

Previsão e construção de indicadores operacionais da produção de cana-de-açúcar por meio de séries temporais

Anderson da Silva Santos¹, João Vitor da Silva¹,
Victor Wanderley Costa de Medeiros¹, Glauco Estácio Gonçalves¹

Departamento de Estatística e Informática – Universidade Federal
Rural de Pernambuco (UFRPE), Recife, PE – Brasil

{anderson.silvasantos, joao.vitorsilva,
victor.wanderley, glauco.goncalves}@ufrpe.br

Abstract. *With each harvest, millions of liters of ethanol and thousands of tons of sugar are exported by the Brazilian sugar and alcohol sector. Idleness and work overload in production due to the lack of predictability about the amount of raw material available are in the sector. This work aims to build operational indicators, using the waste distribution of the SARIMA model (Seasonal Autoregressive Integrated Moving Average) as a basis, to inform the probability of a reduction in the production of sugarcane. The database was taken from the quarterly results published by the São Martinho Group. The results obtained show the potential of the proposed model for decision-making.*

Resumo. *A cada safra, milhões de litros de etanol e milhares de toneladas de açúcar são exportados pelo setor sucroalcooleiro brasileiro. A ociosidade e sobrecarga de trabalho na produção devido à falta de previsibilidade sobre a quantidade da matéria-prima disponível são alguns problemas no setor. Este trabalho tem como objetivo construir indicadores operacionais, utilizando como base a distribuição dos resíduos do modelo SARIMA (Seasonal Autoregressive Integrated Moving Average), para informar a probabilidade de redução na produção de cana-de-açúcar. A base de dados foi retirada dos resultados trimestrais publicados pelo Grupo São Martinho. Os resultados obtidos evidenciam o potencial do modelo proposto para a tomada de decisão.*

1. Introdução

O Brasil é o líder mundial na produção de cana-de-açúcar e etanol de cana-de-açúcar. Consequentemente, a indústria sucroalcooleira possui papel de destaque na economia brasileira [Pereira et al. 2019]. Apesar desta liderança, o setor sucroalcooleiro brasileiro também é impactado, como os demais setores produtivos, pelas diversas mudanças nos processos de produção trazidas pelas tecnologias digitais [Silva 2019]. A moagem da cana-de-açúcar, por exemplo, que antes era feita por animais, hoje é amparada por maquinário, sensores e atuadores em um processo completamente definido e automatizado [Silva 2019].

Devido a tais mudanças, surgiram outros desafios para a melhoria contínua nas atividades. Métricas dos processos e ferramentas para auxílio na tomada de decisões estratégicas são requisitos essenciais para obtenção de vantagem competitiva no mercado.

Pelo fato da indústria sucroalcooleira apenas funcionar durante uma parte do ano, é fundamental garantir a otimização do tempo e da força de trabalho ao longo do período da safra [Silva 2019]. Problemas, como a queda na chegada de insumos, podem provocar a parada de toda a linha de produção. Além disso, conforme [Saath and Fachinello 2018], as novas demandas do setor agrícola deverão ser atendidas por meio do crescimento da produtividade. Portanto, as soluções capazes de melhorar a previsibilidade da produção são necessárias para o setor agrícola e conseqüentemente para a sociedade.

Diante disso, este trabalho propõe um modelo de estimação da produção de cana-de-açúcar com base em dados de séries temporais. Uma heurística foi criada para determinar a escolha dos melhores valores dos hiperparâmetros do modelo. Além disso, dois indicadores operacionais relacionados com a probabilidade da redução da produção foram gerados para agregar valor no planejamento estratégico de uma usina sucroalcooleira.

Neste sentido, [Suresh and Priya 2011] construíram modelos a fim de prever a área, produção e produtividade da cana-de-açúcar em Tamil Nadu, Índia. Utilizando dados coletados entre 1950 e 2007, os modelos foram usados para previsão dos valores para os anos seguintes. O presente trabalho, assim como [Suresh and Priya 2011], desenvolveu modelos capazes de prever a produção de cana-de-açúcar no Brasil. [Hossain and Abdulla 2015] utilizaram modelos ARIMA nos dados de produção anual de cana-de-açúcar em Bangladesh durante o período de 1971 a 2013 para efetuar previsões da produção. A comparação entre a série original e a série prevista mostrou similaridade razoável.

