

Eixo Temático: Inovação e Sustentabilidade

**PREVISÃO DO ÍNDICE DE VOLUME DE VENDAS NO VAREJO -
COMBUSTÍVEIS E LUBRIFICANTES NO ESTADO DO RIO GRANDE DO SUL.**

**PREDICTION OF VOLUME OF RETAIL SALES INDEX - FUELS AND
LUBRICANTS IN THE STATE OF RIO GRANDE DO SUL.**

Claudia Aline de Souza Ramser, Tatiana Fernanda Mousquer dos Santos, Adriano Mendonça Souza
Adenise Clereci, Antonio Vanderlei dos Santos e Maria Aparecida Brum Trindade

RESUMO

Objetivou-se realizar previsões para o ano de 2014, referente à venda no varejo de combustíveis e lubrificantes, no Estado do Rio Grande do Sul. Pois sabemos que cada vez há mais veículos trafegando no estado, também sabemos que o petróleo é um recurso natural não renovável, a partir daí surgiu o interesse em demonstrar a quantidade de combustíveis e lubrificantes utilizados no estado. Os dados utilizados do índice de volume vendido no varejo foram utilizados uma média mensal, do ano de 2000 a 2014. Para realização das previsões dos combustíveis e lubrificantes no Estado do Rio Grande do Sul foi utilizada Metodologia Box & Jenkins. Através desta metodologia calculou vários modelos, mas por meio dos Critérios AIC e BIC, escolheu-se o melhor modelo para a série, sendo ideal o modelo SARIMA (1,1,1)₁₂.

Palavras-chave: previsão, combustíveis, modelos, Box & Jenkins.

ABSTRACT

Make predictions for the year 2014 The objective of referring the sale at retail of fuels and lubricants in the state of Rio Grande do Sul For we know that every day there are more vehicles traveling in the state, we also know that oil is a natural resource non-renewable thereafter the interest shown me the amount of fuels and lubricants used in the state emerged. The data used in the index volume sold at retail were used a monthly average of 2000 to 2014. To perform the predictions of fuels and lubricants in the state of Rio Grande do Sul Box & Jenkins methodology was used. Through this methodology, calculated several models, but by the AIC and BIC criteria, we chose the best model for the series, ideal SARIMA (1,1,1)₁₂ model.

Keywords: weather, fuel, models, Box & Jenkins.

INTRODUÇÃO

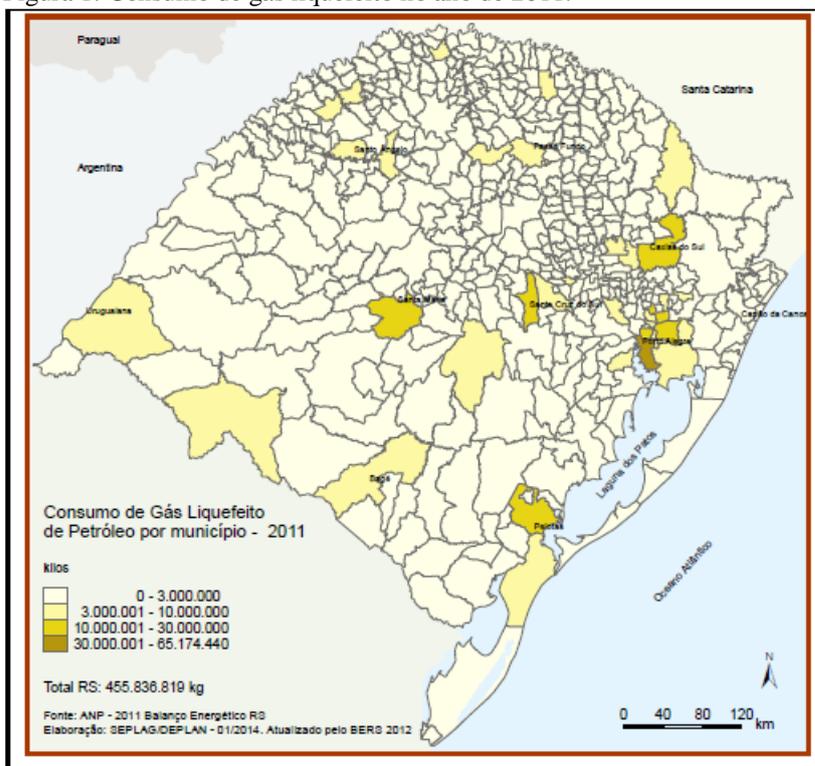
Conforme informações disponibilizadas pela Secretaria de Planejamento, Gestão e Participação Cidadã do Rio Grande do Sul (SEPLAG):

