

Modelagem matemática aplicada ao rendimento de grãos do milho

Mathematical modeling applied to corn grain yield

RESUMO

Rodrigo Matheus Ritter

ritter.rodrigo4@gmail.com

Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil

Araceli Ciotti de Marins

araceli@utfpr.edu.br

Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil

Anderson Alves Miguel

ander.alves.miguel123@gmail.com

Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil

Luiz Gabriel Martins

luizgabrielmartins2014@hotmail.com

Universidade Tecnológica Federal do Paraná, Toledo, Paraná, Brasil

O objetivo deste trabalho foi encontrar, por meios probabilísticos, um modelo matemático que melhor representasse o comportamento do rendimento de grãos do milho em função de atributos químicos do solo. Foram utilizadas técnicas de regressão linear e não linear, análise de componentes principais e métodos de qualidade de ajustamento, além de pesquisas sobre o comportamento da produtividade das culturas agrícolas a depender da fertilidade do solo. Foram encontrados um modelo linear e quatro modelos não-lineares cujos coeficientes de ajuste foram relativamente próximos. Escolheu-se o modelo mais adequado através do teste de significância dos parâmetros, o qual indicou o modelo linear como o mais consistente para ilustrar a relação entre rendimento e o par de atributos químicos do solo mais influentes no rendimento, sendo estes atributos o Fósforo e o Potássio.

PALAVRAS-CHAVE: Modelos de regressão. Produtividade agrícola. Fertilidade do solo.

ABSTRACT

The objective of this work was to find, by probabilistic means, a mathematical model that best represents the behavior of corn grain yield in function of soil chemical attributes. Linear and non-linear regression techniques, principal component analysis and adjustment quality methods were used, in addition to research on the productivity behavior of agricultural crops dependent on soil fertility. One linear model was found and four nonlinear models fit coefficients were relatively close. The most appropriate model was chosen through the parameter significance test, which indicated the linear model as the most consistent to illustrate the relationship between yield and pair of soil chemical attributes most influential in yield, these attributes being Phosphorus and Potassium.

KEYWORDS: Regression Models. Agricultural productivity. Soil fertility.

Recebido: 19 ago. 2020.

Aprovado: 01 out. 2020.

Direito autorial: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



INTRODUÇÃO

Conhecendo a perspectiva da sociedade, sabe-se que o setor produtivo busca meios de produção com sustentabilidade para obter lucro e, ao mesmo tempo, satisfazer a demanda. No setor agrícola, identificar os fatores relacionados ao rendimento e compreender como eles interferem na produtividade, bem como aprimorar técnicas de manejo do solo, são de suma importância, fundamentais para que se produza em maior escala sob a mesma proporção de terra, visto que, por princípios de proteção ambiental, não é possível a expansão das lavouras.

O rendimento das culturas, seja de grãos ou de massa fibrosa, está ligado à diversos fatores, como pluviosidade, aspectos físicos e, principalmente, a fertilidade do solo. Esta última pode ser verificada pela presença de atributos químicos específicos.

Na região Sul do Brasil, maior produtora de milho do país, embora tenha havido um aumento de 2,14% da área cultivada, houve uma redução de 11,98% na produção, passando de 11.813.000 toneladas na primeira safra 2018/2019 para 10.389.000 toneladas na primeira safra 2019/2020, observando-se assim uma redução de 13,82% por hectare (CONAB, 2020). Deste modo, considera-se importante o conhecimento acerca dos fatores que contribuem para a rentabilidade da lavoura, em especial do milho, um dos cereais mais produzidos no Brasil e um dos pilares da economia brasileira.

Estudos realizados por Andrade et al. (2012) revelam como os sistemas de plantio, por rotação e sucessão de culturas, influenciam nos teores dos elementos químicos presentes no solo a cada safra. Portanto, compreender como os atributos químicos do solo interferem no rendimento de grãos de determinada cultura é indispensável, pois este conhecimento pode ser um fator determinante no momento da tomada de decisões sobre o manejo do solo.

Uma das maneiras de se avaliar a influência dos atributos químicos no rendimento de grãos é com a utilização da modelagem matemática, que segundo Silva Jr. et al. (2012), tem auxiliado no planejamento agrícola quando associada ao conhecimento da variabilidade dos atributos do solo, ao mapeamento das formas do relevo e índices de fertilidade e à rotação de culturas.

