminar dos resultados, ou seja, se os resultados pertencem ou não a uma determinada população, se existe uma tendência dos dados, se há uma associação entre as séries, entre outros. Para isso, podem ser usadas técnicas simples de análises de dados que permitam obter informações satisfatórias sobre as características dos resultados analisados (REIS *et al.*, 2016).

O objetivo desse estudo foi observar o histórico de preços do PLD e identificar possíveis conexões e correspondências nos dados por meio da aplicação de análises estatísticas.

2. Materiais e métodos

Os preços médios referentes a cada mês são divulgados no portal da Comércio de Comercialização de Energia Elétrica (CCEE, 2019), onde estão disponíveis para acesso livre. São divulgados os preços para os quatro submercados: Sudeste/Centro-Oeste, Sul, Nordeste e Norte. O histórico é um conjunto de dados extenso, uma vez que há dados mensais de cada submercado do ano de 2003 até o último mês disponível. No caso deste trabalho o mês de novembro de 2019.

Devido a extensão dos elementos, é necessário utilizar algum meio para auxiliar a observação desses dados e transformá-los em variáveis efetivamente úteis. A estatística permite a coleta, organização, apresentação, resumo, análise e interpretação de dados. Assim, através desta é possível aumentar a análise crítica e ter argumentos para a tomada de decisões (SANTOS *et al.*, 2016).

Para isso, inicialmente, os dados foram inseridos no software Excel, onde puderam ser feitas análises estatísticas simples, como média e desvio padrão. Com essas observações foi possível identificar certos padrões de comportamento e, a partir disso, optou-se por fazer a análise de variância e análises multivariadas.

Para realizar a análise de variância e Teste de Tukey foi selecionado o software *Sisvar*®. A análise de variância, também conhecida como ANOVA, é uma abordagem usada para comparar vários grupos de interesse, identificando diferenças consideráveis entre os grupos. O Teste de Tukey permite a formação de intervalos de confiança comparando a diferença dos pares (REIS, 2016).

No software *Tanagra*® foram feitas a análise de componentes principais análise discriminante e agrupamentos. A análise de componentes principais é usada para transformar um conjunto de variáveis originais em outro conjunto denominado de componentes principais. Cada componente principal representa uma combinação linear das variáveis originais, buscando reduzir a massa de dados, mas sem perda da informação (VARELLA, 2008). O número de componentes da Análise de Componentes Principais é definido com base na porcentagem da variabilidade total de cada componente, escolhendo, geralmente, obter 80% ou mais da variabilidade total (NEISSE; HONGYU, 2016).

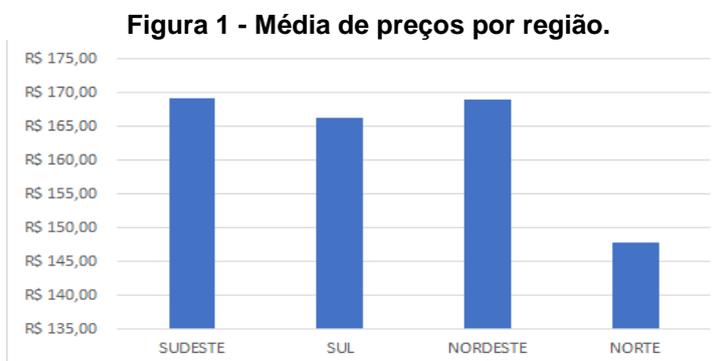
Visando ainda reduzir as dimensões do conjunto de dados, optou-se por utilizar o agrupamento. Assim, para verificar as possibilidades, foram testados os agrupamentos com o método Hierárquico (HAC) e por K-médias.

O HAC foi feito utilizando o método de Ward, ou seja, um agrupamento hierárquico que se baseia na similaridade para juntar agrupamentos. Em cada estágio, combinam-se dois agrupamentos que apresentarem o menor aumento na soma global de quadrados (SEIDEL et al., 2008).

O K-média é uma técnica de agrupamento não hierárquico, que utiliza um processo iterativo para adequar os objetos dentro de um número de classes pré-definidas, comparando a variabilidade intraclasse e interclasse (LINDEN, 2009)

3. Resultados e discussões

Para iniciar as análises optou-se por calcular as médias dos preços para diferentes variáveis. Foram calculadas as médias do histórico de preços de 2003 a 2019 para cada submercado, mostrado na Figura 1.

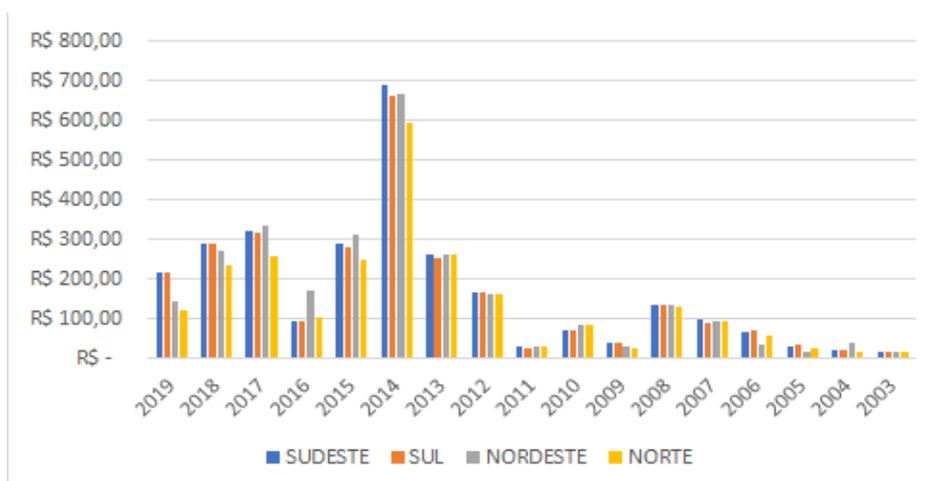


Fonte: Autoria própria (2019).