A Refinaria Alberto Pasqualini - REFAP em Canoas e a Refinaria Riograndense em Rio Grande, são as responsáveis pelo total do refino do petróleo no Rio Grande do Sul. Em 2011, ambas processavam 165.147 barris/dia, o que correspondia a 8,8% da capacidade nominal nacional. Os principais derivados extraídos são: óleo diesel, óleo combustível, gasolina, gasolina de aviação, gás liquefeito de petróleo - GLP e querosene. (CARGNIN et al, s/a, s/p).

O maior gasto com combustíveis derivados do petróleo, com exceção do óleo diesel, acontece nos principais centros urbanos do Estado onde é utilizada a maior parte da frota de veículos. O consumo de óleo diesel distribui-se mais uniformemente pelos municípios, sobretudo, a sua utilização nas atividades rurais, com destaque para a lavoura arrozeira.

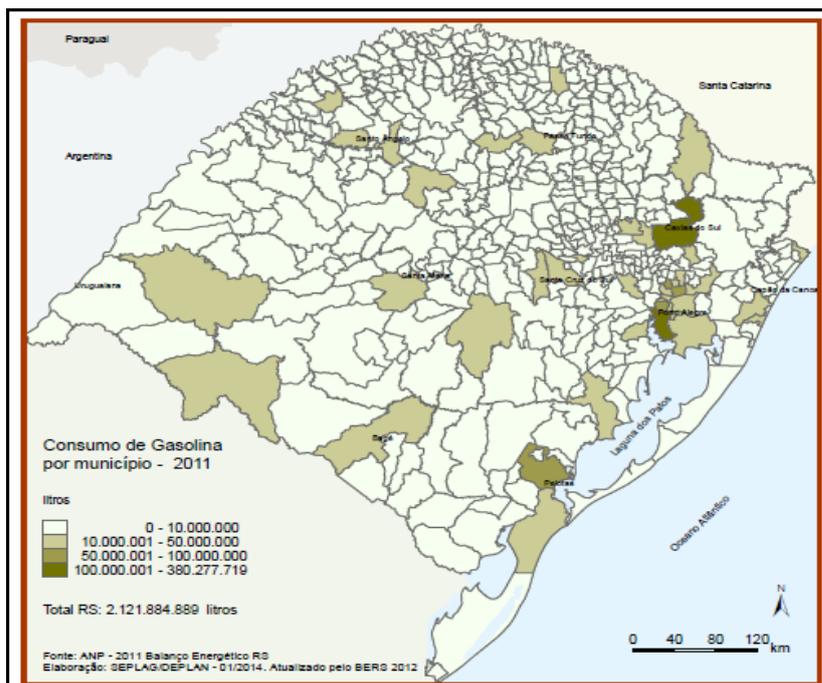
Conforme informações coletadas em diversos níveis de agrupamento das informações geográficas foi demonstrado pelos autores Cargnin et al (s/a) mapas do uso de combustíveis no RS.

Figura 1: Consumo de gás liquefeito no ano de 2011.



