

MÉTODOS DE PREVISÃO APLICADOS AO CONSUMO BRASILEIRO DE COMBUSTÍVEIS LEVES

FORECAST METHODS APPLIED TO THE BRAZILIAN CONSUMPTION OF LIGHT FUELS

Autor(es): Lucas Rodrigues e Luciano Rodrigues

Filiação: Escola Superior de Agricultura Luiz de Queiroz (ESALQ/USP)

E-mail: lucas2.rodrigues@usp.br; lurodrig@usp.br

Grupo de Trabalho (GT): GT1. Mercados Agrícolas e Comércio Exterior

Resumo

Neste estudo foram avaliados diferentes métodos de previsão para projetar o consumo de combustíveis leves no Brasil (combustíveis utilizados pelos veículos com motor de combustão interna). Para tanto, 11 diferentes métodos foram implementados, tomando como referência o erro de previsão e o procedimento de validação cruzada empregado para séries temporais. A avaliação foi conduzida para período de previsão de 3, 6 e 12 meses. Os testes realizados indicam a superioridade dos modelos avaliados frente a um método ingênuo de previsão. Em linhas gerais, os resultados indicam que, a despeito das mudanças tecnológicas, das alterações de mercado e dos movimentos institucionais observados no setor de combustíveis leves nos últimos anos, os modelos de previsão implementados permitiram previsões bom nível de acurácia.

Palavras-chave: Demanda de Combustíveis; Métodos de Previsão; Séries Temporais; Etanol; Gasolina

Abstract

In this study, different forecasting methods were evaluated to project the consumption of light fuels in Brazil (fuels used by vehicles with an internal combustion engine). Eleven different methods were implemented, based on the forecast error and the cross-validation procedure used for time series. The evaluation was conducted for a forecast period of 3, 6 and 12 months. The tests performed indicate the superiority of the evaluated models compared to a naive forecasting method. In general, the results indicate that, in spite of technological changes, market changes and institutional movements observed in the light fuel sector in recent years, the forecasting models implemented allowed forecasts of a good level of accuracy.

Key words: Fuel Demand; Forecasting Methods; Time Series; Ethanol; Gasoline

1. Introdução

A regulação e a política energética exigem previsibilidade para garantir o abastecimento regular e atender a demanda doméstica de um país. Do ponto de vista econômico, essa premissa é fundamental para a melhor decisão na alocação de recursos no curto e longo prazos. Na perspectiva social, a regularidade do suprimento energético deve incorporar os efeitos de uma crise de abastecimento sobre o bem-estar da sociedade.

O gerenciamento da demanda por energia se tornou uma questão mandatória para agentes privados que atuam na área energética e para representantes do poder público, com finalidade de planejar a necessidade de recursos futuros, promover medidas de conservação ambiental, otimizar a utilização de recursos e fontes energéticas, auxiliar a tomada de decisão dos agentes e garantir regularidade do suprimento energético (SUGANTHI; SAMUEL, 2012).



Nesse contexto, o uso de modelos quantitativos para a realização de previsões é um importante ferramental que auxilia nessa tarefa.

Dentre as diversas fontes de energia, encontra-se os combustíveis líquidos utilizados por veículos a combustão, que ainda são majoritariamente utilizados nos dias de hoje, para atender as necessidades de transporte privado.

No Brasil, o setor de combustíveis líquidos passou nos últimos anos por intensas mudanças em relação as políticas econômicas que ditam o funcionamento dessa indústria (RODRIGUES; RODRIGUES, 2018). Tem-se, portanto, uma série de fatores que tornam complexa e incerta a formação de preços, a lógica de produção e, conseqüentemente, a dinâmica do consumo de combustíveis leves no País (gasolina, etanol e GNV).

Nesse contexto, o objetivo do presente estudo compreende a avaliação da performance de diversos métodos de previsão para projetar o consumo de combustíveis do ciclo Otto no mercado brasileiro.

Em linhas gerais, os resultados devem trazer indicações importantes para as seguintes questões: qual a melhor metodologia do ponto de vista de erro de previsão? É satisfatória a performance dos modelos a ponto de utilizá-los como ferramenta de apoio à tomada de decisão? Existe diferença significativa na performance dos diferentes métodos?

Para atender o objetivo proposto, o artigo foi dividido em 3 sessões. A próxima seção concentra-se em descrever as principais características do mercado de combustíveis brasileiro e levanta os principais estudos realizados que estimaram a demanda por combustíveis no Brasil e em outras regiões do mundo. Em seguida, são descritos os métodos avaliados no presente estudo, sendo estes: Modelos Sarima (SAR) e Arima com variáveis binárias para sazonalidade (SAR.Bin), Vetor Autoregressivo com Mecanismo de Correção de Erro (VECM), Modelo de Decomposição de Séries Temporais por Regressão Local (STL), Modelos Lineares Dinâmicos (DLM), Redes Neurais Artificiais univariadas e multivariadas (ANN e MANN), Máquina de Vetor de Suporte (SVM), Hybrid Neural Fuzzy Inference System (HyFIS), Random Forest (RF) e modelos híbridos (HM). Por fim, são apresentados nos resultados uma comparação da performance de previsão segundo métricas especificadas na seção metodológica.

2. O Mercado de Combustíveis Brasileiro

O mercado brasileiro de combustíveis leves possui peculiaridades derivadas da complexidade tecnológica, produtiva e institucional que afetam por um lado, as decisões de consumo e, por outro, a dinâmica de produção e abastecimento.

No tocante a tecnologia, a inserção dos veículos flexíveis no início de 2003 deu ao consumidor brasileiro a oportunidade de decidir dentre o etanol e gasolina, ou uma mistura de ambos no momento do abastecimento. Essa mudança afetou a dinâmica de formação de preços e, conseqüentemente, do próprio consumo de combustíveis do país por tornar mais elástica a demanda pelos combustíveis individualmente (DE FREITAS; KANEKO, 2011).

Na esfera produtiva, devido ao uso da cana-de-açúcar como matéria-prima para o etanol brasileiro, a produção do combustível fica sujeita as condições apresentadas no mercado de açúcar e a intempereis climáticas. O abastecimento doméstico de gasolina, por sua vez, é realizado pelo refino de petróleo por parte da Petrobrás, empresa estatal com quase monopólio da atividade.



Por fim, no âmbito institucional, o mercado de combustíveis passou por inúmeras alterações ao longo da última década, incluindo mudanças na política de precificação dos derivados, alterações de tributos sobre os combustíveis e, mais recentemente, a aprovação da Lei nº 13.576 de dezembro de 2017, que instituiu a Política Nacional de Biocombustíveis no Brasil, *RenovaBio*¹.

O referido Programa entra em vigor em 2020 e marca uma mudança nas diretrizes da matriz energética de combustíveis no Brasil, estabelecendo metas de emissões de gases de efeito estufa que implicam na mudança dos padrões de consumo. Para 2028, as metas já estabelecidas preveem a necessidade de expansão da produção de etanol dos atuais 30 bilhões de litros para 47 bilhões de litros (Comitê *RenovaBio*, 2018; CNPE, 2018)

Nesse contexto, o dimensionamento do consumo doméstico por combustíveis para fins de segurança de abastecimento e, ainda, cumprimento das metas estabelecidas pelo *RenovaBio*, torna-se essencial para o gerenciamento de recursos tanto por parte do setor público (necessidade de fiscalização, infraestrutura para escoamento da produção, etc) quanto para o setor privado (aumento da capacidade produtiva, decisão do mix das usinas e, agora, decisão do mix das distribuidoras).

Assim, como ponto de partida para o presente estudo, foi realizada uma revisão dos trabalhos que avaliaram a demanda por combustíveis. A Tabela 1 apresenta uma síntese dos principais estudos que estimaram a demanda por um ou mais combustíveis líquidos, sumarizando as variáveis dependentes, o local em que o estudo foi realizado, a metodologia e variáveis explicativas.

Tabela 1: Estudos que estimaram a demanda por combustíveis no Brasil e no mundo.

Autores	Local	Método	Variável dependente	Variáveis explicativas
Nasr, Badr e Joun (2003)	Líbano	ANN e Multivariate ANN	Consumo de Gasolina	Preço da gasolina, frota de carros registrados e valores defasados
Zanini, Souza e Pedreira (2001)	Brasil	Modelo híbrido com redes neurais e regressão dinâmica	Consumo de Gasolina	Preço da gasolina, renda e valores defasados
Burnquist e Bacchi (2000)	Brasil	Modelo de cointegração (VECM)	Consumo de Gasolina	Preço da gasolina, renda e valores defasados
Moreira (1996)	Brasil	Modelo de co-integração (VECM)	Consumo de Óleo Diesel	Consumo per capita, PIB per capita e preço do óleo
Figueira e Burnquist (2006)	EUA	Modelos Box e Jenkins (Sarima)	Consumo de Gasolina	Valores defasados
Castro (2012)	Brasil	Modelo de cointegração (VECM)	Consumo de Etanol e Gasolina	Preço dos combustíveis e dimensão da frota
Santiago, De Mattos e Perobelli (2011)	Brasil	Modelo híbrido com cointegração (VECM) e insumo-produto	Consumo de gasolina, óleo diesel, óleo combustível e etanol	Variáveis macroeconômicas

¹ O programa tem por objetivo estabelecer metas de emissões para as distribuidoras de combustíveis tendo em vista o nível de poluição relativa a cada combustível. De forma semelhante ao programa norte-americano *Renewable Fuel Standard* o atingimento das metas deverá ser realizado pela compra de títulos de valores mobiliários em uma espécie de mercado de créditos de carbono.