Pode-se observar que, considerando uma média de todos os preços publicados desde 2003 até 2009 o submercado norte apresenta uma média mais baixa. Os submercados são divididos desta forma em função das suas principais características, que são a capacidade de geração e o fluxo de energia (LEITE *et al.*, 2013). Assim, pode-se estimar que essa diferença seja resultante das particularidades do submercado em questão.

Em seguida, foi verificada a média anual para cada submercado em cada ano do histórico, resultando na Figura 2.

Figura 2 - Gráfico com a média anual do valor do PLD de 2003 a 2019 para cada submercado.



Fonte: Autoria própria (2019).

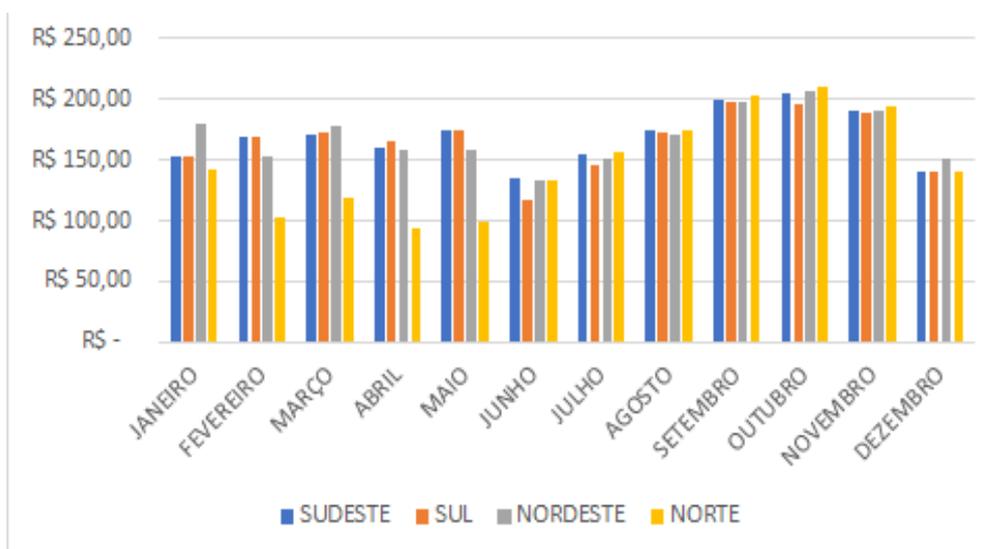
Na Figura 2, calculando as médias do preço do PLD para cada ano e submercado, pode-se notar que o preço não tem uma tendência de crescimento definida. De acordo com (LEITE, 2013), a partir de 2005, devido ao crescimento da economia e, conseqüentemente, um estreitamento da relação entre demanda e oferta, o PLD sofre significativamente com volatilidade.

A precificação da energia elétrica no Brasil ocorre em função da natureza da eletricidade, ou seja, da disponibilidade de água nos reservatórios e do nível de precipitação pluviométrico. Assim, a volatilidade dos preços está relacionada, principalmente, com a dinâmica das aflúncias (LEITE *et al.*, 2016)

Apesar dos preços estarem mais altos a partir de 2012, observa-se, principalmente, uma alta no ano de 2014. De acordo com Bertão (2014), os preços refletem as temperaturas registradas e aposta em menos precipitações no ano citado. Essa previsão de maiores gastos e dificuldades em manter os reservatórios resultou em um pessimismo do mercado e, conseqüentemente, preços mais altos.

Buscando identificar padrões de preços ao longo do ano, calculou-se a média de cada mês com o histórico de preços de 2003 a 2019, gráfico mostrado pela Figura 3.

Figura 3 - Gráfico com a média mensal do valor do PLD de 2003 a 2019 para cada submercado.



Fonte: Autoria própria (2019).

Com a média de todos os anos para cada mês, pode-se notar que, ao contrário do esperado, a energia não necessariamente tem alterações de preços relacionadas ao período úmido e seco. O período seco vai de maio a novembro, quando são esperadas poucas chuvas. No período úmido, com mais chuvas, compreende o período de dezembro a abril (ANEEL, 2012). Porém, os preços no mercado de curto prazo brasileiro não são dependentes somente da relação entre oferta e demanda, mas sim de cálculos matemáticos que incluem fatores complexos como capacidade de geração, hidrologia e pluviometria (CLÍMACO, 2010).

Na análise de variância pode-se observar com mais atenção o relacionamento entre as variáveis ano, mês e preço, conforme mostra a Figura 4.