Neste estudo, as previsões geradas pelo modelo são comparadas com o padrão sazonal da série e com os dados de teste e treino por meio das métricas de erro para selecionar o melhor modelo. Inspirado em [MURTA et al. 2005], foram utilizados os resíduos do modelo para o cálculo das probabilidades, por meio da função de distribuição, para obtenção dos valores dos indicadores de redução de produção de cana-de-açúcar.

Este artigo está organizado da seguinte forma: Seção 2 descreve a base de dados utilizada nos experimentos, explica os processos da heurística para definição do modelo, a escolha dos melhores valores para os hiperparâmetros e a criação dos indicadores operacionais; Seção 3 mostra os resultados obtidos junto com as análises; Por fim, a Seção 4 apresenta as conclusões e trabalhos futuros.

2. Materiais e Métodos

Para a realização dos experimentos e análises neste trabalho, a linguagem R¹ foi escolhida devido a sua simplicidade, eficiência e disponibilidade de bibliotecas para o processamento de séries temporais. Dentre algumas delas, destacam-se a `forecast`², usada para a análise e visualização das previsões das séries temporais com ARIMA, a `tseries`³ para a análise de séries temporais, a `ggplot2`⁴ para a construção de gráficos e a `dplyr`⁵ para uma melhor manipulação nas estruturas de dados das séries temporais.

¹<https://www.r-project.org>

²<https://cran.r-project.org/web/packages/forecast/index.html>

³<https://cran.r-project.org/web/packages/tseries/index.html>

⁴<https://cran.r-project.org/web/packages/ggplot2/index.html>

⁵<https://cran.r-project.org/web/packages/dplyr/index.html>

2.1. Base de dados

A base de dados utilizada foi gerada a partir dos relatórios trimestrais publicados pelo Grupo São Martinho, um dos maiores produtores de açúcar, álcool e energia do Brasil⁶. Nesses relatórios, foram registrados os montantes de cana-de-açúcar processada em milhões de toneladas durante o primeiro trimestre do ano safra 06/07 até o primeiro trimestre do ano safra 19/20, totalizando 53 observações ao longo de 13 anos.

Para realização dos experimentos, foi utilizada a técnica *holdout*, a qual consiste em particionar os dados da base de dados em dois conjuntos (treinamento e teste) para fornecer uma estimativa menos tendenciosa do desempenho do modelo [Sammut and Webb 2010]. Após a divisão, a base de treino foi composta por 45 amostras (1º trimestre da safra 06/07 até o 1º trimestre da safra 17/18) e a base de teste por 8 amostras (2º trimestre da safra 17/18 até o 1º trimestre da safra 19/20).

2.2. Modelo de previsão

A Figura 1 mostra a decomposição da série temporal de produção trimestral de cana-de-açúcar. Observa-se a própria série (*data*), os padrões de sazonalidade (*seasonal*), a tendência (*trend*) e o erro (*remainder*). Analisando a decomposição, percebe-se o padrão sazonal, tanto na própria série quanto na parte sazonal da decomposição. Como os dados utilizados possuem sazonalidade, o modelo de previsão SARIMA, proposto por [Box et al. 2015], foi escolhido devido à sua capacidade de tratar séries temporais com essas características.