Johansson e Schipper (1997)	Países da OCDE	Regressão com dados em painel utilizando efeitos fixos e aleatórios	Consumo de gasolina, diesel, gás liquefeito de petróleo e gás natural comprimido	Preço dos combustíveis, renda, impostos sobre combustíveis, densidade populacional e frota de carros estimada.
Al-Ghandoor <i>et al.</i> (2009)	Jordan	Regressão multivariada	Consumo de combustíveis fósseis e energia elétrica	Preço dos combustíveis e energia elétrica, renda, população e condições climáticas
Smith, Hall e Mabey (1995)	Países da OCDE	Modelo de cointegração (VEC) com adaptação para endogeneizar o progresso técnico	Consumo de combustíveis fósseis	Preço dos combustíveis, PIB e valores defasados
Ediger e Akar (2007)	Turquia	Modelos Box e Jenkins (Sarima)	Consumo de combustíveis fósseis e outras fontes energéticas	Valores defasados
Ediger, Akar e Ugurlu (2006)	Turquia	Modelos Box e Jenkins (Sarima) e regressão simples	Consumo de combustíveis fósseis	Valores defasados
Mu <i>et al.</i> (2004)	China	Modelos de regressão	Consumo de biocombustíveis de áreas rurais	Renda, população da área de plantio, nível educacional, etc.
Islas, Manzini e Masera (2006)	México	Long-range energy alternatives planning system (LEAP) model	Consumo de combustíveis e biocombustíveis	População, PIB, demanda por energia elétrica e fontes disponíveis para transformação em energia
Melikoglu (2014)	Turquia	Modelos semi-empíricos	Consumo de gasolina, diesel, biodiesel, etanol e gás liquefeito de petróleo.	Variável tempo

Os estudos realizados para o mercado brasileiro limitam-se nos seguintes aspectos: i) são essencialmente concentrados no uso de metodologias tradicionais de séries temporais; ii) não buscam verificar se é possível estimar um modelo acurado para a previsão de consumo de combustíveis (no geral, se centram na avaliação do encaixe do modelo nos dados de treinamento e na significância das elasticidades); iii) não testam de forma dinâmica diversas metodologias em uma mesma base de dados para avaliar qual delas se mostra superior na realização de previsões.

Em síntese, os trabalhos que avaliaram a demanda por combustíveis não se dedicaram à realização de previsões. Nesse contexto, existe na literatura inúmeros estudos que estruturaram modelos para a previsão do consumo por energia elétrica. Além de ser um campo correlato ao consumo energético de combustíveis, as séries adotadas nessas pesquisas possuem características semelhantes aquela utilizada no presente trabalho. Desse modo, esses estudos foram avaliados para identificar os modelos mais indicados para a análise aqui proposta.

Na revisão desses trabalhos, as seguintes técnicas de projeção foram as mais recorrentes: modelos de Lógica Fuzzy (CHANG; FAN; LIN, 2011; KUCUKALI; BARIS, 2010; PADMAKUMARI; MOHANDAS; THIRUVENGADAM, 1999; PAI, 2006), Máquinas de Vetores de Suporte (AHMAD *et al.*, 2014; CHE; WANG, 2014; CHEN *et al.*, 2017; DEBNATH; MOURSHED, 2018; FAN; CHEN; LEE, 2008; GARCÍA NIETO *et al.*, 2018; HONG, 2009), Random Forest (DUDEK, 2015; IBARRA-BERASTEGI *et al.*, 2015; LAHOUAR; BEN HADJ SLAMA, 2015; M.J. *et al.*, 2014) e modelos híbridos (BABU;



REDDY, 2014; BARAK; SADEGH, 2016; MAIA; DE CARVALHO; LUDERMIR, 2008; NENGLING; STENZEL; HONGXIAO, 2006; NIE *et al.*, 2011; WEI; CHEN; LEE, 2013).

Embora pouco explorado para previsão do consumo de combustíveis, o modelo de Redes Neurais Artificiais têm sido recorrentemente adotados para a previsão da demanda de diversas fontes energética ou mesmo de outros mercados (ver, por exemplo, EBRAHIMPOUR *et al.*, 2011; KADIR KAVAKLIOGLU, HALIM CEYLAN, HARUN KEMAL OZTURK, 2009; KANDANANOND, 2011; KHASHEI; BIJARI, 2010; RINGWOOD; BOFELLI; MURRAY, 2001; SÖZEN; ARCAKLIOĞLU; ÖZKAYMAK, 2005; XIA; WANG; MCMENEMY, 2010). Outras técnicas de previsão de demanda também foram identificadas a partir dos trabalhos de Deb *et al.* (2017), Debnath e Mourshed (2018) e Suganthi e Samuel (2012).

3. Metodologia

3.1. Fonte de dados e pré-tratamento

A base de dados utilizada é composta pela soma do consumo de gasolina, etanol e GNV convertidos em litros de gasolina equivalente, baseada na eficiência energética de cada combustível. Essa conversão é necessária para considerar os rendimentos distintos de cada combustível na composição da demanda (RODRIGUES; BACCHI, 2017). Os dados foram obtidos a partir de publicações da ANP (2019) e da ABEGÁS (2019). A série, que se inicia em agosto de 2002 e se encerra em agosto de 2018, pode ser visualizada na Figura 1.

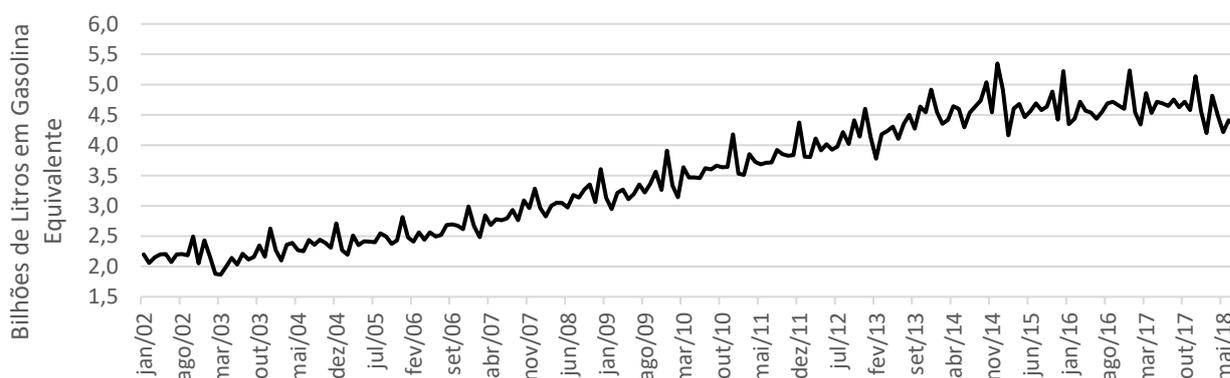


Figura 1: Consumo de Etanol Hidratado, Gasolina e GNV em litros de gasolina equivalente do período de agosto de 2002 a agosto de 2018.

Para o modelo de cointegração e ANN multivariado, outras duas séries foram tomadas como endógenas e projetadas em conjunto ao consumo de ciclo Otto. A primeira consiste no PIB *per capita* estimado a partir do PIB mensal divulgado pelo BACEN (2019) e da população residente pelo IBGE (2019). A segunda é preço médio ao consumidor final ponderado pelo volume de consumo dos combustíveis (ANP, 2019) como uma *proxy* do preço médio pago pelo consumo de combustíveis do ciclo Otto. Essas séries são as mais utilizadas em estudos que buscaram explicar a demanda por combustíveis, como pode ser observado na Tabela 1.

3.2. Determinação de hiperparâmetros dos modelos

Cada modelo possui um conjunto de hiperparâmetros que deve ser determinado *a priori* e outro conjunto de parâmetros estimados por um processo de otimização pautada em algum



procedimento estatístico paramétrico ou não-paramétrico. A Tabela 2 apresenta cada um desses elementos para os diferentes modelos adotados.

Tabela 2: Hiperparâmetros a serem determinados *a priori* em cada modelo.

Modelo	Hiperparâmetros
Sarima	Número de termos autoregressivos (p), médias móveis (q), ordem de integração (d), número de termos autoregressivos sazonais (P), médias móveis sazonais (Q) e ordem de integração sazonal (D).
Arima ^{c/} Binárias	Número de termos autoregressivos (p), médias móveis (q), ordem de integração (d)
VECM	Número de termos autoregressivos (p) e vetores de cointegração (r).
Modelo Estrutural	Fator de desconto para estimação das matrizes de covariâncias (δ).
STL	Número de passagens pelo <i>loop</i> interno $n_{(i)}$, número de iterações do <i>loop</i> externo $n_{(o)}$, parâmetros que suavizam o componente de baixa frequência presente no componente sazonal $n_{(l)}$ e d_l , parâmetros que suavizam o componente de tendência $n_{(T)}$ e d_T , parâmetros que suavizam o componente sazonal $n_{(s)}$ e d_s .
ANN	Número de termos autoregressivos (p), número de neurônios na primeira e segunda camada oculta (k_1, k_2).
Multi ANN	Número de termos autoregressivos (p) das variáveis de saída, i.e, consumo ciclo Otto, preço médio dos combustíveis e PIB <i>per capita</i> , número de neurônios na primeira e segunda camada oculta (k_1, k_2).
SVM	Número de termos autoregressivos (p), erro insensível (ϵ), constante C e o kernel para transformação do espaço de entrada no espaço de características.
Hyfis	Número de termos autoregressivos (p), número de características do universo de discurso (L).
Random Forest	Número de termos autoregressivos (p), número de árvores (B), número de variáveis explicativas utilizadas por árvore (m) e número mínimo de observações em um nó terminal (n_{min}^B).
Modelos Híbridos	Número de termos autoregressivos (p_L), médias móveis (q) e ordem de integração (d) do modelo da componente linear e número de termos autoregressivos (p_{NL}), erro insensível (ϵ), constante C e o kernel para transformação do espaço de entrada no espaço de características do modelo da componente não linear.