Figura 4 - Tabela Análise de Variância obtida pelo software SISVAR®

FV	GL	SQ	QM	Fc	Pr>Fc
ANO	16	19690701.377996	1230668.836125	248.618	0.0000
MES	11	427577.954286	38870.723117	7.853	0.0000
REGIÃO	3	62360.589139	20786.863046	4.199	0.0127
ANO*MES	170	5456838.270739	32099.048651	6.485	0.0000
ANO*REGIÃO	48	242523.100715	5052.564598	1.021	0.4825
ANO*MES*REGIÃO	510	657306.213133	1288.835712	0.260	1.0000
erro	33	163351.568288	4950.047524		
Total corrigido	791	26700659.074297			
CV (%) =	43.14				
Média geral:	163.0900631	Número de observações:	792		

Fonte: Autoria própria (2019).

Os resultados obtidos indicaram que as variáveis ano e mês apresentaram o indicador p-valor baixo, de modo que pode se rejeitar a hipótese H₀, ou seja, as amostras são diferentes. Essa observação também pode ser feita para a interação entre “ano * mês”, como esta análise apresentou um valor de F maior, pode-se considerá-la como mais significativa.

O coeficiente de variação obtido foi de 43,13%. Este índice representa a variabilidade dos valores avaliados, sendo útil na definição da repetição dos resultados e estão relacionados com o erro residual nas análises de variância. O resultado obtido pode ser considerado muito alto.

Para comparar as médias dos preços com relação a variação de submercado, mês e ano foi utilizado o teste de Tukey. As Figuras 5, 6 e 7, apresentam as médias de preço para as variáveis região, mês e ano, respectivamente.

Figura 5 - Teste de Tukey para a região pelo software SISVAR®.

Tratamentos	Médias	Resultados do teste
NORTE	147.858030	a1
SUL	166.235657	a1 a2
NORDESTE	169.102626	a2
SUDESTE	169.163939	a2

Fonte: Autoria própria (2019).

As médias de preços por submercado não apresentam diferenças significativas, confirmando as informações obtidas anteriormente. Essa proximidade indica que não seria possível agrupar as informações de acordo com o submercado.

Figura 6 - Teste de Tukey para a variável mês pelo software SISVAR®.

Teste Tukey para a FV MES		
Tratamentos	Médias	Resultados do teste
MES_5	129.277941	a1
MES_11	143.113906	a1 a2
MES_3	144.369219	a1 a2
MES_1	148.283594	a1 a2
MES_4	151.803676	a1 a2 a3
MES_6	151.990735	a1 a2 a3
MES_12	157.079063	a1 a2 a3 a4
MES_2	160.110000	a1 a2 a3 a4
MES_7	173.018529	a2 a3 a4 a5
MES_10	191.548750	a3 a4 a5
MES_8	199.650588	a4 a5
MES_9	204.832647	a5

Fonte: Autoria própria (2019).

Para a variável mês, ao contrário do esperado, o Teste de Tukey confirmou que não há uma relação clara entre preço e período seco ou úmido. Pode-se, entretanto, observar que,

historicamente, os preços dos meses de agosto, setembro e outubro são mais altos, enquanto os meses com menor preço são maio, novembro e março.

O Teste de Tukey para a variável ano permite algumas análises, conforme observa-se na Figura 7.

Figura 7: Teste de Tukey para a variável ano pelo software SISVAR®.

Teste Tukey para a FV ANO		
Tratamentos	Médias	Resultados do teste
ANO_2003	18.026250	a1
ANO_2004	24.611875	a1 a2
ANO_2005	26.953125	a1 a2
ANO_2011	28.832292	a1 a2
ANO_2009	33.836458	a1 a2
ANO_2006	56.579583	a1 a2 a3
ANO_2010	77.076250	a2 a3 a4
ANO_2007	94.144792	a3 a4 a5
ANO_2016	115.706458	a4 a5 a6
ANO_2008	134.647083	a5 a6 a7
ANO_2012	163.942917	a6 a7
ANO_2019	175.808000	a7
ANO_2013	260.530625	a8
ANO_2018	270.862708	a8
ANO_2015	283.074792	a8
ANO_2017	308.144375	a8
ANO_2014	653.518542	a9

Fonte: Autoria própria (2019).

O Teste de Tukey reforçou que o ano de 2014 apresentou a maior média histórica. Neste caso, porém, observou-se tendência de agrupamento dos dados através da variável ano.

A análise de componente principal faz o agrupamento das amostras de acordo com a variação de suas características (VARELLA, 2009). Para este trabalho, optou-se por realizar essa análise considerando todas informações de meses e anos. Obteve-se a seguinte a resposta mostrada pela Figura 8.

Figura 8 - Análise de Componentes Principais para todos os dados pelo software Tanagra®.

Axis	Eigen value	Difference	Proportion (%)	Histogram	Cumulative (%)
1	1,512968	1,025937	75,65 %		75,65 %
2	0,487032	-	24,35 %		100,00 %
Tot.	2,000000	-	-	-	-