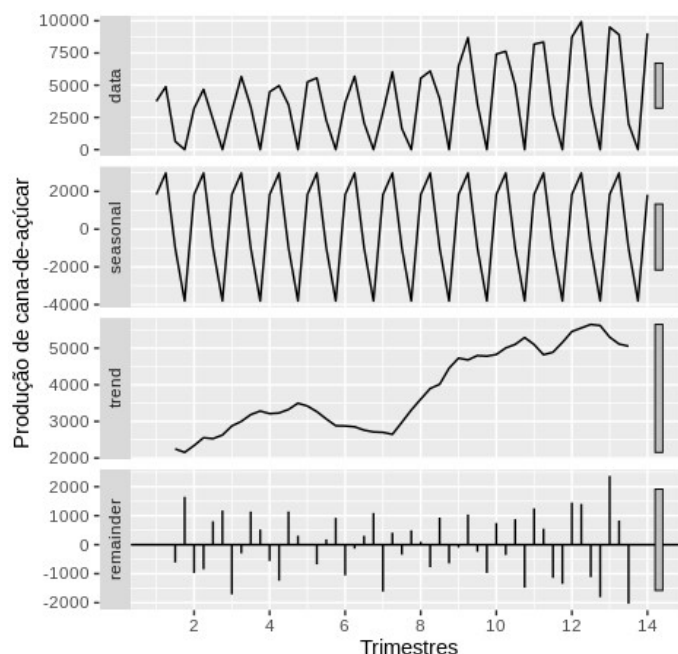


Figura 1. Decomposição da série temporal de produção de cana-de-açúcar.

Tal modelo é conhecido como a função $SARIMA(p, d, q)(P, D, Q)$. O parâmetro p é a ordem (número de defasagens) do modelo auto-regressivo, d é o grau de

⁶Central de Resultados da São Martinho - <https://ri.saomartinho.com.br>

diferenciação, q é a ordem do modelo de média móvel, P é a ordem do modelo auto-regressivo da parte sazonal, D é o grau de diferenciação da parte sazonal e Q é a ordem do modelo de média móvel da parte sazonal. Uma definição mais formal, de acordo com [Haviluddin and Dengen 2016], é representada pela Equação 1.

$$\Phi_p B^S \phi_p(B) (1 - B)^d (1 - B^S)^D Z_t = \theta_q(B) \Theta_q(B^S) a_t \quad (1)$$

onde $\phi_p(B)$ é a auto-regressão não-sazonal, $\Phi_p B^S$ é a auto-regressão sazonal, $(1 - B)^d$ é a diferenciação não sazonal, $(1 - B^S)^D$ é a diferenciação sazonal, $\theta_q(B)$ é a média móvel não-sazonal e $\Theta_q(B^S)$ a média móvel sazonal.

2.3. Otimização dos hiperparâmetros

Neste trabalho, o SARIMA foi considerado um modelo base, uma vez que possui seis hiperparâmetros, os quais podem ser combinados para gerar diferentes modelos. Para a escolha dos melhores valores para os hiperparâmetros, foi definida uma heurística baseada em técnicas de teste de hipótese e análise gráfica. A princípio, buscou-se definir os valores dos hiperparâmetros d (diferenciação da série) e D (padrão sazonal).

Para determinar o valor de d , utilizou-se o teste de Dickey-Fuller [Dickey and Fuller 1979] para identificar a presença de estacionariedade na série. Foram aplicadas diferenciações até o teste de hipótese informar quando a série se tornou estacionária e, conseqüentemente, detectar a quantidade de diferenciações necessárias para tornar a série estacionária. Aplicando o teste de Dickey-Fuller à série temporal, sem qualquer diferenciação, obteve-se um valor-p de 0,71, i.e., não foi possível refutar a hipótese de não-estacionariedade. Após uma diferenciação, obteve-se um valor-p de 0,01, confirmando a estacionariedade da série. Desta forma, concluiu-se que $d = 1$.

Já para escolher D , utilizou-se o decaimento do gráfico da função de autocorrelação da série temporal como indicador da estacionariedade na sazonalidade. Foram aplicadas diferenciações sazonais até o decaimento da autocorrelação tornar-se aparentemente exponencial, indicando quando a série se tornou estacionária na parte sazonal e definindo o número necessário de diferenciações sazonais. A Figura 2a mostra a autocorrelação da série temporal de produção de cana-de-açúcar. A análise gráfica indica a presença de um padrão sazonal, visto que a autocorrelação nos *lags* (atrasos) sazonais foi significativa. Além disso, a função de autocorrelação tem um decaimento linear, indicando a falta de estacionariedade da série na parte sazonal. Ao aplicar uma diferenciação sazonal à série, obteve-se uma nova função de autocorrelação em que o padrão de sazonalidade foi eliminado, conforme mostra a Figura 2b. Nota-se que o decaimento apresentado é exponencial. Portanto, foi necessário apenas uma diferenciação sazonal ($D = 1$) para a série se tornar estacionária.