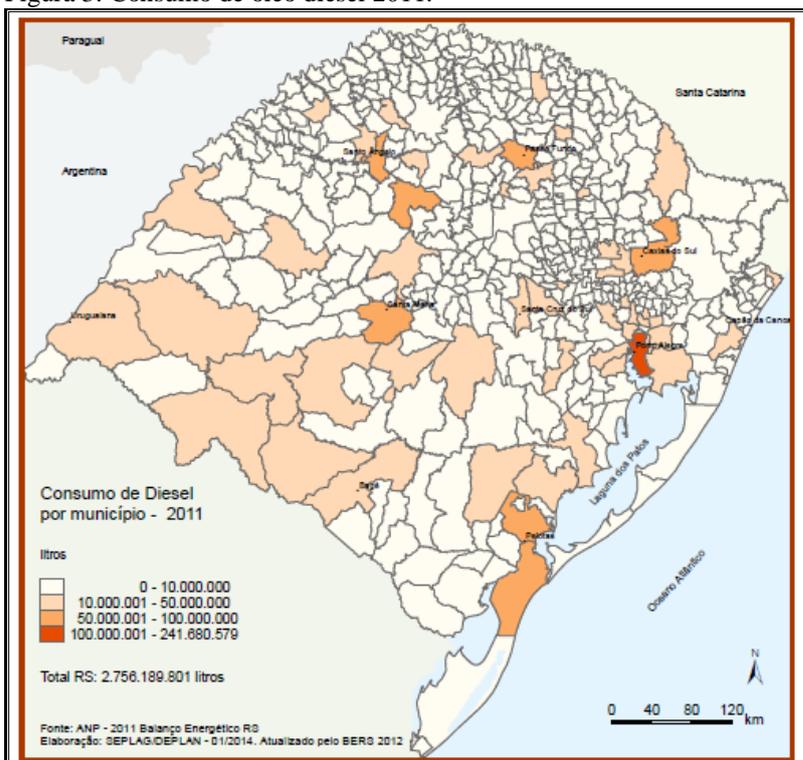
Fonte: Cargnin et al (s/a)

Figura 2: Consumo de gasolina no ano de 2011.



Fonte: Cargnin et al (s/a)

Figura 3: Consumo de óleo diesel 2011.



Fonte: Cargnin et al (s/a)

Por meio dos mapas acima demonstrados na figura 1, 2 e 3, podemos identificar as regiões onde se concentra o maior consumo de combustíveis.

Material e métodos

Conjunto de dados

O conjunto de dados utilizado neste trabalho provém do site do IBGE e refere-se a uma série mensal de venda mensal (em litros) de combustíveis e lubrificantes no período de janeiro de 2000 até abril de 2014.

Conforme abordagem de Morettin e Tolo, (2004), uma série temporal (ST) é um conjunto de observações ordenadas no tempo, que devido a essa ordenação cronológica, surge o efeito da autocorrelação entre as observações. Esta amostragem devem ser colhidas em intervalos de tempo iguais.

A ST (Z_t) pode ser discreta, contínua, univariada, ou multivariada e o tempo (T) identificado como unidimensional ou multidimensional.

A análise de séries temporais apresenta como objetivo principal, a realização de previsões futuras, adotando por base seus valores presentes e/ou passados, por meio da correlação temporal que existe geralmente entre os valores exibidos pela série.

Segundo o autor Lima, a finalidade da análise de séries temporais é a verificação de previsões. Sendo assim, utiliza-se uma metodologia que estabelece mecanismos onde:

Valores futuros de uma série possam ser previstos com base apenas em seus valores presentes e passados. As técnicas desse processo de previsão fundamentam-se na exploração da correlação temporal que pode existir entre os valores exibidos pela série. (LIMA *et al*, 2010, p.189)

Diante do objetivo de conhecer o processo gerador da série temporal, encontram-se necessárias a identificação da série, para a realização das previsões a horizontes futuros, na qual se utilizam os dados históricos.

Recomenda-se a utilização de ST estacionárias, pois só assim tornam-se os efeitos sazonais aditivos e a variância residual constante garantindo a ergodicidade (representatividade) do processo. (SOUZA, s/a, s/p).

De acordo com Morettin e Tolo, (2004), a série a ser estudada deve ser estacionária, isto quer dizer, a série deve se distribuir aleatoriamente em torno de uma média constante, admitindo certa estabilidade ao longo do tempo. Dessa forma, garante-se que os parâmetros estimados para essa série sejam válidos, e confiáveis.

Para auxiliar os estudos e análises das séries temporais George Box e Gwilyn Jenkins em 1976 desenvolveram um método de análise de séries temporais que se tornou mundialmente conhecido como metodologia de Box & Jenkins, sendo que a previsão é um dos principais fatores da popularidade deste modelo. (Souza et al., 2011).

Conforme Bueno (2008), para distinguir os possíveis modelos devemos analisar a função da autocorrelação (FAC) e a função de autocorrelação parcial (FACP). Conforme tabela 1.

Tabela 1- Identificação dos modelos AR(p) e MA(q).

Modelo	FAC	FACP
AR(p)	Decaimento exponencial	Truncada na defasagem p
MA(q)	Truncada na defasagem q	Decaimento exponencial
ARMA (p, q)	Decai exponencialmente se $j > q$	Decai exponencialmente se $j > q$

A tabela acima explica como se identifica o modelo por meio da função de autocorrelação (FAC) e função de autocorrelação parcial (FACP).