A determinação do conjunto de hiperparâmetros de um modelo deve ocorrer pela minimização da estimativa do erro de generalização ou erro de previsão (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Para modelos lineares, uma forma simples de avaliar a capacidade de generalização consiste na comparação dos critérios de informação de Akaike (AIC) ou Bayesiano (BIC) (HASTIE; TIBSHIRANI; FRIEDMAN, 2009; HYNDMAN; ATHANASOPOULOS, 2018). Esses são os casos dos modelos Sarima, VEC (os quais ainda possuem outros procedimentos descritos nas próximas sessões) e STL. No presente estudo, o critério utilizado foi o AIC.

No entanto, quando se tratam de modelos não-lineares, recorre-se a formas empíricas de estimação do erro de generalização. A mais comum destas é a validação cruzada (HASTIE; TIBSHIRANI; FRIEDMAN, 2009; KUHN; JOHNSON, 2013). A validação cruzada envolvendo séries temporais é diferente dos problemas envolvendo dados em painel devido a dinâmica autoregressiva da série (BERGMEIR; HYNDMAN; KOO, 2018; HYNDMAN; ATHANASOPOULOS, 2018). Seguindo lógica semelhante a proposta por Hyndman e Athanasopoulos (2018), o procedimento foi estruturado da seguinte forma:

1. Um conjunto de parâmetros α_i é selecionado inicialmente para o modelo.



- Um corte inicial é realizado na série de consumo de ciclo Otto restando apenas os dados de janeiro de 2002 a dezembro de 2011, i.e, 120 observações do total de 200. A escolha desse ponto de corte se deve a mudança de tendência na série percebida a partir de 2013.
- Essa subsérie é utilizada para estimativa dos modelos e realização das previsões (\hat{Y}) h passos-à-frente. Nesse caso, o vetor de erro das previsões é calculado e armazenado.
- Uma nova observação é adicionada na base de treino e o passo 3 é repetido.
- Os passos 3 e 4 são repetidos até que a previsão h passos-à-frente tenha como último período previsto a última observação conhecida da série. Assim, o passo 3 e 4 são realizados $81 - h$ vezes e este mesmo número de vetores de erros são gerados, cada um contendo h erros de previsão. A raiz quadrada do erro quadrático médio (RMSE) é calculada para o modelo com parâmetros α_i , conforme a Eq. 1.

$$RMSE_i = \sqrt{\frac{1}{(81-h)h} \sum_{j=1}^{(81-h)h} (Y_j - \hat{Y}_j)^2} \quad (1)$$

Com os erros de previsão de todas as subséries obtidas, é possível estimar métricas de erro de generalização do modelo e selecionar aquele com melhor desempenho. Esse procedimento foi realizado e a melhor estrutura foi selecionada para previsões (h) de 3, 6 e 12 meses à frente. Essa lógica, além de evitar super encaixe do modelo aos dados de treinamento (*overfitting*) será utilizada como base para comparação dos algoritmos, conforme descrito da seção 3.4.

Os trabalhos de Nelson *et al.* (1999) e Zhang e Qi (2005) concluíram que alguns algoritmos podem apresentar dificuldades no tratamento de séries sazonais. Visando evitar essa complicação, três diferentes formas de tratamento da sazonalidade são avaliadas. Na primeira, a série dessazonalizada é utilizada para treinamento do modelo, sendo a dessazonalização realizada por meio de um filtro de médias móveis. Nas duas últimas abordagens, a modelagem e previsão do componente sazonal é realizado pelo próprio modelo, seja pela inclusão de um termo autoregressivo sazonal ou pela inclusão de regressores binários para cada um dos meses.

O mesmo é válido para o pré-tratamento dos dados exigido por alguns algoritmos. As formas de pré-processamento compreendem tomar os dados padronizados pela média e desvio-padrão ou em logaritmo para realizar o treinamento dos modelos. O pré-processamento e forma de tratamento da sazonalidade de dados selecionado para os modelos ANN, SVM, HyFis e Random Forest também foi selecionado a partir do procedimento de validação cruzada.

3.3. Métodos de Previsão Implementados

3.3.1. Modelo Sazonal Autoregressivo Integrado com Médias Móveis (Sarima)

O Sarima $(p,d,q)(P,D,Q)$, ou SAR, é um método muito popular de modelagem paramétrica. A ideia central consiste em ajustar um modelo capaz de explicar uma série temporal estacionária por meio dos valores passados da própria série e dos erros do modelo (BOX; JENKINS; REINSEL, 2008). Além da abordagem tradicional, que utiliza as próprias defasagens da série e dos erros para incorporação dos efeitos da sazonalidade, este estudo propõe uma abordagem alternativa utilizando regressores binários (SAR.Bin). Para determinação da ordem de integração da série, foram realizados os testes de raiz unitária KPSS



(KWIATKOWSKI *et al.*, 1992) e DF-GLS (ELLIOTT; ROTHENBERG; STOCK, 1996) para uma raiz, além do teste DP (DICKEY; PANTULA, 1987) para duas raízes. Adicionalmente, foi implementado o teste de raiz unitária sazonal proposto por Osborn *et al.* (1988).

O diagnóstico para seleção final utiliza-se do critério de informação AIC e dos testes de autocorrelação de Ljung-Box, ARCH-LM para heterocedasticidade e o teste de Jarque-Bera (BUENO, 2018) e Shapiro-Wilk (SHAPIRO; WILK, 1965) para a normalidade dos resíduos. Com exceção dos testes de normalidade dos resíduos, os diagnósticos realizados demonstram boa aderência dos modelos.

3.3.2. Vetor Autoregressivo com Mecanismo de Correção de Erro (VECM)

O Vetor Autoregressivo com Mecanismo de Correção de Erro (VECM) consiste em uma abordagem alternativa para previsão de séries temporais não estacionárias (LÜTKEPOHL, 2005). Este estima um sistema de equações cujas variáveis possuem uma trajetória comum de longo prazo, isto é, são cointegradas. Para isso, introduz o termo chamado vetor de correção de erros em um Vetor Autoregressivo (VAR).

O teste de cointegração entre as variáveis e a estimativa dos parâmetros do modelo foram conduzidos conforme proposto por Johansen (1988, 1991). O teste de máximo autovalor foi utilizado para determinação do número de vetores de cointegração. A seleção dos termos autoregressivos foi baseada no critério de informação AIC e a verificação do modelo se apoiou nos testes de autocorrelação de Portmanteau (LÜTKEPOHL, 2005) e de Edgerton e Shukur (1999), nos testes de normalidade multivariada de Jarque-Bera (LÜTKEPOHL, 2005) e Shapiro-Wilk (SHAPIRO; WILK, 1965) e no teste de heterocedasticidade multivariada ARCH-LM (ENGLE, 1982). Do mesmo modo que no modelo Sarima, os testes de diagnóstico corroboram para conclusão de que o modelo possui um bom ajuste aos dados, com exceção do teste de normalidade dos resíduos.

3.3.3. Modelos Lineares Dinâmicos (Modelo Estrutural)

O modelo linear dinâmico (DLM) é um caso especial dos modelos de espaço-estado. Neste último, as variáveis dependentes são resultado de um sistema dinâmico perturbado por um erro aleatório em seu estado. Os problemas de estimação e previsão são resolvidos computando recursivamente a distribuição a posteriori das variáveis de interesse dada a informação disponível. Nesse sentido, eles são tratados segundo uma abordagem bayesiana (WEST; HARRISON, 1997). Um caso específico do DLM, utilizado neste estudo, é chamado modelo estrutural para séries temporais (HARVEY, ANDREW C.; SHEPHARD, 1993). Neste caso, a série é decomposta em termos da sua tendência, sazonalidade, componente cíclico e aleatório.

Os parâmetros foram estimados utilizando um procedimento de inferência bayesiana conjugada proposta por West e Harrison (1997, p. 199–202). Esse procedimento adapta a estimação padrão executada por meio do filtro de Kalman, introduzindo um mecanismo de fatores de desconto para estimação das matrizes de covariâncias do ruído da equação de observação e de estado. A determinação do fator de desconto foi realizada utilizando o procedimento de validação cruzada descrito na seção 3.4, pois, embora o modelo seja linear nos coeficientes da equação de previsão, ele não é linear em relação ao fator de desconto.



3.3.4. Seasonal e Trend decomposition using Loess (STL)

O método STL (*Seasonal-Trend Decomposition Procedure Based on Loess*) foi desenvolvido por Cleveland *et al.*, (1990) e tem o objetivo de decompor uma série temporal (Y_t) em três componentes: tendência (T_t), sazonalidade (S_t) e resíduo (ϵ_t). O STL parte de uma condição inicial arbitrária para os componentes da série e por meio de regressões locais as estimam iterativamente. O componente sazonal é previsto assumindo-o igual ao último componente estimado pelo processo iterativo anteriormente descrito. A projeção da série dessazonalizada ocorre por qualquer método de previsão não sazonal (HYNDMAN; ATHANASOPOULOS, 2018). Neste caso foi utilizado um modelo ETS, uma metodologia também baseada na decomposição da série a partir de um modelo de espaço-estado. A seleção do melhor modelo foi baseada no critério de informação AIC.