Para a escolha do restante dos hiperparâmetros (p , q , P e Q), foram realizadas combinações para os valores dentro do intervalo $[0, 5]$. Cada modelo gerado foi avaliado pelo critério de informação AICc, o qual consiste em uma variante para pequenas amostras de uma das medidas mais usadas na avaliação de modelos estatísticos chamada AIC (*Corrected Akaike Information Criterion*) [Mondal et al. 2014], e pela medida de erro RMSE (*Root Mean Square Error*). As Equações 2, 3 e 4 representam essas medidas respectivamente.

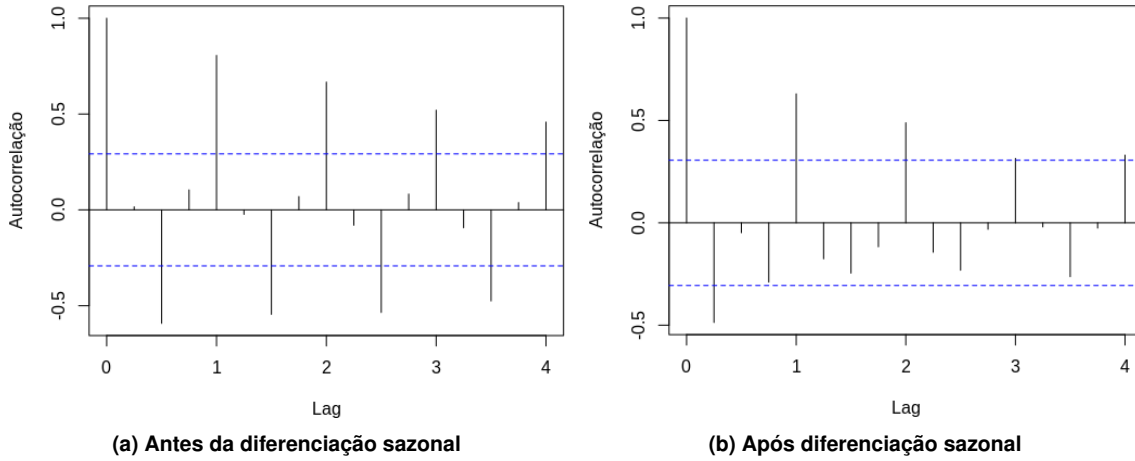


Figura 2. Gráficos da função de autocorrelação dos dados de produção de cana-de-açúcar.

$$AIC_c = AIC + \frac{2k^2 + 2k}{n - k - 1} \quad (2)$$

$$AIC = 2k - 2\ln(\hat{L}) \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} \quad (4)$$

Onde k é o número de parâmetros estimados no modelo, \hat{L} o valor máximo da função de verossimilhança para o modelo, e a diferença entre o esperado e o previsto e n o tamanho da amostra. A seleção foi definida de acordo com a melhor performance dos modelos, considerando essas duas métricas. Como saída dessa etapa, 12 modelos foram selecionados de um total de 1296.

Em seguida, as medidas ME (*Mean Percentege Error*) e MAE (*Mean Absolute Error*) foram utilizadas para a escolha final dos hiperparâmetros. As Equações 5 e 6, respectivamente, apresentam essas métricas. As médias dos dados de treino e teste foram calculadas e os resultados são mostrados na Tabela 1. Analisando os desempenhos dos modelos na base de teste, em relação às métricas ME e MAE, conclui-se que o conjunto de parâmetros mais adequado foi o $SARIMA(0, 1, 1)(0, 1, 0)$.

$$ME = \frac{\sum_{i=1}^n e_i}{n} \quad (5)$$

$$MAE = \frac{\sum_{i=1}^n |e_i|}{n} \quad (6)$$

Tabela 1. Conjunto de hiperparâmetros dos modelos com os menores AICc, RMSE, ME e MAE.

