

Previsões estatísticas de área plantada e colhida do grão de trigo no Estado do Paraná

Carolina Deina (UTFPR) caroldeina@gmail.com

Carla Adriana Pizarro Schmidt (UTFPR) carlaschmidt@utfpr.edu.br

Flavio Trojan (UTFPR) trojan@utfpr.edu.br

Resumo:

O objetivo deste trabalho foi realizar uma análise estatística e previsão dos dados de áreas plantadas e áreas colhidas do grão de trigo no estado do Paraná. Para que os objetivos deste trabalho pudessem ser respondidos, coletou-se dados históricos para um período de 26 anos dessas variáveis, diretamente da base de dados do Instituto Brasileiro de Geografia e Estatística – IBGE. Esses dados passaram por uma análise estatística para identificar a presença de ciclos, sazonalidade, tendência e padrões. Posteriormente, para ambas variáveis analisadas o modelo escolhido para realizar a previsão foi Suavização Exponencial, utilizando o software Action Stat®. Os modelos encontrados para área plantada e área colhida apresentaram nível multiplicativo, tendência e sazonalidade ausentes. Os resultados apresentaram que a maior parte de área plantada foi também colhida ao longo dos anos. As previsões encontradas para todas as variáveis foram satisfatórias, pois estavam dentro dos limites superiores e inferiores da previsão.

Palavras chave: Trigo no Paraná, previsão de área plantada, Previsão de área colhida, séries temporais.

Statistical forecasts of planted and harvested area of wheat grain in the State of Paraná

Abstract

The objective of this work was to perform a statistical analysis and prediction of the data of planted areas and areas harvested from wheat grain in the state of Paraná. In order to answer the objectives of this work, historical data were collected for a period of 26 years of these variables, directly from the Brazilian Institute of Geography and Statistics (IBGE) database. These data underwent a statistical analysis to identify the presence of cycles, seasonality, trend and patterns. Later, for both analyzed variables, the model chosen to perform the prediction was Exponential Smoothing, using the Action Stat® software. The models found for planted area and harvested area presented multiplicative level, tendency and seasonality absent. The results showed that the majority of planted area was also harvested over the years. The predictions found for all variables were satisfactory, since they were within the upper and lower limits of the forecast.

Key-words: Wheat in Paraná, planted area forecast, harvested area forecast, time series.

1. Introdução

O trigo, considerado essencial na alimentação humana, é um dos três cereais mais consumidos no mundo, juntamente com o milho e o arroz. Ele é utilizado em grande escala na elaboração de diversos produtos alimentícios, colas, bebidas, etc (Companhia Nacional de Abastecimento - CONAB, 2017).

De acordo com o Centro de Estudos Avançados em Economia Aplicada – CEPEA (2017), no ano de 2016, o PIB do agronegócio Brasileiro acumulou crescimento de 4,48%, sendo que no segmento agrícola, as produções de café e trigo apresentaram as maiores elevações em volume no período.

Dos estados que mais produziram esse grão na safra de 2014/2015, conforme a CONAB (2017), encontram-se os estados do Paraná, representando 55%, e do Rio Grande do Sul, representando uma parcela de 38% dessa safra.

Sendo assim, a necessidade de conhecer e analisar o futuro deste grão é perceptível, e para isso é preciso adotar ferramentas que auxiliam essa análise e conhecimento.

De acordo com Krajewski, Ritzman e Malhotra (2009) e Tubino (2009) algumas ferramentas que auxiliam antever a produção são os modelos de previsões estatísticos. A partir dos resultados de previsões, é possível adotar estratégias que auxiliarão tomar decisões adequadas em relação a produção do produto.

Sendo assim, esse trabalho tem como finalidade realizar uma previsão do cenário agrícola do trigo no estado do Paraná, utilizando métodos estatísticos de previsão.

2. Métodos de previsão quantitativos

Os métodos de previsão quantitativos estão distribuídos em dois grupos: projeções de séries temporais; e correlação e regressão (THOMAS, 1996; LEE; SONG; MJELDE, 2008).

As previsões baseadas em correlação são conhecidas como uma abordagem causal devido ao fato de buscarem prever o futuro de uma variável dependente em função de uma ou mais variáveis independentes (ELZAYED; BOUCHER, 1994; BALLOU, 2006).

Com base na definição de Elzayed e Boucher (1994) e Ballou (2006), pode-se dizer que o número de sacos de trigos (variável dependente) que um produtor produz, pode estar relacionado com variáveis independentes, que poderiam ser: o clima e solo da região, o preço que o produto é vendido, a renda per capta, entre outros.