Um processo aleatório linear simples é o processo autoregressivo AR(p) em que $p = 1$. O modelo AR(p) correspondente ao processo autoregressivo de 1ª ordem ou 2ª ordem.

O modelo autoregressivo genérico AR(p) escreve-se:

$$\tilde{Z}_t = \varphi_1 \tilde{Z}_{t-1} + \varphi_2 \tilde{Z}_{t-2} + \dots + \varphi_p \tilde{Z}_{t-p} + a_t$$

Na verdade, como o nome indica, AR(p) modela uma autoregressão da variável Y_t com essa mesma variável, defasada ($Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$), para os “p” períodos de defasagem em que a autocorrelação parcial entre as variáveis é significativa. Onde o p indica a ordem do modelo, ou o número de defasagens.

Já o MA(q) quer dizer modelo de médias móveis de ordem “q” significa que o Z_t é uma função soma algébrica ponderada dos a_t , que se movem no tempo, Z_t não é média, visto que os parâmetros θ 's não somam, obrigatoriamente, a unidade. No modelo MA (q) onde “q” determina a ordem do modelo e θ é um parâmetro. (SOUZA, 2006).

ARMA

O modelo ARMA (p, q) é formado por uma parte autoregressivo e a outra de médias móveis, onde o modelo se encontra estacionário. Podendo ser chamado método de previsão de mínimos quadrados.

ARIMA

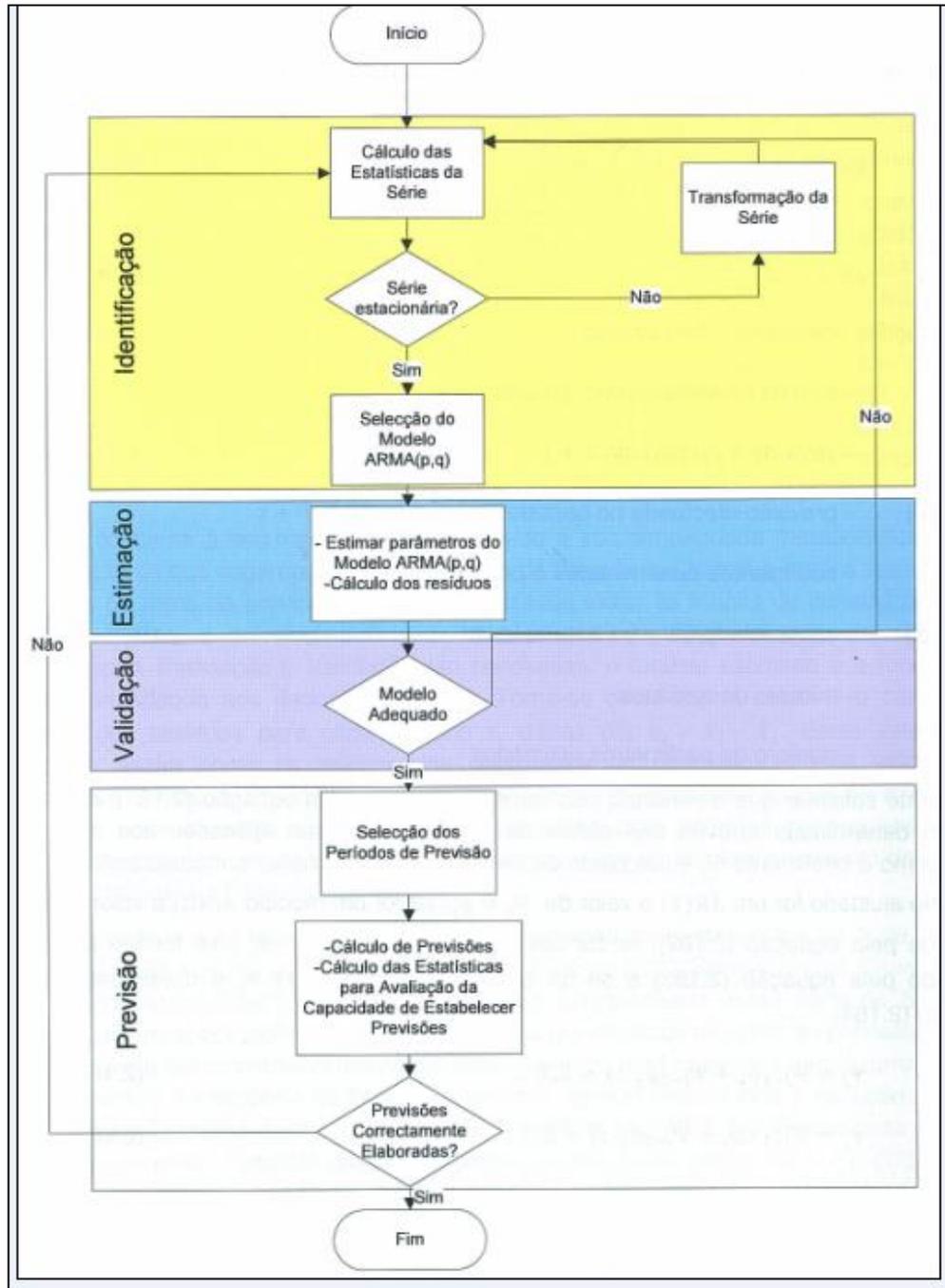
O modelo ARIMA gera da combinação de três componentes Autoregressivo (AR), o componente de Integração (I) e o componente médias móveis (MA). Este modelo não é estacionário.