Assim, este trabalho buscou ajustar modelos de regressão não-linear para ilustrar a produção de grãos do milho de acordo com o par de elementos químicos mais influentes, em condições climáticas correspondentes à média de precipitação de chuvas na lavoura onde foi realizada a coleta dos dados.

MATERIAIS E MÉTODOS

Os dados aqui analisados, provém de análise laboratorial dos dados coletados para a tese de doutorado de Marins (2014) realizado em Cascavel – PR. De posse dos dados (teores dos elementos químicos), foi realizada a técnica de modelagem matemática intitulada Regressão Não-linear, pela qual foram ajustadas funções que estimam a relação entre o rendimento de grãos do milho e o par de atributos químicos mais influentes nessa relação, por processos iterativos de Gauss-Newton (ZEVIANI; RIBEIRO JR; BONAT, 2013).

Para encontrar o par de componentes mais influentes, ou ainda, que mais se correlacionam, foi efetuada a análise de componentes principais, através do *software RStudio*, versão 8.3. Segundo Kassambara (2017), o método trabalha com a combinação dos dados originais e tem como características: identificar padrões ocultos; reduzir a dimensão, através da diminuição da redundância dos dados e identificar variáveis correlacionadas.

Com o *RStudio* foram obtidos os modelos matemáticos e também plotados os gráficos aqui apresentados. Nele, os dados foram distribuídos por característica e efetuou-se regressão não-linear considerando o método dos mínimos quadrados. Utilizou-se o comando *nls ()* para obter os modelos não-lineares e *lm ()* para a obtenção do modelo linear (ZEVIANI, 2012). No primeiro caso, foi preciso efetuar dois chutes iniciais, a função de iteração e os valores iniciais dos parâmetros. Para isso, por meio do *software GeoGebra*, conforme Dantas e Matucheski (2019), os dados foram inseridos como pontos e criados controles deslizantes para cada parâmetro, com o intuito de poder escolher um bom chute inicial para os parâmetros de cada modelo que, por meio do *GeoGebra*, foram construídos e analisados para a escolha das funções de iteração mais convenientes.

Em seguida, os modelos foram levados ao *RStudio* para o início do processo de regressão. Os modelos analisados foram os modelos de crescimento de Mattos (2013), e o Modelo Bidimensional, sugerido por Ritz e Streibig (2008), adequados para modelos com um par de variáveis explicativas pois, para estes dados, os modelos de rendimento necessitam de valores iniciais muito específicos, o que impossibilitou a compilação da programação. Os modelos de analisados foram:

- a) Modelo de Regressão Linear Múltipla;

$$\hat{R}(x, y) = \alpha \cdot x + \beta \cdot y + c \quad (1)$$

- b) Modelo de Mitscherlich;

$$\hat{R}(x, y) = \alpha(1 - 10^{(\beta(x-\gamma(y+\delta)))}) \quad (2)$$

- c) Modelo Bidimensional;

$$\hat{R}(x, y) = \frac{\alpha}{1 + \beta(x + \gamma \cdot y)} \quad (3)$$

- d) Modelo Logístico;

$$\hat{R}(x, y) = \frac{\alpha}{1 + e^{(\beta - \gamma(x + \delta \cdot y))}} \quad (4)$$

- e) Modelo de Gompertz.

$$\hat{R}(x, y) = \alpha e^{(-e^{(\beta - \gamma(x + \delta \cdot y))})} \quad (5)$$

onde \hat{R} é a função de estimação, $\alpha, \beta, \gamma, \delta$ são as estimativas dos parâmetros, x, y as variáveis explicativas (independentes) e c uma constante.

Quando diferentes modelos de regressão não-linear são ajustados a um mesmo conjunto de dados, é importante a utilização de ferramentas estatísticas para compará-los e indicar o melhor modelo, os chamados avaliadores de qualidade de ajuste. Aqui, para avaliar a qualidade do ajuste utilizou-se o R^2 e a estatística t , que, segundo Thomas (2016), mede a significância dos parâmetros do modelo em relação ao erro padrão. Caso o valor da estatística t de um parâmetro não for significativa, recomenda-se a retirada deste parâmetro e que o ajuste seja refeito.

RESULTADOS E DISCUSSÃO

Efetuada a análise de componentes principais, através do *RStudio*, para obter o par de atributos químicos que mais influenciam conjuntamente no rendimento de grãos do milho, obteve-se dos dados da Tabela 1.