Um outro exemplo, com base na obra de Tubino (2009), é poder haver relação da demanda de farinha de trigo (variável dependente) com as vendas das máquinas de fazer pão (variável independente), ou seja, pode-se dizer que as máquinas de fazer pão “causam” as vendas de farinha de trigo.

A análise dessas variáveis, de acordo com Lustosa et al. (2008), é realizada pelo método de regressão. E para que essas previsões sejam realizadas, Tubino (2009) relata que é necessário coletar dois tipos de dados: o histórico de dados da variável dependente, e o histórico de dados da variável independente.

Tubino (2009) ainda aponta que, caso exista apenas duas variáveis, esse modelo é conhecido por regressão simples, e caso exista mais de duas variáveis, ele é conhecido por regressão múltipla.

Para Ballou (2006), a grande dificuldade encontrada nessa técnica de previsão é detectar variáveis que são realmente causais e que conduzem a variável prevista no tempo.

Já as projeções baseadas em séries temporais, diferente dos modelos de correlação e regressão, não recebem influências de outras variáveis (TUBINO, 2009). Essas projeções partem do princípio de que o futuro poderá ser previsto baseando-se em dados passados (ELZAYED; BOUCHER, 1994; FERNANDES; FILHO, 2010).

Segundo Moreira (2008), uma série temporal é uma sequência de observações sobre o que se pretende prever ao longo do tempo, ou seja, o tempo é a única variável que influenciará nesse modelo de previsão.

Conforme Makridakis, Wheelwright e Hyndman (1998) para que possa escolher um dos métodos de previsão baseado em séries temporais, inicialmente deve-se reconhecer o padrão de comportamento da série.

O comportamento da série é reconhecido plotando-se os dados passados em um gráfico de dispersão (FERNANDES; FILHO, 2010). Esse gráfico, segundo Corrêa e Corrêa (2012), poderá conter tendência, sazonalidade e variações irregulares.

O efeito tendência em uma série temporal, é definido por Moreira (2008), como um comportamento dos dados crescerem ou decrescerem ao longo do tempo. O efeito sazonal, segundo Tubino (2009), refere-se as variações cíclicas que a série sofre em um período de tempo. As variações irregulares, são definidas por Moreira (2008), como variáveis resultantes de causas não identificadas.

3. Metodologia

Para atender os objetivos deste trabalho, coletou-se dados brutos de área plantada e área colhida entre os anos de 1990 e 2015, diretamente da base de dados do IBGE, para o estado do Paraná.

Esses dados foram analisados estatisticamente e graficamente para identificar o comportamento da série e escolher o modelo adequado de previsão.

Para a análise estatística da série temporal e realização das previsões, utilizou-se o software Action Stat® (2017), este software de previsão é pago, foi desenvolvido pela empresa Estatcamp. Dentre suas diferentes funções, para este estudo o software foi utilizado para prever séries temporais.

Em seguida realizou-se uma análise descritiva e de normalidade dos dados e os resultados estatísticos de previsão foram apresentados.

4. Resultados

Nesta seção contém os resultados obtidos com a análise dos dados históricos de área plantada e área colhida do grão de trigo para o estado do Paraná. Inicialmente foi realizado uma análise estatística que descreve os dados e posteriormente apresentou a escolha do modelo de previsão e os resultados dessa previsão.

4.1 Análise descritiva da área plantada

Os dados de hectares de área plantada do trigo no Estado do Paraná foram coletados da base de dados do IBGE. O Quadro 1 apresenta os dados de área plantada do grão de trigo da lavoura temporária no estado do Paraná, expresso em hectares, referente aos anos de 1990 até 2015.

Variável = Área Plantada (Hectares)			
Produto das lavouras temporárias = Trigo (em grão)			
Ano	Hectares	Ano	Hectares
1990	1826626	2003	1255216
1991	1144471	2004	1358692
1992	1183143	2005	1275869
1993	722504	2006	885163
1994	739824	2007	821789
1995	639828	2008	1123807
1996	1107079	2009	1308792
1997	959714	2010	1172820
1998	963100	2011	1062543
1999	755828	2012	793448
2000	850633	2013	997656
2001	961649	2014	1388388
2002	1142078	2015	1320984

Fonte: IBGE (2017)

Quadro 1 – Área plantada do grão de trigo para o Estado do Paraná, expressa em hectares

Com base nos dados apresentados na Quadro 1, realizou-se uma análise descritiva dos mesmos (Tabela 1).