A metodologia de (Box- Jenkins, 1970) determina em ajustar os modelos ARIMA (p, d, q) a um conjunto de observações em um determinado período, onde p é a parte autoregressiva, d o número de diferenças aplicados na série para deixa-lá estacionária e q é a parte de médias móveis.

Já o modelo SARIMA é um processo sazonal autoregressivo integrado de média móvel de ordem (p, d, q) $X(P, D, Q)_s$ denotado por SARIMA (p,d,q) $X(P,D,Q)_s$ onde d é a diferenciação e o D é o grau de diferenciação sazonal.

A tática usada para construção do modelo será feita por meio de um ciclo interativo conforme fluxograma da metodologia Box e Jenkins (1970) adaptada de Zulema e Requeijo (2008):

FIGURA 1 – Fluxograma da metodologia Box e Jenkins (1970) adaptada de Zulema e Requeijo (2008)



Fonte: Zulema e Requeijo (2008)

Identificação, estimação, verificação e previsão. Com a identificação saberemos se a série é estacionária ou não-estacionária, para confirmar a estacionariedade usaremos os testes Dick-Fuller (ADF) e KPSS. A seleção do modelo será feita pela análise da função de autocorrelação (FAC) e função de autocorrelação parcial (FACP) A estimação pode ser analisada pelo AIC e BIC. Com a verificação saberemos qual o melhor modelo será usado para por fim fazer a previsão.

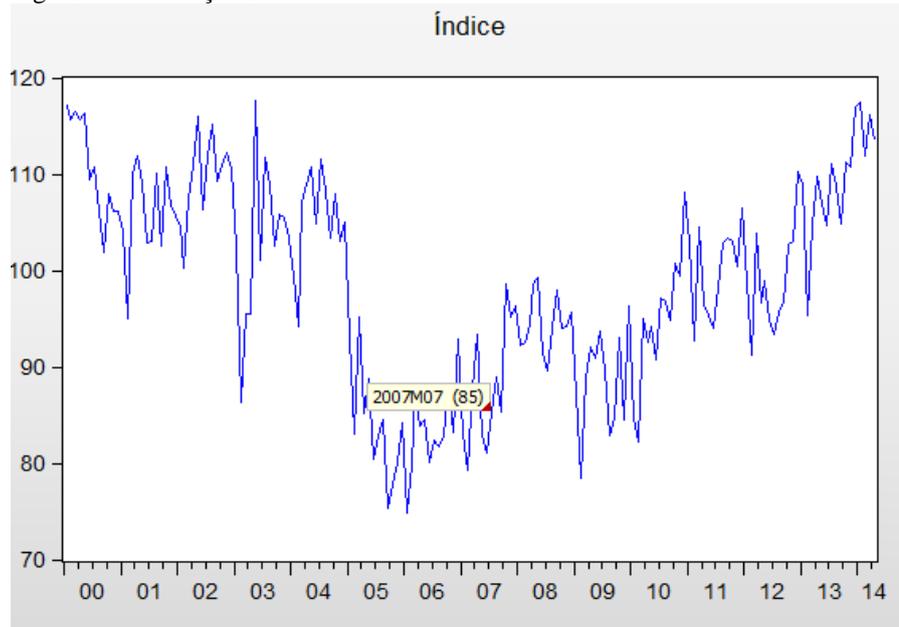
RESULTADO

Para analisar o comportamento da Série do Índice de volume de vendas no varejo - Combustíveis e lubrificantes - Rio Grande do Sul, de janeiro de 2000 a abril de 2014, no Rio

Grande do Sul, desenvolveu-se um Gráfico da Série Original, que possibilitou observar o comportamento ao longo do período de análise.