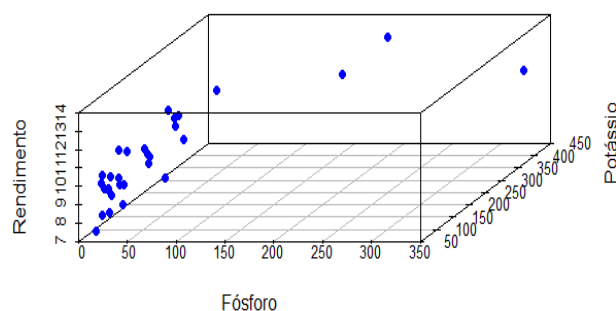
Tabela 1 – Dados da análise de componentes principais

Componente	Autovalor	Variância (%)	Variância acumulada (%)
Rendimento	2,8931004	48,218340	48,218340
P	1,3436083	22,393472	70,61181
K	0,8274557	13,790928	84,40274
M.O.	0,4252072	7,086786	91,48953
Al	0,3293404	5,489007	96,97853
Ca	0,1812880	3,021466	100,0000

Fonte: Autores (2020).

Ao desconsiderar os valores atribuídos ao rendimento, pois é uma variável dependente, fica claro que o par de elementos com maior influência são **P** (Fósforo) e **K** (Potássio), cuja relação com o rendimento está ilustrado na Figura 1.

Figura 1 – Plotagem dos dados amostrais



Fonte: Autores (2020).

Posteriormente foi efetuado o processo de regressão linear múltipla conforme a Eq. (1), obtendo-se os resultados expressos na Tabela 2.

Tabela 2 – Resultados da regressão linear múltipla

Coefficientes	Estimativa	P
(Intercepto)	8,093651	1,22e-15***
P	0,006916	0,0633''
K	0,006771	0,0183*

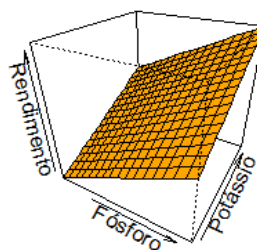
Significado dos códigos: 0'***' 0,001'***' 0,01'**' 0,05 '' 0,1 ''''

Fonte: Autores (2020).

Segundo Wagner (1998), o teste *t*, expresso estatisticamente por *p*, é um teste de significância estatística que traduz o quanto a hipótese testada é compatível com os dados observados. Nele, os resultados são considerados significativos quando o teste *t* é menor ou igual à 0.05, ou seja, quanto menor é

resposta do teste, mais o parâmetro em teste é influente na variação da variável resposta. Isto nos mostra que os parâmetros relacionados à P e K exercem influência significativa sobre o modelo. Portanto, este modelo indica que o rendimento depende de maneira semelhante de ambos os atributos, com um coeficiente de explicação R^2 de 51,2737%. Na Figura 2 está o gráfico desta relação.

Figura 2 – Modelo de regressão linear múltipla



Fonte: Autores (2020).

Porém, como o rendimento não apresenta retorno linear e contínuo, buscou-se alternativas por meio de regressão não-linear múltipla. Partindo da Eq. (2), foi refeito o processo iterativo de ajuste obtendo-se os resultados da Tabela 3.

Tabela 3 – Resultados da regressão pelo modelo de Mitscherlich

Coeficientes	Estimativa	p
α	1,275e+01	7,85e-14***
β	5,122e-03	0,298
γ	1,312e-01	0,381
δ	5,755e+02	0,160

Significado dos códigos: 0'***' 0,001'***' 0,01'***' 0,05'***' 0,1'***'

Fonte: Autores (2020).

Porém, no teste t verifica-se que somente o primeiro parâmetro é significativo para o modelo em relação aos dados. Entretanto, há apenas um parâmetro significativo para o modelo, se for retirado algum parâmetro, o estudo perde o sentido, pois todos estão diretamente ligados às variáveis independentes e, mesmo apresentando R^2 de 59,0965% o modelo não expressa com convicção a relação entre o rendimento e os atributos. Partindo então para o modelo da Eq. (3), foram obtidos os resultados expressos na Tabela 4.