Área Plantada em Hectares	
Média	1.067.756
Mediana	1.084.811
Mínimo	639.828
Máximo	1.826.626
Desvio Padrão	265.832
Coefficiente de Variação	0,248963
Amplitude	1.186.798

Fonte: Autoria própria

Tabela 1 – Análise Descritiva dos dados de área plantada do trigo no Estado do Paraná

Através da análise descritiva do preço do trigo para o estado do Paraná, percebe-se que as medidas de tendência central estão bem próximas. Quanto as medidas de variabilidade, a diferença entre o maior e o menor valor do conjunto de dados (amplitude) é de 1.186.798 hectares, o valor do desvio padrão em relação aos dados observados foi de 265.832, indicando a existência de dispersão nos dados. O coeficiente de variação dos dados (0,248963) também deixa evidente que existe variação.

Com o objetivo de analisar a normalidade dos dados, realizou-se também a representação gráfica do conjunto de dados através do Histograma e do Box-plot (Figura 1).

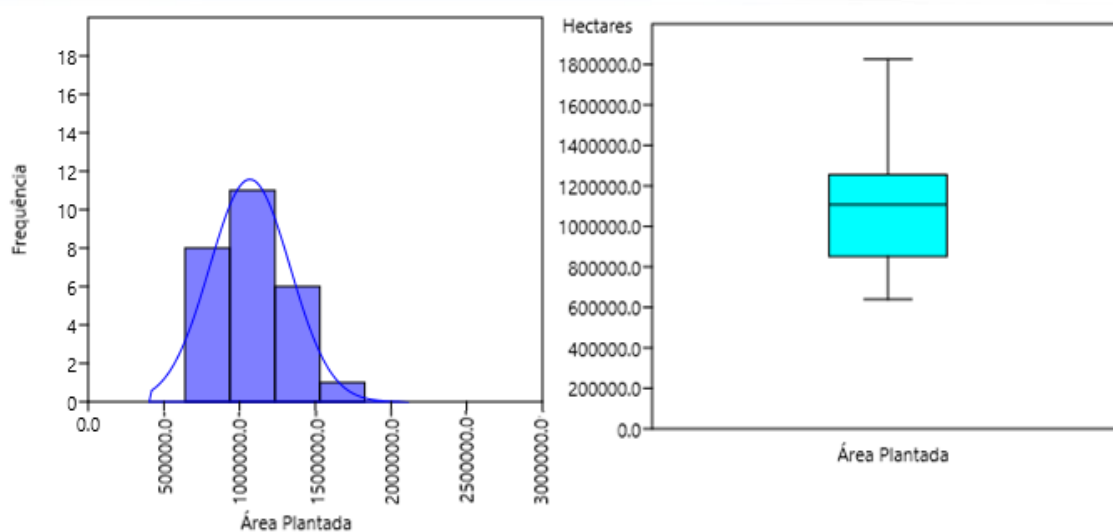


Figura 1 – Histograma com curva de distribuição normal e Box-plot dos dados de área plantada do trigo no Paraná

Fonte: Autoria Própria

Observando o Histograma da Figura 1 sobreposto por uma curva normal percebe-se que este encontra-se em formato de sino, porém a posição da linha mediana do Box-plot não está exatamente no meio do retângulo, indicando uma leve presença de assimetria nos dados.

Analisando o Box-plot, identifica-se que esta leve assimetria se encontra na parte inferior da figura, portanto existem dados discrepantes com valores menores, isso explica a média ser menor que a mediana.

4.2 Escolha do método de previsão para a área plantada

Para a escolher o modelo de previsão adequado a série de dados analisada foi necessário identificar se os dados apresentavam ciclicidade, tendência e sazonalidade. Portanto, para possível identificação desses parâmetros, os dados foram plotados em um gráfico e podem ser visualizados na Figura 2.