Na figura 2, foi possível observar que as Séries não assumem um comportamento estacionário por apresentar uma tendência crescente e picos. Como a análise visual não é comprobatória recorreu-se aos testes Dick-Fuller (ADF) e KPSS, através da análise das Funções de Autocorrelação de cada Série Original, e aos testes de significância, para se concluir se as Séries apresentam tendência.

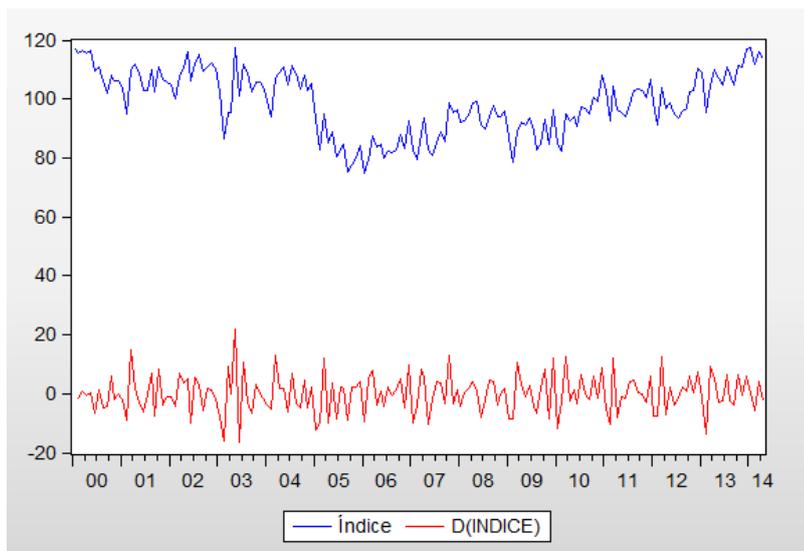
Figura 2: Verificação de estacionariedade.



Fonte: Elaborado pelos autores

Na figura 2 visualmente identificamos que a série não é estacionária, portanto devemos aplicar uma diferença, mas para confirmarmos esta não estacionariedade fizemos o teste de estacionariedade Kwiatkowski, Phillips, Schmidt and Shin (KPSS), e o Teste da raiz unitária Augmented Dickey-Fuller (ADF).

Figura 3: Paralelo do gráfico da verificação de estacionariedade, com aplicação da diferença.



Fonte: Elaborado pelos autores

Após aplicada esta diferença podemos observar na figura 3 em paralelo com o gráfico da série sem a diferença, que a série ficou estacionária, mas para confirmação aplicaremos os testes ADF e KPSS, para assim comprovar que a série se tornou estacionaria.

Tabela 2- Teste da raiz unitária Augmented Dickey-Fuller (ADF)

	Em nível	Em primeira diferença
Valor crítico do ADF	-0,9665	-3,9229
5%	-1.9427	-2,8796
p-valor	0,3352	0,0024

Fonte: Elaborado pelos autores

Tabela 3- Teste de estacionariedade Kwiatkowski, Phillips, Schmidt and Shin (KPSS).