Tabela 4 – Resultados da regressão pelo modelo Bidimensional

Coeficientes	Estimativa	P
α	8,2807351	< 2e-16***
β	-0,0002825	0,270
γ	2,2900153	0,399

Significado dos códigos: 0'***' 0,001'***' 0,01'***' 0,05'***' 0,1'***'

Fonte: Autores (2020).

Este modelo apresentou R^2 de 48,2867%. Repare que, assim como o caso do modelo de Mitscherlich, não expressa a relação com convicção.

Outro modelo avaliado foi o modelo logístico, Eq. (4), que apresentou o melhor resultado de R^2 sendo este 59,6434% de explicação dos dados.

Tabela 5 – Resultados da regressão pelo modelo Logístico

Coeficientes	Estimativas	P
α	12,72504	4,04e-15***
β	-0,43800	0,052''
γ	0,01475	0,266
δ	0,14401	0,382

Significado dos códigos: 0'***' 0,001'***' 0,01'***' 0,05 '.' 0,1 ''''
 Fonte: Autores (2020).

Entretanto, é possível observar que mesmo havendo dois parâmetros com maior significância, são parâmetros não ligados diretamente às variáveis, o que também torna este modelo inconsistente.

Por fim, analisou-se o modelo de regressão não-linear de Gompertz, Eq. (5), o qual teve como resultados de saída os dados abaixo:

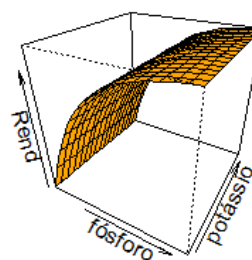
Tabela 6 – Resultados da regressão pelo modelo de Gompertz

Coeficientes	Estimativa	P
α	12,73583	1,57e-14***
β	-0,67141	0,000913***
γ	0,01326	0,279608
δ	0,13733	0,381285

Significado dos códigos: 0'***' 0,001'***' 0,01'***' 0,05 '.' 0,1 ''''
 Fonte: Autores (2020).

Este modelo apresentou R^2 de 59,38% e dois parâmetros muito significativos, porém, estes parâmetros também não estão vinculados às variáveis independentes. Observe que, comparado aos dados, é um modelo muito atraente.

Figura 3 – Modelo de Gompertz



Fonte: Autores (2020).

Perceba que, mesmo que os modelos aparentem ilustrar bem o comportamento do rendimento dos grãos de milho, os coeficientes de explicação e as medidas do teste baixos podem ser explicados pela densidade de pontos com índices baixos de F, fazendo com que os pontos com os maiores teores de Fósforo tenham menor influência sobre o ajuste no modelo, aumentando os resíduos.

De acordo com Machado et al. (2007), grandes áreas agrícolas são consideradas homogêneas na agricultura brasileira, sendo aplicadas doses de corretivos considerando apenas média necessária, ou seja, há um desbalanço no uso de fertilizantes, podendo comprometer o rendimento de grãos. Assim, justifica-se o conhecimento do comportamento do rendimento de grãos em função dos atributos químicos pois, de posse desse conhecimento, se pode tomar medidas mais regionalizadas para evitar desperdícios e gastos desnecessários.

Para eles, além de conhecer as características biofísicas das plantas, é preciso conhecer que regiões da lavoura estão debilitadas e quais atributos químicos estão faltando, para que, ao se adquirir fertilizantes, o produtor possa mapear sua lavoura e aplicar a correção de teor de maneira mais precisa, intensificando a aplicação nas regiões de maior necessidade, de modo a evitar desperdícios e maximizar o rendimento e qualidade do cereal.

Conforme Andrade et al. (2012), os métodos e sistemas de plantio influenciam de maneira direta no rendimento, uma vez que proporcionam características físicas e químicas diferentes à próxima cultura e, como consequência, cada método e sistema de plantio requer um manejo específico. Em seus estudos, os autores constataram que, justamente o par de elementos mais influentes aqui apresentados, são os mais controláveis considerando a rotação e sucessão de culturas, e que o sistema de rotação por plantio direto foi o método com resultados mais significativos para isto.

CONCLUSÕES

Um manejo adequado do solo é essencial para garantir a produtividade da lavoura. Como visto, o fósforo e o potássio são essenciais para isso. Na dose inadequada deles, mesmo que se tenha condições pluviométricas e físicas ideais, o rendimento de grãos do milho pode estar comprometido.