(p,d,q)	(P,D,Q)	RMSE	AICC	ME*	ME**	MAE*	MAE**
(0,1,1)	(0,1,0)	962.28	675.81	49.12	25.11	662.08	582.00
(0,1,2)	(0,1,0)	954,01	677,66	46,87	28,92	656,86	588,26
(0,1,1)	(1,1,0)	940,82	676,82	60,26	-77,58	664,62	605,24
(0,1,2)	(1,1,0)	959,54	677,99	14,05	375,01	696,90	683,77
(0,1,1)	(3,1,0)	867,88	676,71	86,75	-11,81	633,12	861,93
(0,1,2)	(3,1,0)	879,21	678,92	40,48	-27,84	637,96	916,76
(0,1,1)	(4,1,0)	854,18	678,75	78,75	-90,12	622,05	805,88
(0,1,1)	(0,1,1)	933,61	676,54	77,15	-42,83	679,28	746,64
(0,1,2)	(1,1,1)	930,10	678,68	11,79	532,51	697,01	951,18
(0,1,1)	(3,1,1)	823,27	677,61	75,25	-3,70	609,84	772,01
(0,1,1)	(0,1,2)	924,71	678,63	80,13	-93,40	682,85	682,83
(0,1,1)	(1,1,3)	816,30	677,81	69,27	66,02	610,58	821,36

* dados de treino

** dados de teste

2.4. Indicadores operacionais

Nesta etapa, foram gerados dois indicadores operacionais denominados I_1 e I_2 . Sendo o primeiro a probabilidade da produção de cana-de-açúcar ser acima da média de produção histórica com base no valor previsto pelo modelo, e o segundo representando a probabilidade da produção prevista ser reduzida em um dado percentual p' ou mais.

Os resíduos do modelo selecionado foram utilizados na construção de ambos indicadores. A Figura 3a mostra a distribuição dos resíduos gerados pelo modelo. Devido à pequena amostra de resíduos, realizou-se uma reamostragem para obter uma melhor representação da distribuição dos resíduos por meio da simulação Monte Carlo. Após a simulação, a nova amostragem foi composta por 1000 observações. A Figura 3b mostra a distribuição gerada pela simulação. Os dados dos resíduos simulados foram submetidos ao teste de hipótese Shapiro-Wilk [Shapiro and Wilk 1965] para verificar se seguiam uma distribuição normal. Como o valor-p foi menor que 0,05, obteve-se a função de distribuição acumulada empiricamente.

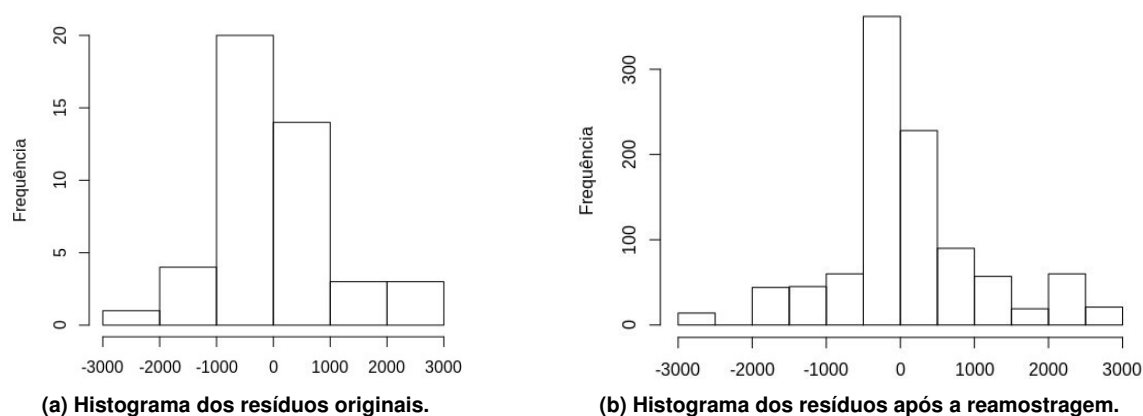


Figura 3. Gráficos de distribuição dos resíduos do modelo SARIMA(0, 1, 1)(0, 1, 0).