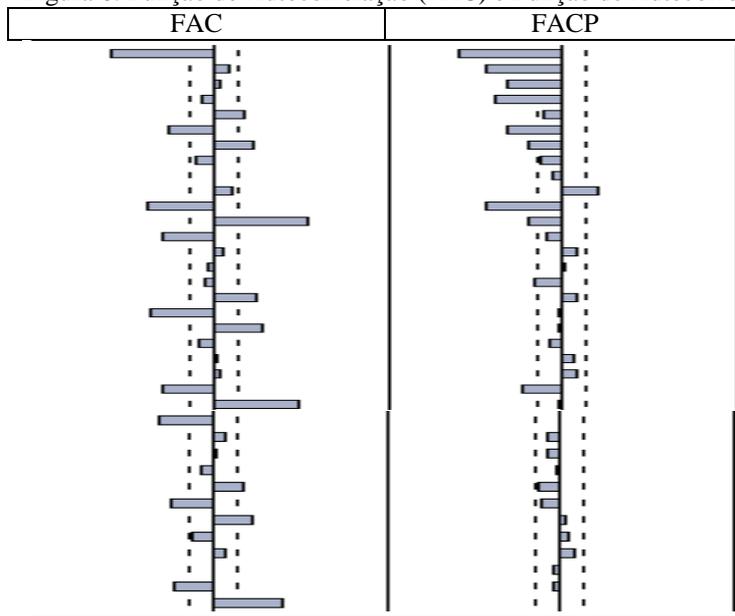
	Em nível	Em primeira diferença
Valor crítico do KPSS	0,4015	0,3703
5%	0,4630	0,4630
p-valor	0,0000	0,9668

Fonte: Elaborado pelos autores

Analisada a série diferenciada, o próximo passo é identificação dos modelos significativos que representam o comportamento do Preço das vendas no varejo de combustíveis e derivados. Através da estimação de vários modelos, observou-se os modelos significativos encontrados bem como seus parâmetros. Como regra de decisão para optar pelo melhor modelo, usou os Critérios o AIC e o BIC, que definem os modelos parcimoniosos, por meio dos menores valores encontrados. Tais valores encontram-se na tabela 2. Encontrados os modelos significativos, o próximo procedimento é a verificação das Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP) de cada modelo verificando se os

resíduos não se apresentam autocorrelacionados. As FAC e FACP (Figura 6) satisfizeram esta condição e encontraram-se dentro dos limites de confiabilidades.

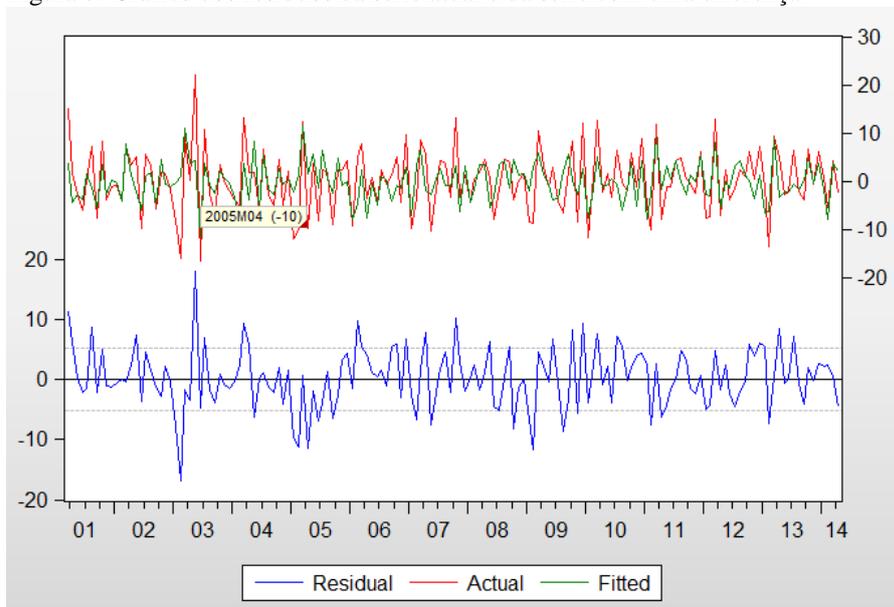
Figura 6: Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP)



Fonte: Elaborado pelos autores

A FAC apresentou ϕ_1 negativo e a FACP decai exponencialmente, mas esta apresentando sinais positivos e negativos denominando uma parte do modelo como autorregressivo .