A interpretação do estudo aqui apresentado e a escolha do modelo mais adequado para a estimação do rendimento do milho fica a cargo do leitor, porém, estatisticamente, sugere-se o modelo linear como mais conciso para ilustrar o rendimento de grãos, pois, além de os parâmetros ligados aos teores dos atributos químicos serem significativos, o coeficiente de explicação do modelo não está distante dos demais modelos.

Analisando o gráfico do modelo de regressão linear múltipla, é possível concluir que depende de forma semelhante de ambos os atributos, e que o rendimento máximo está onde o teor de Fósforo é 2,1% maior que o teor de Potássio.

AGRADECIMENTOS

Agradecemos os nossos familiares pelo apoio na iniciativa e pela compreensão com os momentos de dedicação, à prof.^a Dr.^a Daniela Trentin Nava pela ajuda com o *software RStudio*, à orientadora Prof.^a Dr.^a Araceli Ciotti de Marins pela dedicação, à UTFPR pelo fomento da bolsa, tornando possível a realização deste trabalho.

REFERÊNCIAS

- ANDRADE, et al. Atributos químicos de um Cambissolo Húmico após 12 anos sob preparo convencional e semeadura direta em rotação e sucessão de culturas. **Ciência Rural**, Santa Maria, v.42, n.5, p. 814-821, mai. 2012.
- CONAB. **Acompanhamento da safra brasileira de grãos**. v.5 – SAFRA 2019/2020 – n.11 – Décimo primeiro levantamento, 2020.
- DANTAS, S. C; MATUCHESKI, S. Resolução de um problema com o uso de diferentes ferramentas do GeoGebra. **Pesquisa e debate em educação**, v.9, p. 588-602, 2019.
- KASSAMBARA, A. 2017. **Practical guide to principal component methods in R**. 1 ed. USA: STHDA.
- MACHADO, et al. Variabilidade espacial de atributos químicos do solo em áreas sob Sistema Plantio Convencional. **Revista brasileira de ciência do solo**. v. 31, p. 591-599, 2007.
- MARINS, A. C. **Rendimento de grãos e de óleo do crambe em um latossolo: análise espacial e modelos de correlação**. 2014. 161 f. Tese (Doutorado) – Curso de Ciência do Solo, Programa de Pós-Graduação em Ciência do Solo, Universidade Federal de Santa Maria, Santa Maria, 2014. Disponível em: <https://repositorio.ufsm.br/bitstream/handle/1/3359/MARINS%2C%20ARACELI%20CIOTTI%20DE.pdf?sequence=1&isAllowed=y> . Acesso em: 25 abr. 2019.
- MATTOS, T. B. **Modelos não lineares e suas aplicações**. 58 f. Monografia (Curso de Estatística). UFJF, Juiz de Fora, 2013.
- RITZ, C; STREIBIG, J. C. **Nonlinear regression with R**_Use R! Copenhagen: Springer, 2008.
- SILVA JR, et al. Classificação numérica e modelo digital de elevação na caracterização espacial de atributos dos solos. **Revista brasileira de engenharia agrícola e ambiental**, v.16, n.4, p.415–424, 2012.
- THOMAS, G. **Regressão não linear**. 17 f. Relatório Final (Disciplina de Regressão e Covariância – Pós-Graduação em Estatística e Experimentação Agropecuária). ESALQ-USP, 2016.

WAGNER, M. B. Significância ou confiança? **Jornal de pediatria**. Rio de Janeiro, p. 343-346. ago. 1998. Disponível em:
<https://www.lume.ufrgs.br/bitstream/handle/10183/54344/00096544.pdf?sequence=1&isAllowed=y> . Acesso em: 06 jun. 2020.

ZEVIANI, W. M. et al. Modelos não-lineares para a liberação de potássio de esterco animal em latossolos. **Ciência Rural**, Santa Maria, v.42, n.10, p. 1789-1796, out. 2012. Disponível em:
http://www.scielo.br/scielo.php?script=sci_arttext&pid=S01038478201200100012&lng=en&nrm=isso. Acesso em: 12 abr. 2020.

ZEVIANI, W. M.; RIBEIRO JR, P. J.; BONAT, W. H. **Modelos de regressão não linear**. Laboratório de Estatística e Geoinformação. Departamento de Estatística. UFPR, 2013. Disponível em: <http://leg.ufpr.br/~walmes/cursoR/mrnl2013/>. Acesso em: 25 abr. 2020.