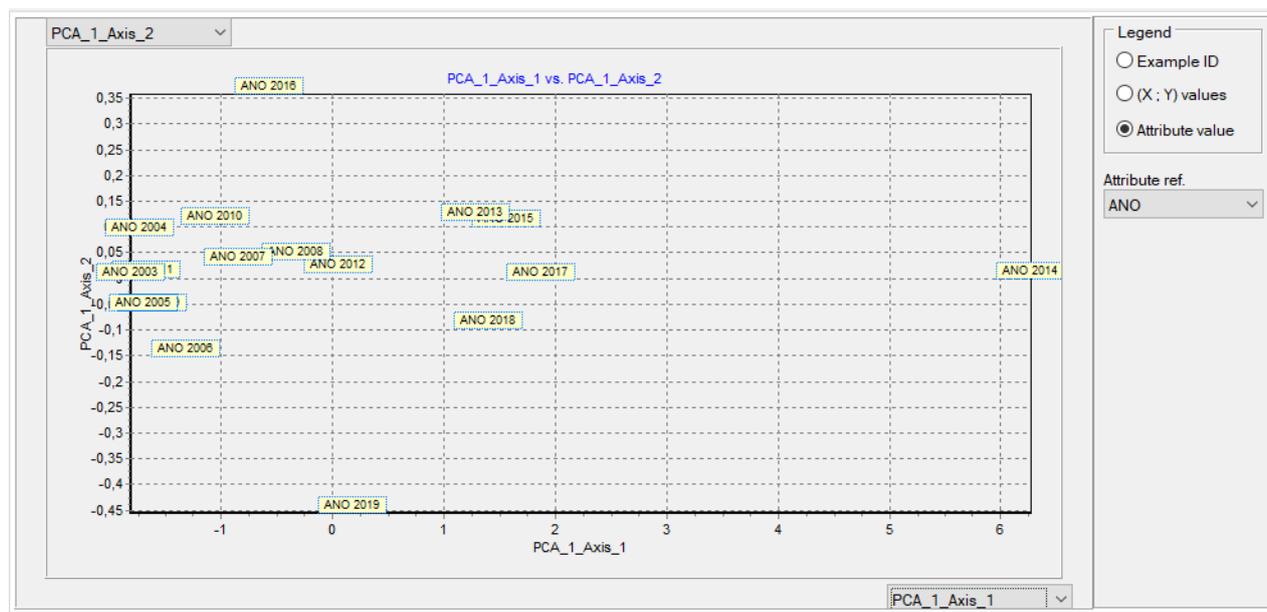
Fonte: Autoria própria (2019).

Neste caso, para todos os dados, a Análise de Componentes Principais permite que as informações sejam explicadas totalmente em dois eixos, que são formados pelas variáveis ano e preço. Como os preços não apresentam uma grande variação, o número de componentes é mais baixo. O eixo 1 corresponde a componente principal de maior peso de representação desse conjunto de dados, uma vez que tem o maior autovalor associado (1,51).

Os critérios de Kaiser-Guttman no valor 1 confirma que a transformação para dois eixos ainda representa uma parcela suficiente da população, conforme critério indicado por Neisse e Hongyu (2016).

A Figura 9 mostra a relação entre os anos e os dois eixos formados pela Análise de Componentes Principais

Figura 9: Gráfico da Análise de Componentes Principais para todos os dados pelo software Tanagra®.



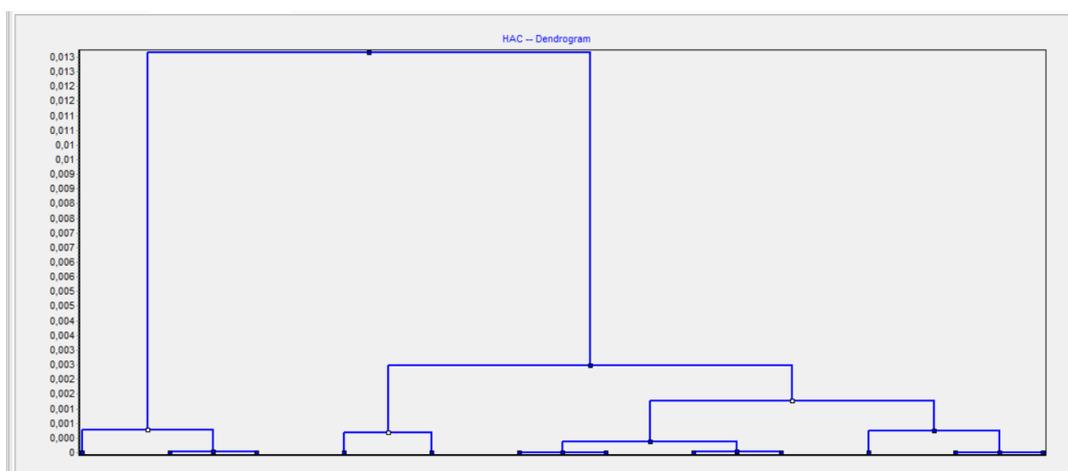
Fonte: Autoria própria (2019).

O eixo 2, representado verticalmente, não apresenta grandes mudanças e até mesmo sua escala é pequena. Porém, o eixo 1, mostrado horizontalmente, exibe variações maiores, uma vez que representa 75,65% da variabilidade.

Observa-se que, mais uma vez, o ano de 2014 mostra valores discrepantes e permanece isolado dos grupos. O posicionamento e distribuição dos anos na Figura 9 permite uma visualização de um agrupamento relacionado ao ano. Assim, para buscar um agrupamento desses dados foram realizados dois tipos de agrupamento, um hierárquico e outro não hierárquico.

No Tanagra®, a combinação para o agrupamento pelo método hierárquico é mostrada através de um dendrograma, processos exibidos na Figura 10.

Figura 10 - Dendrograma para método HAC com todos os dados do PLD pelo software Tanagra®.