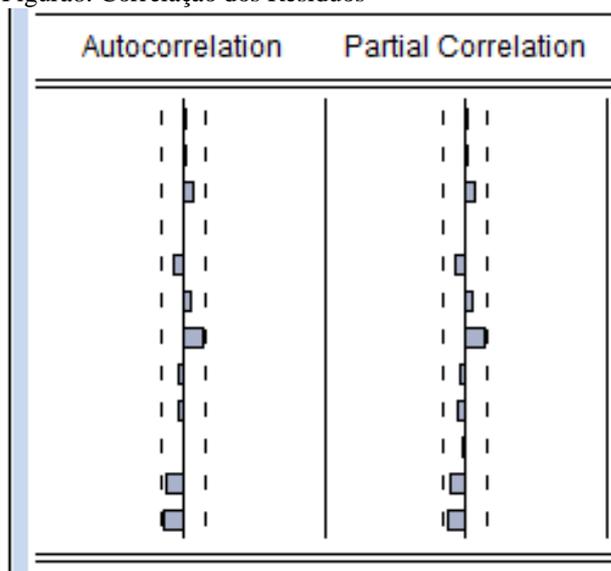
Figura 7: Gráfico dos resíduos da série atual e da série com uma diferença



Fonte: Elaborado pelos autores

A figura 7 esta demonstrando o comportamento do resíduo da série atual e do ajustado.

Figura8: Correlação dos Resíduos



Fonte: Elaborado pelos autores

Conforme a figura 8 ela descreve que os resíduos são ruídos branco, foi realizada a análise dos resíduos assim como no primeiro modelo e todas as suposições foram verificadas. Auxiliando na escolha do melhor modelo. Portanto o modelo representativo da série está em negrito na tabela abaixo:

Tabela 2 - Modelos significativos para a Série

MODELO	AIC	BIC	PROB.
SARIMA (1,1,1)₁₂	6,11	6,17	0,42
MA(1)	6,32	6,34	0,00
ARIMA (1,1,1)	6,34	6,37	0,34
ARI (1)	6,41	6,46	0,00

Fonte: Elaborado pelos autores

CONCLUSÃO

O estudo realizado teve como objetivo encontrar um modelo de previsão confiável para a venda no varejo de combustíveis e lubrificantes, no Estado do Rio Grande do Sul no ano de 2014, pois como se sabe cada vez há mais veículos trafegando no estado, também e o petróleo é um recurso natural não renovável, a partir daí surgiu o interesse em demonstrar a quantidade de combustíveis e lubrificantes utilizados no estado.

Assim através dos resultados obtidos verifica-se que a previsão de melhor qualidade ou que melhor se ajustou aos dados, segundo os critérios utilizados neste presente estudo, foi o modelo SARIMA (1,1,1)₁₂.

Com isso chega-se à conclusão de que a venda de combustíveis e lubrificantes pode se beneficiar da utilização de métodos de previsão a partir de séries temporais, pois esta metodologia possibilitou efetuar a análise das informações tentando retirar conhecimento útil das mesmas e assim prever o futuro, por meio da análise dos dados estudados e com isso construir um modelo que permita antever a evolução futura da série temporal.

REFERENCIAS BIBLIOGRAFICAS

BOX, G. E. O.; JENKINS, G.M. **Time series Analysis: Forecasting and Control**. San Francisco: Holden-Day, 1970.

BUENO, R. L. S. **Econometria de Séries Temporais**. São Paulo: Cengage Learning, 2008.

LIMA, C, R; Góis, R M; Ulises, C. **Previsão de preços futuros de Commodities agrícolas com diferenciações inteira e fracionária, e erros heteroscedásticos**. Rev. Econ. Sociol. Rural vol.45 no.3 Brasília July/Sept. 2007. Disponível em:
http://www.scielo.br/scielo.php?pid=S0103-20032007000300004&script=sci_arttext. Acesso em 06 de junho de 2014.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. 1ª edição. Editora Edgard Blucher. 2004.

PEREIRA, Zulema L.; Requeijo, José G., **Planejamento e Controlo Estatístico dos Processos**, Prefácio, 2008.

SOUZA, Francisca Mendonça; ALMEIDA, Silvana Gonçalves; SOUZA, Adriano Mendonça; LOPES, Luis Felipe Dias; ZANINI, Roselaine Ruviano; **Previsão do preço da gasolina para a região sul do Brasil**. IJIE – Ibero american Journal of industrial Engineering. v.3, n.1, pg 234-248, julho, 2011.

SOUZA. Mendonça, Francisca. Modelos Box & Jenkins Aplicados a Previsão de Demanda de Leitos Hospitalares. Monografia de Especialização. UFSM, 2006.

VASCONCELLOS, M. A. S.; ALVES, D. **Manual de econometria**. São Paulo. Atlas, 2002.