O STL é robusto a *outliers* e capaz de tratar qualquer tipo de sazonalidade, visto que o componente sazonal e de tendência podem se alterar ao longo do tempo e a taxa de mudança pode ser controlada pelo analista (HYNDMAN; ATHANASOPOULOS, 2018).

3.3.5. Redes Neurais Artificiais (ANN)

O método de Redes Neurais Artificiais (ANN) é baseado em dinâmica que remete ao funcionamento do sistema nervoso humano. O modelo é representado por uma rede de elos orientados que são interligados em certos pontos chamados de nós ou neurônios. Quando os valores defasados de uma série são utilizados como variáveis de entrada em uma ANN com estrutura de um perceptron de múltiplas camadas, esta pode ser chamada *time lagged feedforward network-TLFN* (SIMON HAYKIN, 2001). Nessa estrutura, cada neurônio de cada camada possui um conjunto de vetor de coeficientes que ponderam as variáveis de entrada daquele nó. No presente estudo, o número de camadas é igual a 3 e o número de neurônios da camada de saída é igual ao número de variáveis endógenas (m) do modelo.

A inferência quanto aos pesos ocorre pela minimização do erro quadrático médio dos nós na camada de saída. Para tanto, o algoritmo BFGS foi utilizado (SIMON HAYKIN, 2001). Neste estudo, duas estruturas de ANN foram estimadas. O primeiro é um modelo univariado ($m = 1$) cuja variável de saída é o consumo de combustíveis do ciclo Otto e p termos autoregressivos são as variáveis de entrada. O segundo é um modelo multivariado ($m = 3$), cuja saídas são o consumo de ciclo Otto, o preço médio dos combustíveis e o PIB *per capita* previstos por meio de p termos autoregressivos de cada uma dessas variáveis como vetor de entrada.

Os modelos ANN são capazes de realizar um mapeamento não-linear entre os vetores de entrada e saída, contudo, um número de interações excessivas no processo de treinamento pode causar *overfitting* do modelo. Buscando evitar este último, além da própria validação cruzada para determinação de hiperparâmetros, um procedimento de parada antecipada foi aplicado, conforme descrito por Simon Haykin (2001).

3.3.6. Máquinas de Vetores de Suporte (SVM)

Máquina de Vetor de Suporte (SVM) é um modelo não paramétrico capaz de estimar relações não lineares de um vetor de entrada para uma variável de saída, utilizando o chamado truque de kernel (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Este consiste em aplicar nas



variáveis de entrada $\{\{x_{i,t}\}_{i=1}^{m_0}\}_{t=1}^n$ uma transformação não-linear $\{\{\varphi_j(\mathbf{x}_t)\}_{j=1}^{m_1}\}_{t=1}^n$ de modo com que a dimensionalidade aumente de m_0 para m_1 . O objetivo do método é transformar uma relação não linear de $\mathbf{x}_t \rightarrow Y_t$ em uma relação linear de $\boldsymbol{\varphi}(\mathbf{x}_t) \rightarrow Y_t$.

A estimativa do vetor de pesos $\boldsymbol{\omega}$ que permita a construção do hiperplano $\hat{f}(\boldsymbol{\omega}, \mathbf{x}_t) = \boldsymbol{\omega}^T \boldsymbol{\varphi}(\mathbf{x}_t) = \sum_{j=0}^{m_1} w_j \varphi_j(\mathbf{x}_t)$ ocorre pela resolução de um problema de otimização quadrática. Uma etapa fundamental é a determinação de $K(\mathbf{x}_t, \mathbf{x}_j) = \sum_{j=0}^{m_1} \varphi_j(\mathbf{x}_t) \varphi_j(\mathbf{x}_j) = \boldsymbol{\varphi}^T(\mathbf{x}_t) \boldsymbol{\varphi}(\mathbf{x}_j)$ chamado função kernel. Na prática, para estimar um SVM não é necessário conhecer a função $\boldsymbol{\varphi}(\mathbf{x})$, apenas a função kernel e que esta atenda ao teorema de Mercer (SIMON HAYKIN, 2001, p. 366). Existem uma série de kernels previamente indicados pela teoria cuja escolha deve prezar pela melhor generalização do modelo. Assim, o procedimento de validação cruzada proposto no trabalho serviu como critério para escolha deste parâmetro do modelo.

O SVM possui boa capacidade de generalização (evita *overfitting*), possibilita mapeamento não linear dos vetores de entrada e saída e faz uso de programação quadrática para estimação dos parâmetros – esse procedimento é mais rápido na comparação com os métodos iterativos para problemas não lineares.

3.3.7. Hybrid Neural Fuzzy Inference System (HyFIS)

A teoria *fuzzy* se iniciou com o trabalho de Zadeh (1965), que visava introduzir um aspecto de incerteza em variáveis linguísticas. A lógica fuzzy, ao contrário da *booleana*, assume que um conjunto A do universo de discurso Ω é definido por uma função de pertinência $\mu_A: \Omega \rightarrow [0,1]$. A função $\mu_A(x)$ indica o grau de compatibilidade entre x e o conceito expresso por A . Os controles nebulosos são formas de estabelecer relações entre variáveis linguísticas a partir da modelagem de regras ligando uma condição (Se) a uma consequência (Então).

No presente estudo foi implementado o método Hybrid Neural Fuzzy Inference System-HyFIS (KIM; KASABOV, 1999). Neste caso, o sistema de controle fuzzy é estruturado como um perceptron de múltiplas camadas e o método de inferência é o de Mamdani (1974). O treinamento divide-se em 2 etapas: aprendizagem das regras, que busca encontrar, a partir dos dados, a base de regras do sistema fuzzy; e, aprendizagem de parâmetros, que busca identificar os parâmetros ótimos.

A etapa de aprendizagem de regras é realizada utilizando o procedimento proposto por Wang e Mendel (1992). Para a etapa de aprendizagem de parâmetros, minimiza-se o erro quadrático médio dos neurônios da camada de saída por meio do algoritmo de retropropagação de erro.

O modelo possibilita o mapeamento não linear dos vetores de entrada e saída e a simplicidade de interpretação das regras. No entanto, é exigido uma grande capacidade computacional para execução das etapas de aprendizado e também para determinação dos hiperparâmetros.

3.3.8. Random Forest

O modelo Random Forest pode ser entendido como um modelo de aprendizagem conjunta, o qual utiliza árvores de decisão para realização da previsão de uma variável de saída em função de um vetor de entrada. Neste caso, diversas árvores de regressão são estimadas e a



previsão final do algoritmo é dada pela média das previsões das árvores (BREIMAN, 2001; HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Os parâmetros de cada árvore são as variáveis de entrada selecionadas, os pontos de divisão ótimo de cada nó e as observações em cada nó terminal. A determinação desses parâmetros ocorre conforme o algoritmo CART (BREIMAN *et al.*, 1984).

O modelo Random Forest possui uma fácil interpretação e a procura por hiperparâmetros ótimos não requer menos esforço computacional na comparação com os outros algoritmos. Por ser uma espécie de modelo de aprendizagem conjunta, costuma gerar previsões com variância do erro relativamente menor.

3.3.9. Modelos Híbridos

Modelos híbridos são aqueles que combinam paradigmas de diferentes algoritmos de previsão para formulação de um modelo único. Neste caso, o processo de estimação dos parâmetros (aprendizagem) ocorre de forma conjunta. Comumente uma série pode apresentar atributos lineares e não-lineares simultaneamente. O uso de modelos híbridos permite capturar diferentes padrões dentro de uma mesma série (ZHANG, 2003). Uma aplicação comum de modelos híbridos para previsão de uma série Y_t consiste na divisão desta em dois componentes: um linear (L_t) e outro não-linear (N_t).

No presente estudo, a separação das componentes foi realizada baseado em Babu e Reddy (2014), com algumas adaptações. Utilizando uma decomposição STL, a série é separada em um componente de tendência, assumido como linear, e outro contendo o resíduo e sazonalidade da série, assumidos como não-lineares. O componente linear é modelado por um modelo Arima selecionado a partir de critérios de informação e testes descritos anteriormente. O componente não-linear é modelado por uma Máquina de Vetor de Suporte (SVM), em que os hiperparâmetros são determinados por procedimentos de validação cruzada. A previsão final é obtida pela soma das duas componentes estimadas.

3.4. Métricas de avaliação e testes estatísticos para comparação de algoritmos

Para avaliação dos diferentes métodos de previsão, um processo iterativo foi conduzido a fim de testar estatisticamente a diferença de acurácia dos modelos. As seguintes etapas foram executadas:

1. Um corte inicial é feito na série de consumo de ciclo Otto deixando apenas os dados de janeiro de 2002 a dezembro de 2011.
2. Essa subsérie é utilizada para estimativa dos modelos e realização da previsão h passos-à-frente. O RMSE é calculado, conforme a Eq. 4, e armazenado como medida de acurácia no período. Neste caso, cada período de corte possui um valor de RMSE para a previsão h passos-à-frente.
3. Uma nova observação é adicionada a base de treino e o passo 3 é repetido.