O processo para obter I_1 foi definido da seguinte forma: depois da simulação, adiciona-se a média histórica de produção de cana-de-açúcar, obtida a partir dos dados de produção, aos resíduos simulados. Desta forma, a função de distribuição acumulada é ajustada ao comportamento estocástico da média de produção histórica. A partir dos dados simulados, a função de distribuição acumulada pode ser obtida empiricamente. Assumindo que a média de produção histórica é uma variável aleatória P_h , que a função acumulada empírica é dada por $\widehat{F}_{P_h}(x)$ e que o valor previsto pelo modelo SARIMA para um trimestre futuro é \hat{p} , então a probabilidade da produção ser acima da média (p_m) é dada por:

$$p_m = \widehat{F}_{P_h}(\hat{p}) \quad (7)$$

Já para o indicador I_2 , o processo foi estabelecido desta maneira: após a simulação, cria-se a função de distribuição acumulada empírica destes resíduos. Considerando que os resíduos são uma variável aleatória R , que a função acumulada empírica é dada por \widehat{F}_R e que o valor previsto pelo modelo para um trimestre futuro é \hat{p} , então a probabilidade de redução da produção $p_r(p')$ em um dado percentual p' ou mais é dada por:

$$p_r(p') = \widehat{F}_R(-p' \times \hat{p}) \quad (8)$$

Como a distribuição dos resíduos do modelo SARIMA tende a ter um conjunto de dados centrados em torno de zero, o uso do valor negativo para o percentual permite capturar os casos em que há a redução da produção de cana-de-açúcar. Além disso, a própria definição da função acumulada de probabilidade captura os casos que vão além do percentual de redução p' definido. Para fins de experimentação e demonstração do funcionamento do indicador I_2 , p' foi definido como 10%, 20% e 30%.

3. Resultados e Discussões

A Tabela 2 mostra os resultados, em cada trimestre, da previsão dos dados da base de teste com o modelo $SARIMA(0, 1, 1)(0, 1, 0)$ junto com os valores reais de produção. A Figura 4 representa uma outra visualização desses resultados.

Tabela 2. Previsões do modelo $SARIMA(0, 1, 1)(0, 1, 0)$. Valores dados em milhares de toneladas de cana-de-açúcar.

Trimestre	Previsão	Produção
T1	8603,51	9933
T2	3006,51	3534
T3	257,51	0
T4	8996,51	9508
T5	8861,02	8921
T6	3264,02	2021
T7	515,02	0
T8	9254,02	9042

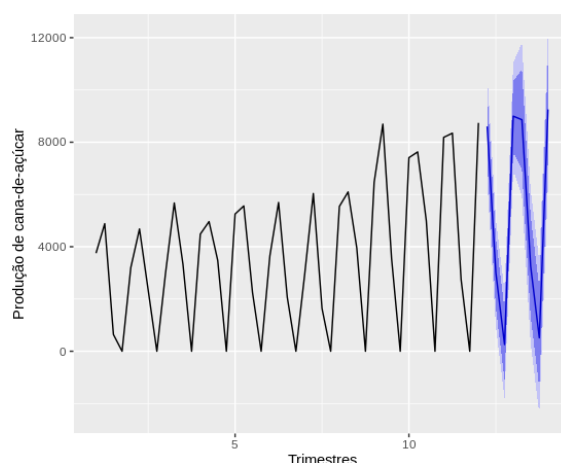


Figura 4. Gráfico de previsão dos dados de produção de cana-de-açúcar. Em destaque, os trimestres previstos pelo modelo $SARIMA(0, 1, 1)(0, 1, 0)$.