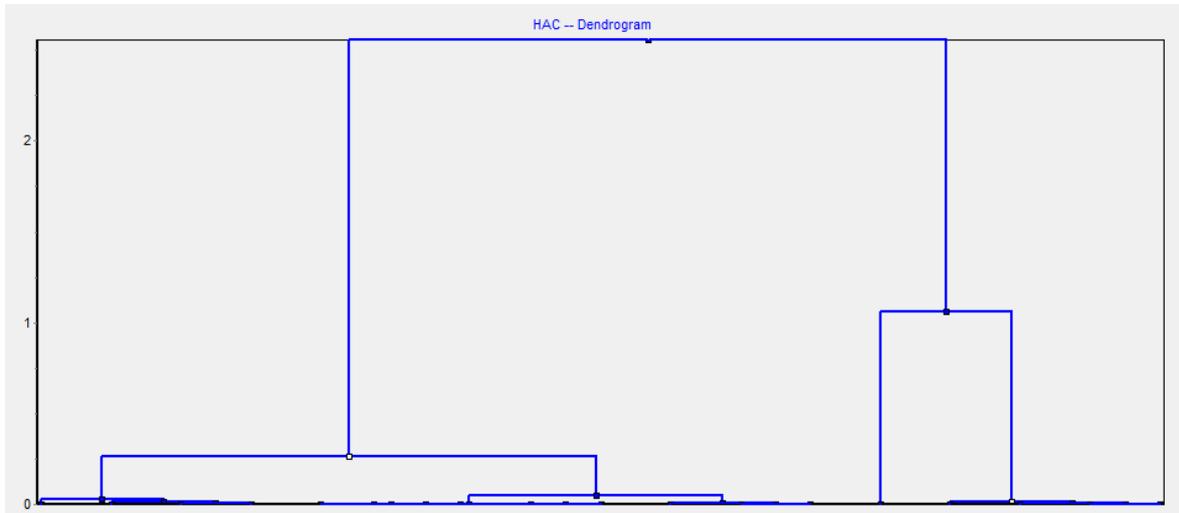


Fonte: Autoria própria (2019).

Observa-se, porém, que esse agrupamento não conseguiu reduzir satisfatoriamente o número de variáveis, principalmente, devido ao grande número de amostras e dados semelhantes.

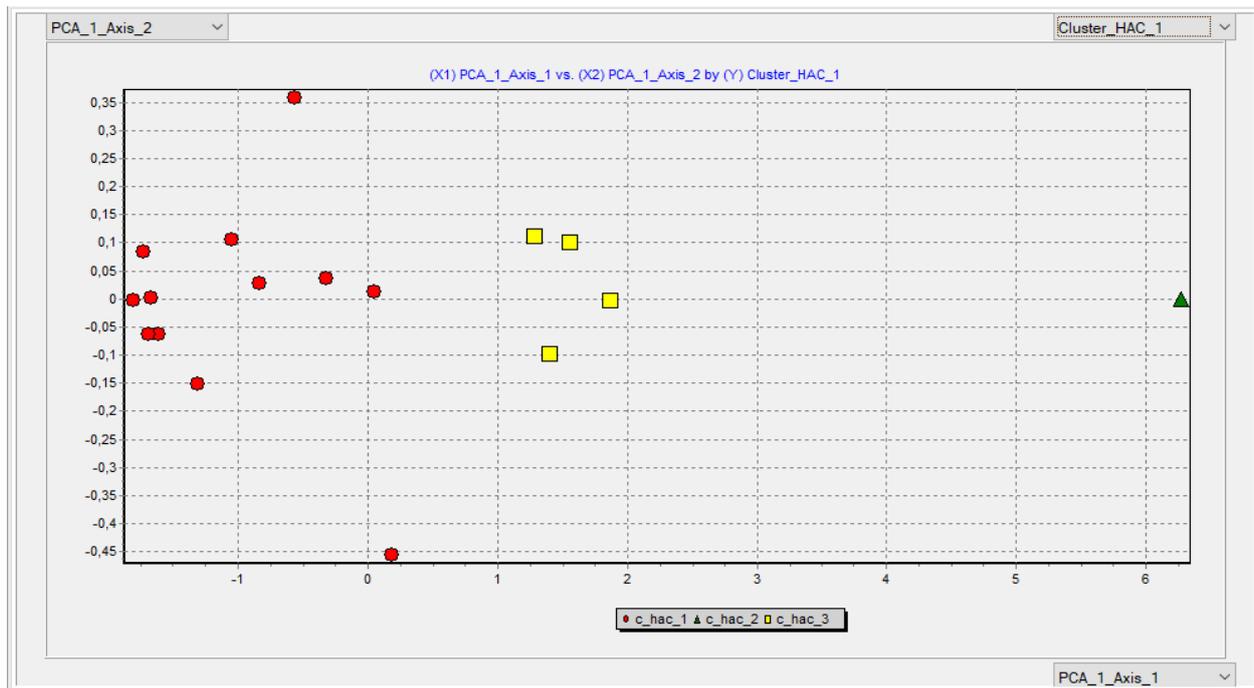
Porém, utilizando o método HAC para as médias dos preços anuais, obteve-se um resultado aceitável, como mostra o dendrograma da Figura 11 e o gráfico da Figura 12.

Figura 11 - Dendrograma para método HAC com as médias anuais do PLD usando o software Tanagra®.



Fonte: Autoria própria (2019)

Figura 12 - Agrupamento pelo método HAC para as médias de cada ano usando o software Tanagra®.