Os passos 3 e 4 são repetidos até que a previsão h passos-à-frente tenha como último período previsto a última observação conhecida da série. Assim obtém-se uma série com o RMSE para cada ponto no tempo e para cada modelo. As performances são ranqueadas para cada período e o teste de Friedman é realizado para verificar se existe diferença entre o ranking médio de um algoritmo i , dado por \bar{r}_i , e o desempenho médio de todos os modelos \bar{r} . Em



seguida, o teste de Nemenyi é realizado para comparação individual dos métodos. Este teste também é baseado no ranqueamento dos modelos, segundo uma métrica de erro, para cada período analisado. O método cujo o ranking médio for o menor será apontado como o de melhor performance. Para mais informações sobre os testes no contexto de comparação de métodos de previsão ver DEMSAR (2006) e SHINGALA (2016). As Eqs. 2 e 3 apresentam as hipóteses nulas e alternativas de cada teste.

$$\begin{array}{l} \text{Friedman} \\ H_0: \bar{r}_i = \bar{r}, \text{ não há diferença entre o ranking dos métodos no geral.} \\ H_A: \bar{r}_i \neq \bar{r}, \text{ há diferença entre o ranking dos métodos no geral.} \end{array} \quad (2)$$

$$\begin{array}{l} \text{Nemenyi} \\ H_0: \bar{r}_i = \bar{r}_j \forall i \neq j, \text{ não há diferença entre o ranking dos métodos } i \text{ e } j. \\ H_A: \bar{r}_i \neq \bar{r}_j \forall i \neq j, \text{ há diferença entre o ranking dos métodos } i \text{ e } j. \end{array} \quad (3)$$

Uma metodologia ingênua de previsão é proposta como comparativo com o intuito de verificar a superioridade dos métodos em relação a um modelo com baixo custo de implementação. Neste caso, os meses do último ano serão utilizados como previsão para um ano subsequente. Esse método é recomendado para dados com sazonalidade.

O RMSE foi utilizado como critério de seleção dos modelos no procedimento de validação cruzada uma vez que quase todos os métodos utilizam uma função de erro quadrático médio para estimativa do vetor de parâmetros. No entanto, será apresentado o *Mean Absolute Square Error* ou Erro Absoluto Percentual Médio (MAPE) para cada ponto no tempo (isto é, para a previsão h passos à frente a partir daquela data), calculado conforme a Eq. 5, com o objetivo de facilitar a interpretação dos resultados. Adicionalmente, o valor máximo e mínimo do Erro Percentual (EP) considerando o horizonte de previsão h passos à frente também é apresentado. A seguir estão as fórmulas que foram utilizadas no cálculo desses indicadores.

$$RMSE_t = \sqrt{\frac{1}{h} \sum_{j=1}^h (Y_j - \hat{Y}_j)^2} \quad (4) \quad MAPE_t = \frac{100\%}{h} \sum_{j=1}^h \frac{|Y_j - \hat{Y}_j|}{Y_j} \quad (5)$$

$$\max EP_t = \max \left\{ \frac{|Y_j - \hat{Y}_j|}{Y_j} \right\}_{j=1}^h \quad (6) \quad \min EP_t = \min \left\{ \frac{|Y_j - \hat{Y}_j|}{Y_j} \right\}_{j=1}^h \quad (7)$$

4. Resultados e Discussão

A Tabela 3 apresenta os resultados da escolha dos hiperparâmetros. Os modelos cujos parâmetros são escolhidos em função do horizonte de previsão apresentam uma quebra para 3, 6 e 12 meses.

Tabela 3: Hiperparâmetros selecionados a partir dos procedimentos de validação cruzada e critérios de informação.

Modelo	Tratamento da Sazonalidade	Pré-processamento dos Dados	Hiperparâmetros Ótimos	
Sarima	Termo autoregressivo sazonal	Logaritmo	$p = 2, d = 1, q = 3$ $P = 1, D = 0, Q = 1$	
Arima c/ Binárias	Variáveis binárias	Logaritmo	$p = 5, d = 1, q = 5$	
VEC	Variáveis binárias	Logaritmo	$p = 5, r = 2$	
Modelo Estrutural	Variáveis binárias	Logaritmo	3 meses	$\delta = 0,81$
			6 meses	$\delta = 0,82$
			12 meses	$\delta = 0,83$



Modelo	Processamento	Transformação	Horizonte	Parâmetros
STL	Separação pelo próprio STL	Logaritmo		$n_{(p)} = 12, n_{(l)} = 1, n_{(o)} = 15, n_{(t)} = n_{(s)} = 13, n_{(T)} = 21, d_l = d_T = 1, d_s = 0$
ANN	Filtro de Médias Móveis	Padronização pela média e desvio-padrão	3 meses	$p = 3, k_1 = 1, k_2 = 2$
			6 meses	$p = 5, k_1 = 1, k_2 = 5$
			12 meses	$p = 4, k_1 = 1, k_2 = 2$
Multi ANN	3 meses	Filtro de Médias Móveis	3 meses	$p = 4, k_1 = 10, k_2 = 10$
	6 meses	Filtro de Médias Móveis	6 meses	$p = 3, k_1 = 4, k_2 = 0$
	12 meses	Variáveis binárias	12 meses	$p = 3, k_1 = 10, k_2 = 9$
SVM	Filtro de Médias Móveis	Padronização pela média e desvio-padrão	3 meses	$p = 3, \epsilon = 0,1, C = 6, \text{kernel}=\text{gaussiano}$
			6 meses	$p = 3, \epsilon = 0,1, C = 10, \text{kernel}=\text{gaussiano}$
			12 meses	$p = 3, \epsilon = 0,1, C = 10, \text{kernel}=\text{gaussiano}$
Hyfis	Filtro de Médias Móveis	Padronização pela média e desvio-padrão	3 meses	$p = 3, L = 22$
			6 meses	$p = 3, L = 22$
			12 meses	$p = 3, L = 16$
Random Forest	Filtro de Médias Móveis	Sem transformação	3 meses	$p = 6, B = 100, m = 5, n_{min}^B = 6$
			6 meses	$p = 6, B = 100, m = 4, n_{min}^B = 3$
			12 meses	$p = 6, B = 100, m = 5, n_{min}^B = 5$
Modelos Híbridos	Separação pelo STL	Logaritmo	3 meses	$p_L = 3, d = 1, q = 0, p_{NL} = 2, \epsilon = 0,1, C = 1, \text{kernel}=\text{gaussiano}$
			6 meses	$p_L = 3, d = 1, q = 0, p_{NL} = 1, \epsilon = 0,9, C = 1, \text{kernel}=\text{gaussiano}$
			12 meses	$p_L = 3, d = 1, q = 0, p_{NL} = 1, \epsilon = 0,1, C = 10, \text{kernel}=\text{gaussiano}$

A Tabela 4, por sua vez, traz os rankings médios de cada método e, concomitantemente, o resultado do teste de Friedman. Nesse teste, a hipótese nula é dada pela homogeneidade do desempenho dos modelos, avaliando se os rankings individuais são estatisticamente iguais ao ranking médio.

Foram avaliadas as previsões dos 11 métodos e do método ingênuo. Essas previsões envolveram 69 pontos no tempo para o cenário de 12 meses, 75 pontos no caso de 6 meses e 78 pontos na projeção de 3 meses.

Tabela 4: Ranks médios dos modelos e resultado do teste de Friedman com hipótese alternativa de que a performance dos algoritmos é estatisticamente diferente da performance média.

Modelo	Rank Médio (12 meses)	Rank Médio (6 meses)	Rank Médio (3 meses)
SVM	4,33	4,87	5,49
SAR.Bin	4,46	5,04	5,03
VECM	5,59	5,24	4,92
ANN	5,84	5,21	5,19
RF	5,61	5,85	6,67



HYFIS	5,88	6,37	6,40
SAR	6,16	6,11	5,79
STL	6,74	6,52	6,44
HM	7,17	8,00	7,85
DLM	7,57	8,00	7,73
Seas.Naive	9,26	9,93	9,74
MANN	9,36	6,85	6,76
Média Geral	6,50	6,50	6,50
p-valor Teste Friedman	0,00	0,00	0,00
(Ha: Desempenhos diferentes)			

Na sequência, o resultado do teste de Nemenyi é apresentado nas Figuras 2 a 4 visando a comparação individual dos modelos. Os valores mostram que os métodos SVM, SAR.Bin, VECM e ANN se alternam como os 3 melhores modelos. O melhor algoritmo do ponto de vista do teste foi o SVM para a previsão de 12 e 6 meses. No entanto, considerando um intervalo de confiança de 95%, não é possível afirmar que os primeiros 7 modelos com melhor *ranking* possuem performance estatisticamente inferiores ao SVM na previsão de 12 meses. O mesmo é válido na previsão 6 e 3 meses à frente. O intervalo de confiança é representado nas Figuras 2 a 4 pela área cinza.