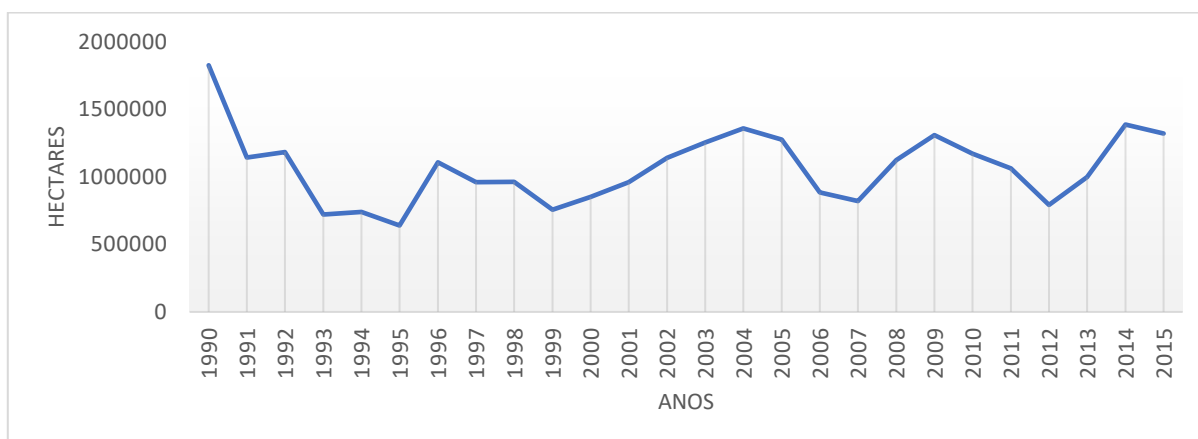


Figura 2 – Gráfico dos dados de área plantada do trigo referente aos anos de 1990 a 2015

Fonte: Autoria Própria

A Figura 2 apresentou leves ciclos de variação de área plantada do trigo para cada ano, não há evidências de tendências de crescimento ou declínio de área, como também de sazonalidade. Com a observação dessas características, os métodos escolhidos para a previsão do modelo são os métodos de suavização exponencial simples.

Para a escolha do método adequado dentre os modelos de suavização exponencial utilizou-se o Software Action Stat.

4.3 Previsão para área plantada

A partir da escolha do método de suavização exponencial, como os dados eram anuais o Action Stat gerou a previsão da demanda para o próximo ano, a fim de compará-la aos dados reais de área plantada do trigo. As previsões consideraram os valores para α (0,0001), β (0), γ (0) (Tabela 2).

Constantes da Suavização	
Nível	0,00010018
Tendência	0
Sazonalidade	0
Amortecimento	0

Fonte: Autoria própria

Tabela 2 – Constantes de suavização utilizadas para gerar as previsões de área plantada do grão de trigo para o Estado do Paraná

O valor de α indica uma mínima variação de nível, o valor de β indica que não existe variação de tendência, e a valor de γ indica que não há variação nos padrões de sazonalidade. Os dados para previsão são anuais e isso explica a ausência de sazonalidade no modelo.

As previsões realizadas foram comparadas com os dados reais para identificar a confiabilidade do modelo, portanto, foram coletados dados reais do ano de 2016 diretamente da base de dados do IBGE, sobre áreas plantadas. O gráfico da Figura 3 ilustra essa comparação.

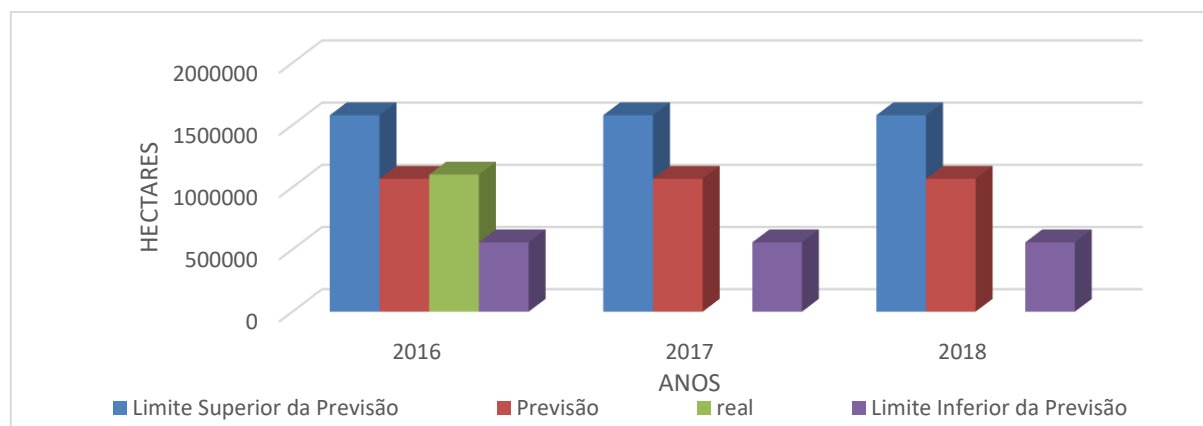


Figura 3 – Previsão de área plantada X real

Fonte: Autoria Própria

Através da Figura 3, é notável que os valores previstos estão bem próximos do valor real da área plantada, os valores previstos não estão abaixo dos limites inferiores e nem acima dos limites superiores, confirmando que o modelo escolhido foi eficaz para a previsão.