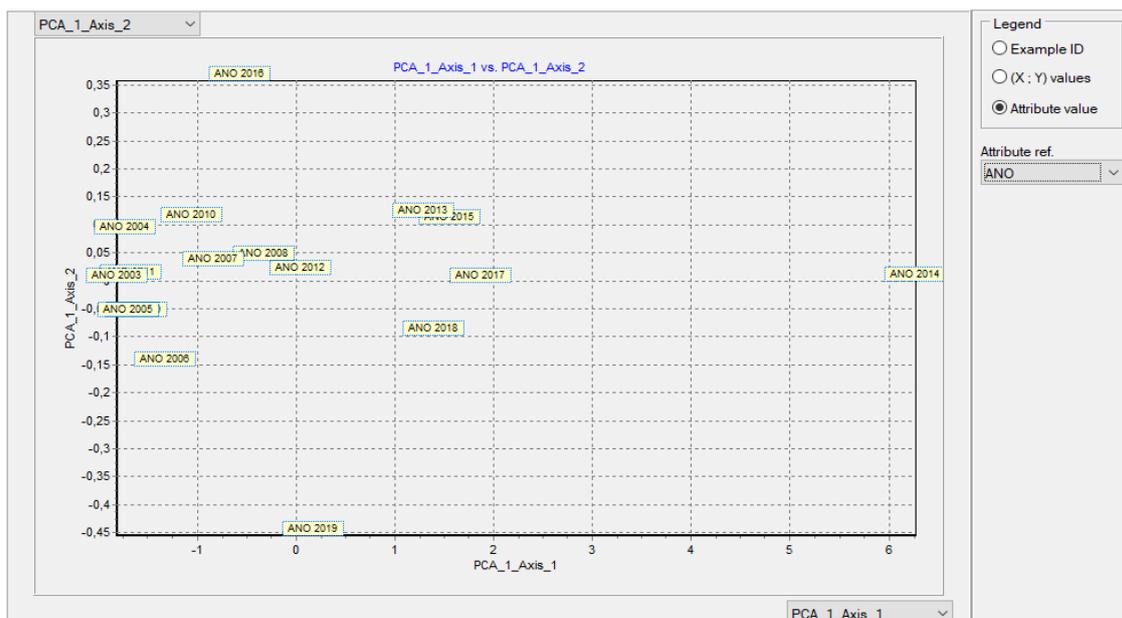


Fonte: Autoria própria (2019).

A Figura 12 mostra que, utilizando como entrada o preço médio dos anos, foi possível segregar os dados em três conjuntos, que considera-se como grupo de preços baixos, intermediários e altos.

A Figura 13 permite a visualização desses dados identificados a quais anos se referem cada ponto:

Figura 13 - Agrupamento pelo método HAC para as médias de cada ano pelo software Tanagra®.



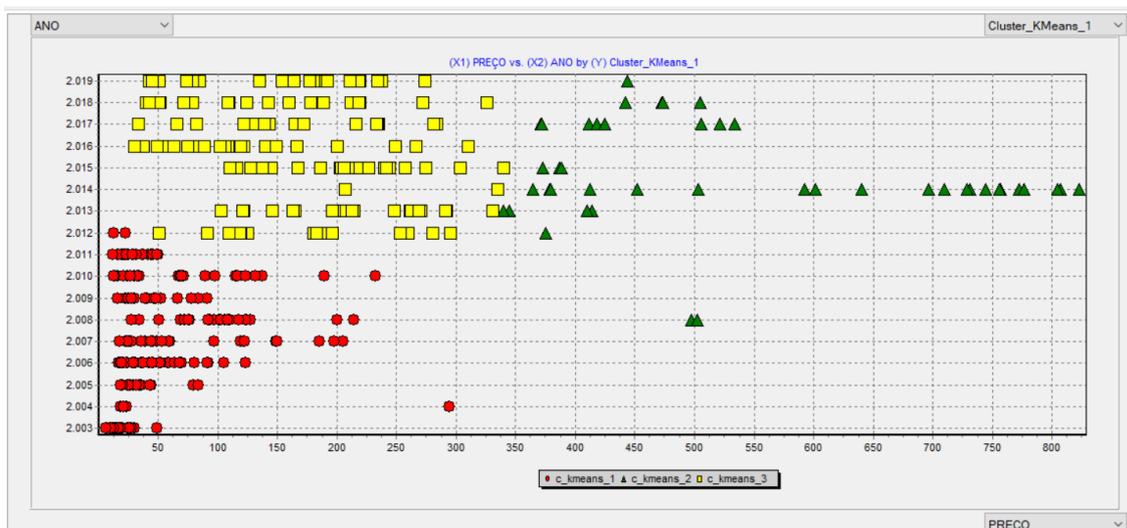
Fonte: Autoria própria (2019).

Dessa forma, na Figura 13 foi possível avaliar que, mais uma vez, o ano de maiores preços foi a no de 2014. Os anos de 2013, 2015, 2017 e 2018 apresentaram um comportamento intermediário e os outros anos apresentaram preços mais baixos que estes anos.

Para avaliar todos os dados, sem restrição de número de amostras, para variáveis com correlação alta, como o preço do PLD, é recomendado o método K-médias, uma vez que busca por amostras que tenha pouca variação dentro de um mesmo grupo.