A Tabela 2 mostra que o modelo de previsão aproximou razoavelmente os valores relativos aos trimestres entre os anos 2017 e 2019 para a base de teste, alcançando um erro médio absoluto de aproximadamente 662 milhares de toneladas. Além disso, a Figura 4 mostra que o padrão sazonal da produção de cana-de-açúcar do Grupo São Martinho foi preservado no modelo. Outra característica que pode ser observada nesses resultados é uma tendência de subestimação dos valores reais. De certa forma, o modelo é pessimista e tal característica desfavorece decisões que estejam muito acima da média.

A Tabela 3 mostra os resultados para os indicadores desenvolvidos, nela pode-se ver a produção prevista para cada trimestre de teste; o valor do indicador I_1 , considerando a média de produção histórica de 3,86 milhões de toneladas de cana-de-açúcar; e as probabilidades referentes ao indicador I_2 com o percentual de redução configurado para 10%, 20% e 30%.

Tabela 3. Previsões do Modelo $SARIMA(0, 1, 1)(0, 1, 0)$ e indicadores I_1 e I_2 com percentual de redução de 10%, 20% e 30% por trimestre.

	T1	T2	T3	T4	T5	T6	T7	T8
Valor previsto	8603,51	3006,51	257,51	8996,51	8861,02	3264,02	515,02	9254,02
I_1	100%	10%	0%	100%	100%	15%	0%	100%
$I_2(10\%)$	11,3%	29,2%	48,7%	11,3%	11,3%	21,1%	48,7%	11,3%
$I_2(20\%)$	4,3%	14,5%	48,7%	1,7%	1,7%	14,5%	44,4%	1,7%
$I_2(30\%)$	0%	11,3%	46,2%	0%	0%	11,3%	42,5%	0%

Para análise do indicador I_1 , considerou-se que, para correta operação, a usina necessite de uma produção, no mínimo, igual à média histórica. Assim, analisando a Tabela 3, percebe-se que o funcionamento da produção de cana-de-açúcar é garantida em condições satisfatórias nos trimestres T1, T4, T5 e T8, pois a probabilidade da produção ser acima da média é de 100%. Já nos trimestres T2, T3, T6 e T7 a fábrica precisaria rever seu esquema de funcionamento, tendo em vista que o indicador aponta grandes chances da produção não atingir a média, provavelmente por serem trimestres de final de safra onde a produção é pequena ou inexistente.

Em relação ao indicador I_2 , as probabilidades indicam pequena ou nenhuma redução da produção nos trimestres T1, T4, T5 e T8. Já nos trimestres T2, T3, T6 e T7 as probabilidades se mantêm maiores, principalmente no T3 e T7 onde a probabilidade passa de 40%.

De outra forma, observa-se que há uma alta probabilidade, acima de 40%, da produção reduzir em até 30% nos trimestres finais da safra. Esse resultado tem relação com o já explicado no indicador I_1 : trimestres finais de safra têm sua produção muito baixa ou quase nula, provocando assim a probabilidade de redução de produção maior que em outros trimestres.

Outro ponto a destacar é a diminuição do indicador operacional conforme o percentual de produção aplicado é aumentado, indicando uma improvável chance da probabilidade de produção de cana-de-açúcar reduzir mais do que 30%, na maior parte dos trimestres.

4. Conclusão e trabalhos futuros

Este trabalho realizou um estudo sobre aplicação do modelo de previsão SARIMA e a construção de indicadores operacionais de estimativas em dados de produção de cana-de-açúcar. Os dados utilizados foram retirados dos relatórios trimestrais de produção publicados pela São Martinho desde 2007 até o início de 2019. Foram utilizados 45 trimestres para treinamento e 8 trimestres para teste.

Os dados foram estudados sob a perspectiva da sazonalidade, tendência e resíduos. Utilizando diferentes critérios, o modelo SARIMA mais adequado aos dados foi escolhido. O modelo $SARIMA(0, 1, 1)(0, 1, 0)$ foi utilizado então nas previsões de produção.