Como o modelo apresenta dados estacionários, apesar de diminuir ou aumentar os dados de um ano para o outro, a tendência da área de plantio é se manter constante ao longo dos anos dentro dos limites de previsão. Portanto, para casos como este onde os dados são anuais é importante que a previsão seja feita a cada ano para manter a aproximação com os dados reais.

4.4 Análise descritiva da área colhida

Como no item 4.1 deste trabalho, os dados de hectares de área colhida do trigo no Estado do Paraná foram coletados da base de dados do IBGE, para posteriormente serem agrupados e retirado informações importantes para realizar a análise estatística dos mesmos.

No Quadro 2, é possível ver os dados de área colhida, em hectares, do grão de trigo da lavoura temporária no estado do Paraná, os dados são dos anos de 1990 até 2015.

Variável = Área Colhida (Hectares)			
Produto das lavouras temporárias = Trigo (em grão)			
Ano	Hectares	Ano	Hectares
1990	1197149	2003	1254125
1991	1138302	2004	1358692
1992	1183143	2005	1275869
1993	683390	2006	767566
1994	630314	2007	821789
1995	636516	2008	1106396
1996	1085444	2009	1308792
1997	954269	2010	1172820
1998	952047	2011	1027936
1999	753513	2012	793448
2000	489920	2013	862000
2001	961649	2014	1388111
2002	1115455	2015	1318214

Fonte: IBGE (2017)

Quadro 2 – Área colhida do grão de trigo para o Estado do Paraná, expressa em hectares

Com base nos dados do Quadro 2, realizou-se uma análise descritiva, esta que por sua vez é apresentada na Tabela 3.

Analisando a Tabela 3, vê-se que as medidas de tendência central estão relativamente próximas. Porém, com a análise das medidas de variabilidade, percebe-se a presença de variabilidade nos dados, onde: a amplitude da amostra é de 898.191, o desvio padrão dos dados é de 253.920, e o coeficiente de variação, se é 0,251627.

Área Colhida em Hectares	
Média	1.009.110
Mediana	1.056.690
Mínimo	489.920
Máximo	1.388.111
Desvio Padrão	253.920
Coefficiente de Variação	0,251627
Amplitude	898.191

Fonte: Autoria própria

Tabela 3 – Análise Descritiva dos dados de área colhida do trigo no Estado do Paraná

Para confirmar a existência de oscilações nos dados, fez-se uma representação gráfica através de um histograma e um box-plot (Figura 4).

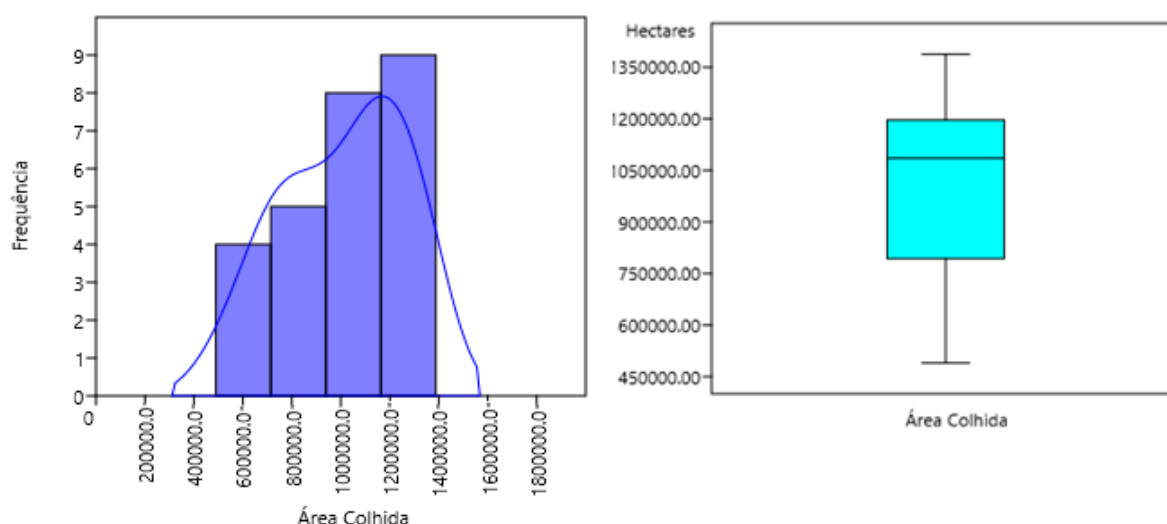


Figura 4 – Histograma com curva de distribuição normal e Box-plot dos dados de área plantada do trigo no Paraná

Fonte: Autoria Própria

O Histograma sobreposto por uma curva normal e o Box-plot da Figura 4 apresentam-se assimétricos.