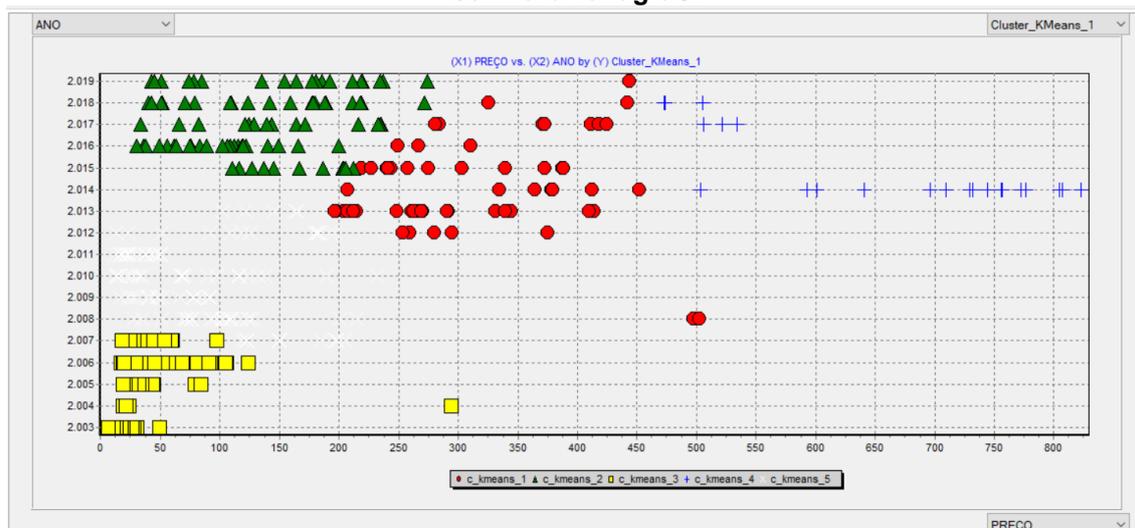
Optou-se por testar reunir os dados em 3 e 5 grupos. O gráfico 14 mostra as associações feitas para aglomerar os dados em três grupos. O gráfico 15 apresenta a compilação de informações em cinco grupos.

Figura 14: Agrupamento de todos os dados utilizando o método K-médias com 3 grupos pelo software Tanagra®.



Fonte: Autoria própria (2019).

Figura 15: Agrupamento de todos os dados utilizando o método K-médias com 5 grupos pelo software Tanagra®.



Fonte: Autoria própria (2019).

O coeficiente de determinação (R^2) foi de 0,8771 para o primeiro caso e 0,7646 para o segundo. Assim, o modelo que melhor se ajusta é o modelo com 5 grupos. O agrupamento dos valores em 5 clusters dividiu os resultados em:

- **Cluster 01:** 15,7% dos dados o correspondente a 124 valores, centralizados no ano de 2014 e nos preços de R\$ 329,58. Este agrupamento se mescla com o clusters 02 e 4 na variável ano, correspondendo a valores medianos em relação ao mercado livre de energia, espalhado ao longo dos anos de 2012 a 2019. Porém este mesmo agrupamento apresenta dois valores que ficaram dispersos no gráfico, correlacionados ao ano de 2008.
- **Cluster 02:** 18,7% dos dados o correspondente a 148 valores, centralizados no ano de 2017, com valores R\$ 140,47. Este cluster está agrupado com o clusters 01 e 04 na variável ano, e com os clusters 03 e 05 na variável preço. Correspondem aos meses dos anos entre 2012 a 2019 que foram obtidos o menor valor no mercado livre.
- **Cluster 03:** 25,8% dos dados o correspondente a 204 valores, centralizados no ano de 2004, com valores R\$ 33,58. Este agrupamento ficou em uma faixa distinta separada dos outros valores pela variável ano e preço, correspondem aos menores valores, juntamente com o cluster 5. Está mesclado pela variável preço aos clusters 02 e 05. Possui um valor discrepante que pode desestabilizar a média deste agrupamento.
- **Cluster 04:** 8,1% dos dados o correspondente a 64 valores, centralizados no ano de 2015, com valores na faixa de R\$ 636,90. Este agrupamento se mescla com o Cluster 01 e 02 na variável ano, mas corresponde aos valores mais altos com relação ao mercado livre de energia, claramente concentrados no ano 2014, onde obtivemos os valores mais altos.
- **Cluster 05:** 31,8% dos dados o correspondente a 252 valores, centralizados no ano de 2009, com valores R\$ 78,96. Está mesclado com os clusters 02 e 03 na variável preço, porém todos os valores estão agrupados em uma única região do gráfico, na área central. É o menor agrupamento. Corresponde aos valores mais baixos em relação ao mercado livre de energia.

A avaliação dos clusters pelo K-means indicou que os preços no mercado livre de energia tiveram seu ápice nos anos 2012 a 2014. Depois do ano de 2015 os índices variaram de mês a mês alternando entre baixos e medianos, com poucas incidências de valores altos.

Observa-se também que, a partir de 2010, a variação dos preços tornou-se mais expressiva. Ainda assim, mesmo após esse período, todos os anos apresentaram valores de energia baixos em alguns meses.

4. Conclusão

Neste trabalho foi possível observar como os métodos estatísticos podem auxiliar na análise de dados e tomada de decisões. A formação do preço da energia envolve uma série de variáveis, como demanda, oferta, custo de operação, clima, entre outros. Sendo assim, é difícil relacionar os dados observados sem uma análise mais criteriosa de todos atributos envolvidos.

O conjunto de dados iniciais do histórico de PLD era longo e com muitas informações. Contudo, a análise através de métodos estatísticos permitiu um diagnóstico adequado das informações

O cálculo das médias e geração de gráficos com essas variáveis permitiu identificar certos comportamentos. Ao contrário do esperado, observou-se a média dos meses ao longo do ano não teve tanta influência do preço. Todavia, o ano, com expectativas do mercado com relação aos níveis dos reservatórios, custos de operação, entre outras variáveis, mostrou uma relação forte com o preço.