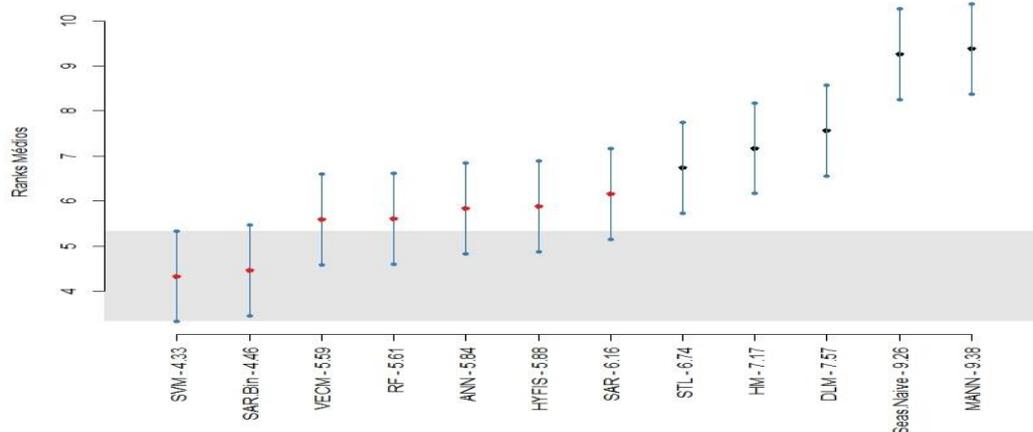


Figura 2: Resultado e intervalo de confiança do teste de Nemenyi para previsão 12 meses à frente.

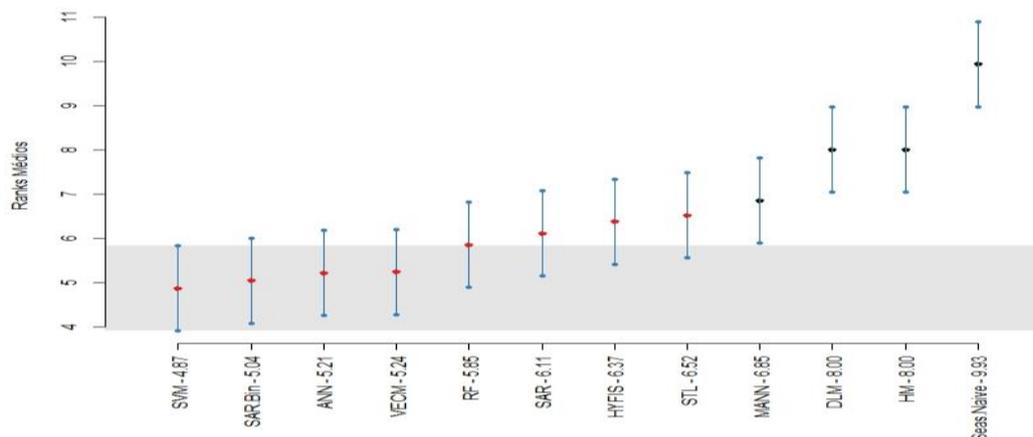


Figura 3: Resultado e intervalo de confiança do teste de Nemenyi para previsão 6 meses à frente.

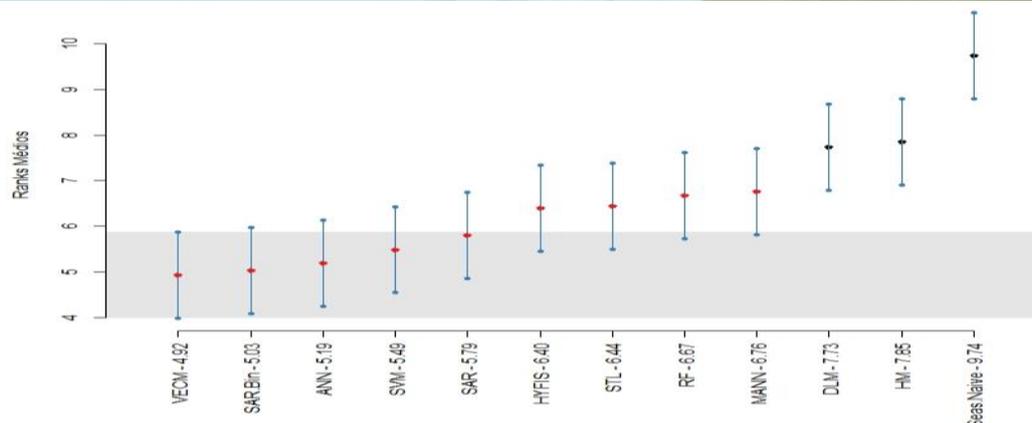


Figura 4: Resultado e intervalo de confiança do teste de Nemenyi para previsão 3 meses à frente.

A performance dos métodos mais bem ranqueados foram estatisticamente superiores ao modelo ingênuo. Esse fato evidencia a robustez dos procedimentos, justificando o maior custo computacional necessário para implementação dos mesmos.

Para uma comparação mais objetiva do desempenho dos modelos, nas Tabelas 5 a 7 são apresentadas a média e desvio-padrão do MAPE da previsão, a média do MAPE da amostra de treinamento, o valor máximo e mínimo do MAPE e o máximo do erro percentual de cada modelo considerando todos os períodos utilizados na validação.

Tabela 5: Resumo da performance dos métodos para previsão de 12 meses à frente.

Modelo	MAPE Médio	MAPE Médio (Treinamento)	Desvio Padrão MAPE	Max MAPE	Min MAPE	Max EP
SVM	2,78%	2,12%	1,11%	5,67%	1,47%	11,61%
SAR.Bin	2,89%	1,92%	1,44%	8,40%	1,22%	16,21%
VECM	2,91%	1,99%	1,21%	7,31%	1,34%	16,78%
RF	3,05%	2,66%	1,33%	6,43%	1,71%	12,03%
ANN	3,25%	2,21%	1,48%	7,24%	1,52%	13,24%
HYFIS	3,26%	2,98%	1,50%	7,14%	1,21%	12,81%
SAR	3,32%	2,15%	1,67%	8,58%	1,25%	19,71%
DLM	3,39%	3,35%	1,47%	8,55%	1,85%	20,70%
STL	3,41%	2,42%	1,53%	8,93%	1,31%	16,36%
HM	3,74%	1,76%	2,42%	16,80%	1,89%	34,34%
Seas Naive	4,62%	-	1,71%	7,37%	1,98%	13,23%
MANN	4,71%	2,31%	1,87%	9,34%	1,35%	21,58%

Tabela 6: Resumo da performance dos métodos para previsão de 6 meses à frente.

Modelo	MAPE Médio	MAPE Médio (Treinamento)	Desvio Padrão MAPE	Max MAPE	Min MAPE	Max EP
SAR.Bin	2,45%	1,92%	0,99%	6,07%	0,97%	11,43%
SVM	2,47%	2,12%	0,81%	4,48%	1,11%	8,94%
VECM	2,53%	1,99%	1,01%	5,17%	0,79%	10,10%
ANN	2,54%	2,20%	0,93%	5,85%	0,69%	9,45%
RF	2,66%	2,64%	0,96%	5,34%	1,20%	9,73%
SAR	2,68%	2,15%	1,20%	6,29%	1,06%	13,81%
HYFIS	2,74%	2,67%	1,02%	5,34%	0,90%	10,22%
STL	2,84%	2,42%	1,04%	6,68%	1,17%	9,95%
DLM	3,14%	3,35%	1,34%	7,67%	0,96%	16,72%
MANN	3,33%	3,25%	1,99%	12,56%	2,00%	18,42%
HM	3,54%	2,39%	1,73%	11,00%	1,23%	17,75%
Seas Naive	4,65%	-	1,94%	8,64%	1,24%	13,23%



Tabela 7: Resumo da performance dos métodos para previsão de 3 meses à frente.

Modelo	MAPE Médio	MAPE Médio (Treinamento)	Desvio Padrão MAPE	Max MAPE	Min MAPE	Max EP
SAR.Bin	2,28%	1,92%	1,17%	6,06%	0,49%	11,79%
VECM	2,28%	1,99%	1,23%	5,63%	0,64%	9,19%
SVM	2,35%	2,12%	1,03%	5,54%	0,68%	7,97%
ANN	2,40%	2,20%	1,09%	5,11%	0,51%	8,05%
SAR	2,42%	2,15%	1,29%	7,31%	0,67%	13,20%
HYFIS	2,56%	2,66%	1,18%	5,45%	0,70%	9,46%
RF	2,58%	2,65%	1,17%	5,68%	0,54%	9,44%
STL	2,65%	2,42%	1,10%	6,72%	0,66%	9,16%
MANN	2,77%	3,26%	1,29%	6,88%	0,79%	10,93%
DLM	3,06%	3,35%	1,57%	7,67%	0,43%	13,76%
HM	3,24%	2,23%	1,82%	9,52%	0,73%	12,87%
Seas Naive	4,72%	-	2,26%	9,54%	0,95%	13,23%

As Tabelas 5 a 7 acabam por refletir os resultados observados no teste de Nemenyi. Os modelos SVM, SAR.Bin e VECM lideram com o menor erro de previsão. Importante destacar que os modelos com menor MAPE médio não necessariamente serão aqueles indicados como superiores pelo teste de Nemenyi, pois este é realizado a partir do ranqueamento dos métodos nos períodos avaliados.

Com exceção do HM e do MANN para 12 meses, aparentemente nenhum método possui problema de *overfitting*, uma vez que os erros de treinamento e validação não destoam de forma considerável. Esse fato é importante para garantir a eficiência do processo de escolha dos hiperparâmetros e garantir da generalização fora da amostra de treino.

O aumento do horizonte de previsão implica em um aumento do MAPE Médio dos modelos que possuem maior acurácia na previsão de 3 meses, seguido da previsão de 6 e então 12 meses. Consequentemente, o MAPE e erros percentuais máximos se reduzem gradativamente conforme o horizonte de previsão também é reduzido. Nota-se, ainda, que o aumento do número de períodos previstos acarreta em uma maior heterogeneidade dos *rankings*, visto que há um aumento considerável do ranking médio do primeiro para o quarto colocado para previsão de 12 meses. O SVM distancia-se dos outros métodos conforme o horizonte de previsão aumenta.