Analisando o histograma percebe-se que assimetria do gráfico está localizada no lado esquerdo, e através do box-plot percebe-se que a assimetria se encontra na parte inferior, portanto existem dados discrepantes com valores menores, essa assimetria com valores menores está confirmada também através das medidas de tendência central, onde a média apresentou valor menor que a mediana.

Em nenhum dos gráficos de box-plot aqui apresentados foram observadas presenças de outliers, ou seja, as séries de dados analisadas não apresentaram valores atípicos com grandes

afastamentos dos demais dados.

4.5 Escolha do método de previsão para os área colhida

Os dados brutos coletados foram plotados em um gráfico, objetivando a identificação de métodos que se adequem melhor ao comportamento da série de dados (Figura 5).

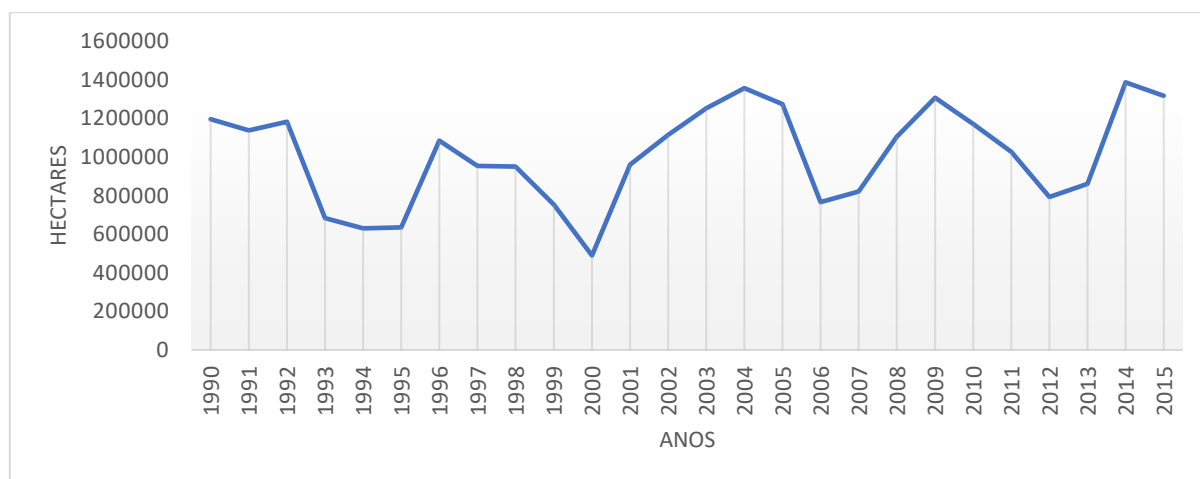


Figura 5 – Gráfico dos dados de área colhida do trigo referente aos anos de 1990 a 2015

Fonte: Autoria Própria

Assim como na Figura 2, ao ser analisada a Figura 5, é identificado a presença de ciclos de variação ao longo dos anos, ausência tendências de crescimento ou declínio de área colhida, e ausência de sazonalidade.

A Figura 6 apresenta uma comparação dos dados de área plantada com os dados de área colhida ao longo dos 26 anos analisados.