Com base nos resultados de previsão do modelo escolhido, foram desenvolvidos dois indicadores operacionais. O primeiro indicador (I_1) representa a probabilidade da produção de um trimestre futuro ser acima da média de produção da série histórica. O segundo indicador (I_2) corresponde à probabilidade da produção prevista ser reduzida para além de um certo percentual p' .

Ambos indicadores utilizam, como base, a distribuição de probabilidade dos resíduos do modelo. Como a amostra de resíduos foi pequena, devido aos dados disponíveis, utilizou-se a reamostragem para aumentar o tamanho da amostra e, por conseguinte, aplicar a função de distribuição acumulada. Para cada previsão, foram calculados os indicadores operacionais de produção.

Esses indicadores capturam aspectos que são essenciais para a operação de usinas sucroalcooleiras. Tais usinas dependem de um fluxo mínimo de insumos para que sua operação seja economicamente viável ao longo de uma safra. A queda imprevista, na produção de cana-de-açúcar, aumenta os custos de operação da usina, gera atrasos e causa prejuízos econômicos. Assim, estes indicadores têm potencial para permitir a tomada de decisão mais cedo, evitando maiores impactos pela diminuição da produção de cana-de-açúcar.

O modelo proposto possui grande relevância para a tomada de decisões estratégicas. Por meio dele, é possível prever reduções da produção que possam levar a paradas dos processos na indústria sucroalcooleira. Apesar dos indicadores desenvolvidos neste trabalho serem limitados a informar a probabilidade de redução da produção,

a estratégia aplicada na construção desses indicadores pode ser aplicada para construção de novos indicadores sobre outros aspectos de tomada de decisão, para além do nível operacional.

Para trabalhos futuros, espera-se poder aplicar o método proposto a outros tipos de dados de produção e a outras escalas de tempo, tais como: diária, semanal e mensal. Desta forma, vislumbra-se que o método aplicado e os indicadores operacionais desenvolvidos podem ser tomados como base para melhoria de processos de uma usina elétrica, de uma produtora de leite ou outros setores produtivos.

Referências

- Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.
- Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.
- Haviluddin and Dengen, N. (2016). Comparison of sarima, narx and bpnn models in forecasting time series data of network traffic. In *2016 2nd International Conference on Science in Information Technology (ICSITech)*, pages 264–269.
- Hossain, M. M. and Abdulla, F. (2015). Forecasting the sugarcane production in bangladesh by arima model. *Journal of Statistics Applications & Probability*, 4(2):297.
- Mondal, P., Shit, L., and Goswami, S. (2014). Study of effectiveness of time series modeling (arima) in forecasting stock prices. *International Journal of Computer Science, Engineering and Applications*, 4(2):13.
- MURTA, R. M., TEODORO, S. M., Bonomo, P., and CHAVES, M. A. (2005). Precipitação pluvial mensal em níveis de probabilidade pela distribuição gama para duas localidades do sudoeste da bahia. *Ciência e Agrotecnologia*, 29(5):988–994.
- Pereira, I. Z., dos Santos, I. F. S., de Castro, H. L., and Barros, S. R. M. (2019). Uma breve revisão sobre a indústria sucroalcooleira no brasil com enfoque no potencial de geração de energia. *Revista Brasileira de Energia— Vol*, 25(2).
- Saath, K. C. d. O. and Fachinello, A. L. (2018). Crescimento da demanda mundial de alimentos e restrições do fator terra no brasil. *Revista de Economia e Sociologia Rural*, 56(2):195–212.
- Sammut, C. and Webb, G. I., editors (2010). *Holdout Evaluation*, pages 506–507. Springer US, Boston, MA.
- Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611.
- Silva, J. V. d. (2019). Suporte a decisão no setor sucroalcooleiro. B.S. thesis, Brasil.
- Suresh, K. and Priya, S. K. (2011). Forecasting sugarcane yield of tamilnadu using arima models. *Sugar Tech*, 13(1):23–26.