A análise de variância confirmou que o grande número de amostras, além da semelhança entre as informações, mas garantindo que são diferentes. Não se observou diferenças entre as regiões apenas entre os anos e meses apontados, mesmo assim os valores estão muito aproximados nos últimos anos do estudo. As médias do Teste de Tukey permitiram identificar a possibilidade de agrupar essas informações.

A análise de componente principal foi satisfatória, uma vez que dois eixos permitiram explicar todas as informações. Porém, apesar da redução do número de informações, os gráficos ainda não apresentam claramente os grupos entre as amostras.

Os agrupamentos foram fundamentais para facilitar a visualização das informações. O método HAC não mostrou bons resultados com um número grande de informações, mas permitiu a identificação de grupos quando as médias dos anos foram utilizadas.

Comparando os métodos utilizados, constatou-se que o método do K-médias forneceu os melhores parâmetros para uma análise correta do conjunto completo de dados. A avaliação permitiu definir os grupos de valores avaliados, e constatar a tendência do mercado livre de energia, que apresenta maior volatilidade ao longo dos anos. Pode-se também observar a inconstância dos valores, e a dependência das estações e as condições climáticas anuais.

Devido a inconstância dos valores, convém manter o histórico de análises nos próximos anos, e tentar estimar os valores por sazonalidade.

Referências

Agência Nacional de Energia Elétrica - ANEEL. Resolução normativa Nº 479 de 3 de abril de 2012. Disponível em:<<http://www2.aneel.gov.br/cedoc/ren2012479.pdf>>. Acesso em 08 de dezembro de 2019.

BERTÃO, N. I. Preço da energia dispara e previsão é de calamidade para o ano. **Revista Veja**. Publicado em 01 de fevereiro de 2014. Disponível em:<>. Acesso em 08 de dezembro de 2019.

Câmara de Comercialização de Energia Elétrica - CCEE. Preços médios. Disponível em:<https://www.ccee.org.br/portal/faces/pages_publico/o-que-

fazemos/como_ccee_atua/precos/precos_medios?_afzLoop=261555840371780&_adf.ctrl-state=14321gj3g1_27#!%40%40%3F_afzLoop%3D261555840371780%26_adf.ctrl-state%3D14321gj3g1_31>. Acesso em 02 de dezembro de 2019.

CLÍMACO, F. (2010). Gestão de consumidores livres de energia elétrica. São Paulo: Universidade de São Paulo.

COSTA, C., & OLIVEIRA, E. E. (2019). O mercado brasileiro de Energia Elétrica Livre: um estudo de caso na indústria de Shopping Center sob a ótica da gestão eficiente dos recursos financeiros. **Revista Multidisciplinar e de Psicologia**, 277-305.

DUTRA, J. G. (2008). Estudo da negociação de contratos bilaterais em sistemas predominantemente hidráulicos. Brasília: Universidade Federal de Brasília.

FLOREZI, G. Consumidores livres de energia: uma visão prática. Universidade de São Paulo. São Paulo, 2009.

GUIDINI, Marilene Bertuol et al. Aplicação do K-means cluster para Classificar estilos gerenciais. **Contextus – Revista Contemporânea de Economia e Gestão**, [s.l.], v. 6, n. 2, p.25-34, 1 jan. 2008. Revista Contextus Universidade Federal do Ceará.

LEITE, A.L.S. ET AL. Causas da volatilidade do preço spot de eletricidade no Brasil. **Ensaio FEE** p 647-668. Porto Alegre, 2013.

LINDEN, R. Técnicas de agrupamento. Revista de Sistemas de Informação da FSMA. pp 18-36, 2009.

NEISSE, A.C. HONGYU, K. Aplicação de Componentes Principais e Análise Fatorial a Dados Criminais de 26 Estados dos EUA. **E & S - Engineering and Science**, 2016.

REIS, B.C.M. ET.AL. Aplicação da análise estatística via teste de Tukey e análise de variância para a avaliação dos parâmetros de qualidade de corte em um processo de fabricação. XXXVI Encontro Nacional de Engenharia de Produção. João Pessoa, 2016.

SANTOS, B.M. ET.AL. A importância do uso da estatística na área empresarial: uma pesquisa de campo com empresas do município de Elói Mendes-MG. Simpósio de Excelência em Gestão e Tecnologia, 2016.

SEIDEL, E.J. ET. AL. Comparação entre o método Ward e o método K-médias no agrupamento de produtores de leite. **Revista Ciência e Natureza** - UFSM, 2008.

SOZZI, G. Formação do preço da energia convencional nas transações entre agentes no mercado de curto prazo brasileiro. Dissertação de mestrado da Universidade de São Paulo. São Paulo, 2014.

VARELLA, C.A.A. Análise de componentes principais. Apostila da disciplina de Análise Multivariada Aplicada as Ciências Agrárias. Programa de pós-graduação da Universidade Federal Rural do Rio de Janeiro. Seropédica, 2008. Disponível em:<<http://www.ufrj.br/institutos/it/deng/varella/Downloads/multivariada%20aplicada%20as%20ciencias%20agrarias/Aulas/analise%20de%20componentes%20principais.pdf>>. Acesso em 03 de dezembro de 2019.