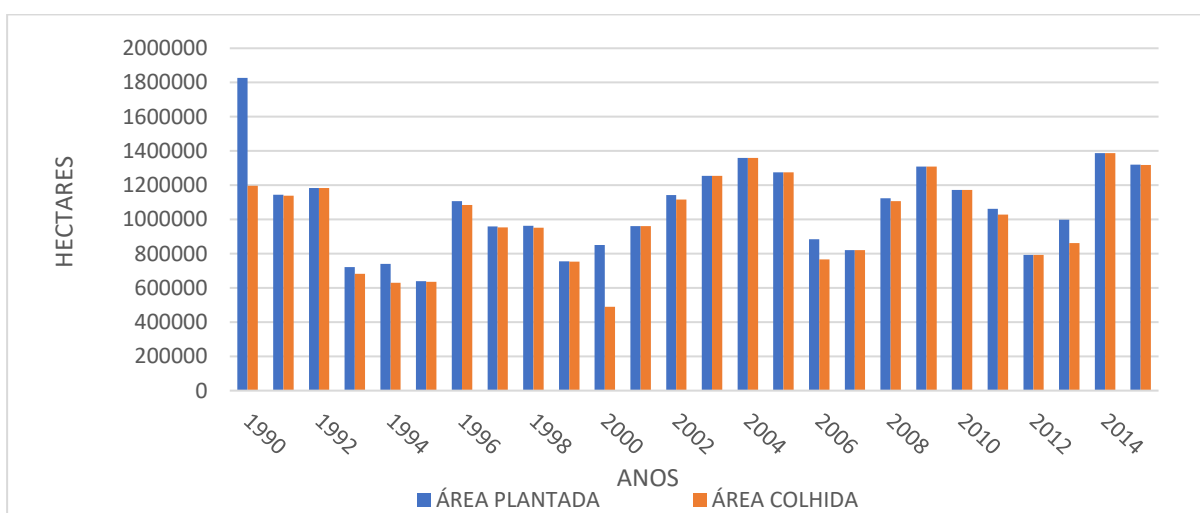


Figura 6 – Gráfico de comparação dos dados de área plantada e área colhida do trigo referente aos anos de 1990 a 2015

Fonte: Autoria Própria

Por meio desse gráfico, é possível identificar que praticamente toda a área plantada ao longo dos anos foram colhidas, isso explica a semelhança entre o comportamento da Figura 2 e Figura 5. Algumas exceções mais relevantes em relação a eficiência da colheita estão nos anos de 1990, 1994, 2000, 2006 e 2013. É possível ver também que os maiores registros de área colhida ocorreram no ano de 2004.

Com a observação dessas características, os métodos escolhidos para a previsão do modelo, assim como para área plantada, são os métodos de suavização exponencial.

4.6 Previsão para área colhida

Para área colhida, o método de previsão escolhido foi o mesmo utilizado para prever a área plantada, devido a semelhança dos dados. As previsões foram geradas considerando α (0,0001), β (0), γ (0), como mostra a Tabela 4.

Constantes da Suavização	
Nível	0,000100032
Tendência	0
Sazonalidade	0
Amortecimento	0

Fonte: Autoria própria

Tabela 4 – Constantes de suavização utilizadas para gerar as previsões de área colhida do grão de trigo para o Estado do Paraná

O valor de α indica uma mínima variação de nível, os valores de β e γ indicam ausência de variação de tendência e sazonalidade.

As previsões realizadas das áreas colhidas foram comparadas com os dados reais para identificar a confiabilidade do modelo, portanto, foram coletados dados reais do ano de 2016 diretamente da base de dados do IBGE. Essa comparação é apresentada na Figura 7.

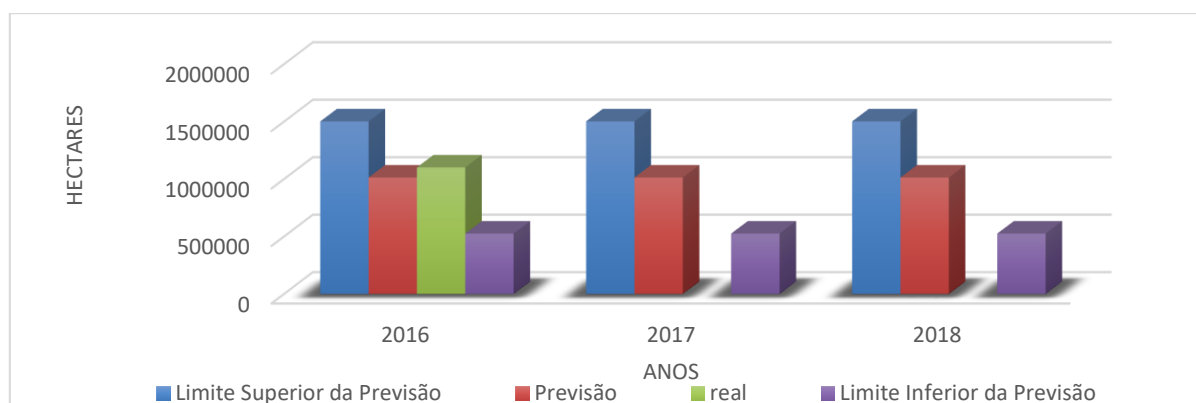


Figura 7 – Previsão de área colhida X real

Fonte: Autoria Própria

Como visto na análise descritiva dos dados e na escolha do método de previsão para área plantada e área colhida, os dados se apresentavam semelhantes, ou seja, praticamente foi colhido toda a área plantada ao longo dos anos.