As previsões realizadas para períodos contemplando os meses de fevereiro de 2014 a agosto de 2016 e abril de 2017 até agosto de 2018 evidenciam um aumento generalizado do erro de previsão de todos os métodos. Em ambos os períodos citados foram observados choques ocasionados pelo início da crise econômica no Brasil e impeachment da presidente Dilma Rousseff, bem como a paralização dos caminhoneiros observada no segundo período.

Buscando avaliar a estabilidade do erro de previsão, pode-se analisar o desvio padrão do MAPE e o máximo erro percentual dentro de um ponto no tempo. O SVM, cujo *ranking* de performance foi o melhor dentre todos os métodos na previsão de 12 e 6 meses, também obteve o menor desvio padrão do MAPE considerando todos os horizontes. Considerando o maior horizonte de previsão, o erro percentual máximo desse modelo foi de 11,6% e o MAPE máximo de 5,7%. Por outro lado, o modelo SAR.Bin, cujo ranking foi o segundo melhor, possui o quinto maior desvio padrão, erro percentual máximo de 16,2% e MAPE máximo de 8,4%.

Vale destacar que o aumento do horizonte de previsão, em linhas gerais, não acarretou em aumento do desvio-padrão do MAPE. Isto é, os modelos não passaram a ter performance mais instável em relação a sua média apenas pela quantidade de períodos previstos. Um modelo com instabilidade de performance ou com previsões que diferem substancialmente do valor real da série em alguns momentos pode vir a ser descartado mesmo quando seu MAPE médio é inferior a outros métodos. A decisão quanto a tolerância máxima deve ser baseada no custo esperado da decisão utilizando uma previsão incorreta.



5. Considerações Finais

O presente estudo comparou diferentes métodos utilizados na previsão do consumo de combustíveis de ciclo Otto no Brasil. O procedimento de validação cruzada permitiu estimar o erro de generalização dos modelos e o MAPE médio obtido foi inferior a 3% em todos os horizontes de previsão. Adicionalmente, os métodos com melhor desempenho se mostraram significativamente superiores ao modelo ingênuo.

Mesmo em períodos de grande incerteza, o desempenho dos modelos se mostrou estável, com erro máximo mensal chegando a 11,61% para previsão de 12 meses e inferior a 10% para 6 e 3 meses.

Em síntese, é possível concluir que, a despeito das mudanças tecnológicas, de mercado e institucionais observadas no setor de combustíveis leves nos últimos 10 anos, os modelos de previsão avaliados indicam que é possível realizar projeções para o consumo de combustíveis leves com bom nível de acurácia.

Esses resultados são especialmente relevantes para o planejamento dos agentes privados. Isso porque, é comum haver cláusula contratual enunciando multas quando existem alterações no volume de combustível comprado ou vendido em relação aquele estabelecido inicialmente em contrato, seja nas operações domésticas ou na importação desses produtos.

A projeção do consumo total de combustíveis leves também é relevante para esses agentes na definição de ações operacionais e de investimento relacionadas à produção, distribuição, transporte e revenda de combustíveis no País.

No caso do setor público, os modelos implementados podem ser úteis no dimensionamento dos recursos para ajuste na capacidade de produção doméstica, bem como de infraestrutura logística necessária para o suprimento da demanda por transporte. Além disso, no caso brasileiro, a projeção do consumo é fundamental na definição das metas decenais da Política Nacional de Biocombustível.

Os resultados obtidos ainda são relevantes na definição e operacionalização de regras e normas de comercialização definidas para o mercado nacional. Entre elas, é possível mencionar os contratos de compra e venda de etanol anidro induzidos pela Resolução ANP N° 67/2011.

Por fim, a análise aqui conduzida não esgota a necessidade de novos estudos relacionados ao tema. Modelos que avaliam a demanda individual para cada combustível são importantes para um entendimento mais profundo do tema. Além disso, do ponto de vista metodológico, a inclusão de modelos combinados ou de métodos de aprendizagem conjunta devem ser avaliados no futuro.

6. Referências

- AHMAD, A. S. *et al.* A review on applications of ANN e SVM for building electrical energy consumption forecasting. *Renewable e Sustainable Energy Reviews*, v. 33, p. 102–109, 2014.
- AL-GHANDOOR, A. *et al.* Residential past e future energy consumption: Potential savings e environmental impact. *Renewable e Sustainable Energy Reviews*, v. 13, n. 6–7, p. 1262–1274, 2009.
- BABU, C. N.; REDDY, B. E. A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data. *Applied Soft Computing Journal*, v. 23, p. 27–38, 2014.
- BARAK, S.; SADEGH, S. S. Forecasting energy consumption using ensemble ARIMA–ANFIS hybrid algorithm. *Electrical Power e Energy Systems*, v. 82, p. 92–104, 2016.
- BERGMEIR, C.; HYNDMAN, R. J.; KOO, B. A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics e Data Analysis*, v. 120, p. 70–83, 2018.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. *Time Series Analysis: Forecasting e Control*. 4ed.



- ed. New Jersey: John Wiley & Sons, 2008. v. 20.
- BREIMAN, L. *et al.* *Classification e Regression Trees*. New York: CRC Press, 1984.
- BREIMAN, L. Random Forests. *Machine Learning*, v. 45, p. 5–32, 2001.
- BUENO, R. DE L. DA S. *Econometria de Séries Temporais*. 2. ed. São Paulo: Cengage Learning, 2018.
- BURNQUIST, H. L.; BACCHI, M. R. P. A demanda por gasolina no Brasil: uma análise utilizando técnicas de co-integração. 2000, Congresso da Sociedade Brasileira de Economia e Sociologia Rural - Sober, 2000. p. 1–13.
- CASTRO, G. R. DE. *Modelo de Previsão de Demanda por Combustíveis Automotivos no Brasil*. 2012. 71 f. Universidade Federal do Rio de Janeiro, 2012.
- CHANG, P. C.; FAN, C. Y.; LIN, J. J. Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach. *International Journal of Electrical Power e Energy Systems*, v. 33, n. 1, p. 17–27, 2011.
- CHE, J.; WANG, J. Short-term load forecasting using a kernel-based support vector regression combination model. *Applied Energy*, v. 132, p. 602–609, 2014.
- CHEN, Y. *et al.* Short-term electrical load forecasting using the Support Vector Regression (SVR) model to calculate the demand response baseline for office buildings. *Applied Energy*, v. 195, p. 659–670, 2017.
- CLEVELAND, R. B. *et al.* STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, v. 6, n. 1, p. 3–73, 1990.
- DE FREITAS, L. C.; KANEKO, S. Ethanol demand under the flex-fuel technology regime in Brazil. *Energy Economics*, v. 33, n. 6, p. 1146–1154, nov. 2011.
- DEB, C. *et al.* A review on time series forecasting techniques for building energy consumption. *Renewable e Sustainable Energy Reviews*, v. 74, n. March, p. 902–924, 2017.
- DEBNATH, K. B.; MOURSHED, M. Forecasting methods in energy planning models. *Renewable e Sustainable Energy Reviews*, v. 88, n. August 2016, p. 297–325, 2018.
- DEMSAR, J. Statistical Comparisons of Classifiers over Multiple Data Sets. *Journal of Machine Learning Research*, v. 7, p. 1–30, 2006.
- DICKEY, D. A.; PANTULA, S. G. Determining the Order of Differencing in Autoregressive Processes. *Journal of Business & Economic Statistics*, v. 5, n. 4, p. 455–461, 1987.
- DUDEK, G. Short-Term Load Forecasting using Random Forests. In: D FILEV *et al.* (Org.). *Intelligent Systems '2014. Advances in Intelligent Systems e Computing*. Cham: Springer, 2015. v. 323. p. 821–828.
- EBRAHIMPOUR, R. *et al.* Mixture of mlp-experts for trend forecasting of time series: A case study of the tehran stock exchange. *International Journal of Forecasting*, v. 27, n. 3, p. 804–816, 2011.
- EDGERTON, D.; SHUKUR, G. Testing autocorrelation in a system perspective testing autocorrelation. *Econometric Reviews*, v. 18, n. 4, p. 343–386, 1999.
- EDIGER, V. Ş.; AKAR, S. ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, v. 35, n. 3, p. 1701–1708, 2007.
- EDIGER, V. Ş.; AKAR, S.; UĞURLU, B. Forecasting production of fossil fuel sources in Turkey using a comparative regression e ARIMA model. *Energy Policy*, v. 34, n. 18, p. 3836–3846, 2006.
- ELLIOTT, G.; ROTHENBERG, T. J.; STOCK, J. H. Efficient Tests for an Autoregressive Unit Root. *Econometrica*, v. 64, n. 4, p. 813–836, 1996.
- ENGLE, R. F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, v. 50, n. 4, p. 987–1007, 1982.
- FAN, S.; CHEN, L.; LEE, W. J. Machine learning based switching model for electricity load forecasting. *Energy Conversion e Management*, v. 49, n. 6, p. 1331–1344, 2008.
- FIGUEIRA, S. R.; BURNQUIST, H. L. Programas para álcool combustível nos Estados Unidos e possibilidades de exportação do Brasil. *Agricultura em São Paulo*, v. 53, n. C, p. 5–18, 2006.
- GARCÍA NIETO, P. J. *et al.* PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA e ARIMA: A case study. *Science of the Total Environment*, v. 621, p. 753–761, 2018.
- HARVEY, ANDREW C.; SHEPHARD, N. Structural time series models. In: MADDALA, G. S.; RAO, C. R.; VINOD, H. D. (Org.). *Handbook of Statistics*. Elsevier Science Publishers, 1993. v. 11. p. 261–302.



- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning*. 2 ed ed. [S.l.]: Springer, 2009.
- HONG, W. Electric load forecasting by support vector model. *Applied Mathematical Modelling*, v. 33, n. 5, p. 2444–2454, 2009.
- HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: Principles e Practice*. 2nd. ed. Melbourne: Otexts, 2018.
- IBARRA-BERASTEGI, G. *et al.* Short-term forecasting of the wave energy flux: Analogues, random forests, e physics-based models. *Ocean Engineering*, v. 104, p. 530–539, 2015.
- ISLAS, J.; MANZINI, F.; MASERA, O. A prospective study of bioenergy use in Mexico. *Energy*, v. 32, n. 12, p. 2306–2320, 2006.
- JOHANSEN, S. Estimation e Hypothesis Testing of Cointegration Vectors in Gaussian Vector. *Econometrica*, v. 59, n. 6, p. 1551–1580, 1991.
- JOHANSEN, S. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics e Control*, v. 12, n. 2–3, p. 231–254, 1988.
- JOHANSSON, O.; SCHIPPER, L. Measuring the Long-Run Fuel Demand of Cars: Separate Estimations of Vehicle Stock, Mean Fuel Intensity, e Mean Annual Driving Distance. *Journal of Transport Economics e Policy*, v. 31, n. 3, p. 277–292, 1997.
- KADIR KAVAKLIOGLU, HALIM CEYLAN, HARUN KEMAL OZTURK, O. E. C. Modeling e prediction of Turkey's electricity consumption using Artificial Neural Networks. *Energy Conversion e Management*, v. 50, n. 11, p. 2719–2727, 2009.
- KANDANANOND, K. Forecasting electricity demand in Thailand with an artificial neural network approach. *Energies*, v. 4, n. 8, p. 1246–1257, 2011.
- KHASHEI, M.; BIJARI, M. An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with Applications*, v. 37, n. 1, p. 479–489, 2010.
- KIM, J.; KASABOV, N. HyFIS : adaptive neuro-fuzzy inference systems e their application to nonlinear dynamical systems. *Neural Networks*, v. 12, p. 1301–1319, 1999.
- KUCUKALI, S.; BARIS, K. Turkey's short-term gross annual electricity demand forecast by fuzzy logic approach. *Energy Policy*, v. 38, n. 5, p. 2438–2445, 2010.
- KUHN, M.; JOHNSON, K. *Applied Predictive Modeling*. New York: Springer, 2013.
- KWIATKOWSKI, D. *et al.* Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, v. 54, p. 159–178, 1992.
- LAHOUAR, A.; BEN HADJ SLAMA, J. Day-ahead load forecast using random forest e expert input selection. *Energy Conversion e Management*, v. 103, p. 1040–1051, 2015.
- LÜTKEPOHL, H. *New Introduction to Multiple Time Series Analysis*. 2. ed. Berlin: Springer, 2005.
- M.J., K. *et al.* Comparison of ARIMA e Random Forest time series models for prediction of avian influenza H5N1 outbreaks. *BMC bioinformatics*, v. 15, p. 276, 2014.
- MAIA, A. L. S.; DE CARVALHO, F. D. A. T.; LUDERMIR, T. B. Forecasting models for interval-valued time series. *Neurocomputing*, v. 71, n. 16–18, p. 3344–3352, 2008.
- MAMDANI, E. H. Application of fuzzy algorithms for control of simple dynamic plant. *Proceedings of the Institution of Electrical Engineers*, v. 121, n. 12, p. 1585–1588, 1974.
- MELIKOGLU, M. Demand forecast for road transportation fuels including gasoline, diesel, LPG, bioethanol and biodiesel for Turkey between 2013 and 2023. *Renewable Energy*, v. 64, p. 164-171, 2014.
- MOREIRA, A. R. B. *Modelos para a Projeção do Consumo Nacional e Regional de Óleo Diesel*. Rio de Janeiro: Instituto de Pesquisa Econômica Aplicada, 1996.
- MU, H. *et al.* Grey relative analysis e future prediction on rural household biofuels consumption in China. *Fuel Processing Technology*, v. 85, n. 8–10, p. 1231–1248, 2004.
- NASR, G. ; BADR, E. ; JOUN, C. Backpropagation neural networks for modeling gasoline consumption. *Energy Conversion e Management*, v. 44, n. 6, p. 893–905, 2003.
- NELSON, M. *et al.* Time series forecasting using neural networks: Should the data be deseasonalized first? *Journal of Forecasting*, v. 18, n. 5, p. 359–367, 1999.
- NENGLING, T.; STENZEL, J.; HONGXIAO, W. Techniques of applying wavelet transform into



- combined model for short-term load forecasting. *Electric Power Systems Research*, v. 76, n. 6–7, p. 525–533, 2006.
- NIE, H. *et al.* Hybrid of ARIMA e SVMs for short-term load forecasting. *Energy Procedia*, v. 16, n. PART C, p. 1455–1460, 2011.
- OSBOM, D. R. *et al.* Seasonality e the Order of Integration for Consumption. *Oxford Bulletin of Economics e Statistics*, v. 50, n. 4, p. 361–377, 1988.
- PADMAKUMARI, K.; MOHANDAS, K. P.; THIRUVENGADAM, S. Long term distribution demand forecasting using neuro fuzzy computations. *International Journal of Electrical Power e Energy System*, v. 21, n. 5, p. 315–322, 1999.
- PAI, P. F. Hybrid ellipsoidal fuzzy systems in forecasting regional electricity loads. *Energy Conversion e Management*, v. 47, n. 15–16, p. 2283–2289, 2006.
- RINGWOOD, J. V.; BOFELLI, D.; MURRAY, F. T. Forecasting Electricity Demand on Short, Medium e Long Time Scales Using Neural Networks. *Journal of Intelligent e Robotic Systems*, v. 31, n. May, p. 129–147, 2001.
- RODRIGUES, L.; BACCHI, M. R. P. Analyzing light fuel demand elasticities in Brazil using cointegration techniques. *Energy Economics*, v. 63, p. 322–331, 2017.
- RODRIGUES, L.; RODRIGUES, L. Economic-financial performance of the Brazilian sugarcane energy industry: An empirical evaluation using financial ratio, cluster e discriminant analysis. *Biomass e Bioenergy*, v. 108, n. October 2017, p. 289–296, 2018.
- SANTIAGO, F. S.; DE MATTOS, R. S.; PEROBELLI, F. S. Um modelo integrado econométrico+insumo-produto para previsão de longo prazo da demanda de combustíveis no Brasil. *Nova Economia*, v. 21, n. 3, p. 423–455, 2011.
- SHAPIRO, S. S.; WILK, M. B. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, v. 52, n. 4, p. 591–611, 1965.
- SHINGALA, M. C. *Comparison of different post HOC tests*. 2016. 188 f. Veer Narmad South Gujarat University, 2016.
- SIMON HAYKIN. *Redes Neurais: Princípios e Prática*. 2 ed ed. Porto Alegre: Bookman, 2001.
- SMITH, C.; HALL, S.; MABEY, N. Econometric modelling of international carbon tax regimes. *Energy Economics*, v. 17, n. 2, p. 133–146, 1995.
- SÖZEN, A.; ARCAKLIOĞLU, E.; ÖZKAYMAK, M. Turkey's net energy consumption. *Applied Energy*, v. 81, n. 2, p. 209–221, 2005.
- SUGANTHI, L.; SAMUEL, A. A. Energy models for demand forecasting — A review. *Renewable e Sustainable Energy Reviews*, v. 16, p. 1223–1240, 2012.
- WANG, L. X.; MENDEL, J. M. Generating Fuzzy Rules by Learning from Examples. *IEEE Transactions on Systems, Man e Cybernetics*, v. 22, n. 6, p. 1414–1427, 1992.
- WEI, C. C.; CHEN, T. T.; LEE, S. J. K-NN based neuro-fuzzy system for time series prediction. *14th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking e Parallel/Distributed Computing*, p. 569–574, 2013.
- WEST, M.; HARRISON, J. *Bayesian Forecasting e Dynamic Models*. 2nd. ed. New York: Springer, 1997.
- XIA, C.; WANG, J.; MCMENEMY, K. Short, medium e long term load forecasting model e virtual load forecaster based on radial basis function neural networks. *International Journal of Electrical Power e Energy Systems*, v. 32, n. 7, p. 743–750, 2010.
- ZADEH, L. A. Fuzzy Sets. *Information e Control*, v. 8, n. 3, p. 338–353, 1965.
- ZANINI, A.; SOUZA, R. C.; PEDREIRA, C. E. *Um modelo híbrido para previsão de curto prazo da demanda de gasolina automotiva no Brasil*. 2001. 403–407 f. Pontifícia Universidade Católica do Rio de Janeiro - PUC-RIO, 2001.
- ZHANG, G. P. Time series forecasting using a hybrid ARIMA e neural network model. *Neurocomputing*, v. 50, n. 0, p. 159–175, 2003.
- ZHANG, G. P.; QI, M. Neural network forecasting for seasonal e trend time series. *European Journal of Operational Research*, v. 160, n. 2, p. 501–514, 2005.