Os valores previstos e o valor real para o ano de 2016 de área colhida estão bem próximos, essa previsão não ultrapassa os limites inferiores e superiores da previsão, sendo assim, este modelo é eficaz para esta previsão.

Como no modelo proposto para área plantada, este de área colhida apresenta dados estacionários, apesar de diminuir ou aumentar os dados de um ano para o outro, a tendência média da área de colheita é se manter constante ao longo dos anos.

5. Conclusão

O objetivo deste trabalho foi realizar previsões de área plantada e área colhida do grão de Trigo no estado do Paraná.

Os resultados de previsões de área plantada e área colhida apresentaram grandes semelhanças, indicando que a maior parte das áreas que foram plantadas ao longo dos anos também foram colhidas, com exceções em alguns períodos que apresentaram quedas nas colheitas.

Nas séries de dados das áreas plantadas e colhidas foi identificada a presença de ciclos, porém com valores muito baixos, ausência de tendência e sazonalidade. A sazonalidade não estava presente na série dados em decorrência dos dados serem coletados anualmente. Os valores das séries analisadas apresentaram comportamento estacionário, a tendência de colheita e plantio é de se manter constante, como foi apresentado também nos gráficos para escolha do modelo de previsão.

Para casos como a previsão de área plantada e de área colhida, em que os dados quantitativos são anuais, recomenda-se que seja realizado uma previsão de demanda também anual, pois assim aumenta a confiabilidade do que se deseja prever.

O software Action Stat mostrou-se eficaz na realização de todas as previsões estatísticas que foram propostas, pois apresentou resultados de previsão dentro dos limites esperados.

Agradecimentos

Agradecimento à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos de Mestrado.

Referências.

ACTION STAT®. *O que é Action Stat?* Disponível em: < <http://www.portalaction.com.br/sobre-o-action> > Acesso em 11 de maio de 2017.

BALLOU, R. H. *Gerenciamento da cadeia de suprimentos/logística empresarial.* Porto Alegre: Bookman, 2006.

CEPEA, Centro de Estudos Avançados em Economia Aplicada. *o PIB do Agronegócio Brasileiro.* Disponível em: <<http://www.cepea.esalq.usp.br/br/pib-do-agronegocio-brasileiro.aspx>> Acesso em 11 de abril de 2017.

CONAB, Companhia Nacional de Abastecimento. *Acompanhamento da safra Brasileira.* Disponível em: <http://www.conab.gov.br/OlalaCMS/uploads/arquivos/16_01_12_09_00_46_boletim_graos_janeiro_2016.pdf> Acesso em 5 de maio de 2017.

CONAB, Companhia Nacional de Abastecimento. *A cultura do trigo.* Organizador Paulo Magncorro Rabelo. – Brasília: Conab, 2013. Disponível em:

http://www.conab.gov.br/OlalaCMS/uploads/arquivos/17_05_03_16_09_46_a_cultura_do_trigo-versao_digital_nova_logo.pdf. Acesso em: Maio, 2017.

CORRÊA, H. L.; CORRÊA, C. A. *Administração de Produção e Operações*. São Paulo: Atlas, 2012.

ELZAYED, E.; BOUCHER, T. *Analysis and controlo f production systems*. 2. Ed., New Jersey: Prentice-Hall, 1994.

FERNANDES, F. C. F.; FILHO, M. G. *Planejamento e controle da produção*. São Paulo: Atlas, 2010.

KRAJEWSKI, L. J.; RITSMAN L. P.; MALHOTRA M. K. *Administração de Produção e Operações*. São Paulo, 2009.

LEE, C. K.; SONG, H. J.; MJELDE, J. W. *The forecasting of international expo tourism using quantitative and qualitative techniques*. Tourism Management. v. 29, p. 1084-1098, 2008.

LUSTOSA, L.; MESQUITA, M. A.; QUELHAS, O.; OLIVEIRA, R. *Planejamento e controle da Produção*. Rio de Janeiro: Editora Elsevier, 2008.

MAKRIDAKIS, S.; WHEELWRIGHT. S. C.; HYNDMAN, R. J. *Forecasting: methods and applications*. 3ª ed., New York: John Wiley & Sons, 1998.

MOREIRA, D. A. *Administração da Produção e Operação*. São Paulo: Cengage Learning, 2008.

THOMAS, R. J. *Estimating demand for services: Issues in combining sales forecasts*. Journal of Retailing and Consumer Services. v. 3, p. 241 – 250, 1996.

TUBINO, D. F. *Planejamento e Controle da Produção: Teoria e Prática*. São Paulo: Atlas, 2009.