

Regressão linear múltipla envolvendo variáveis biológicas e ambientais na simulação de indicadores da composição química de grãos de aveia

Odenis Alessi¹
DCEEng/UNIJUÍ, Ijuí, RS
Rubia Diana Mantai²
DCEEng/UNIJUÍ, Ijuí, RS
José Antonio Gonzales da Silva³
DEAg/DCEEng/UNIJUÍ, Ijuí, RS
Renata Linassi Bárta⁴
DEAg/UNIJUÍ, Ijuí, RS
Vanessa Pansera⁵
DCEEng/UNIJUÍ, Ijuí, RS
Adriana Rosélia Kraisig⁶
DCEEng/UNIJUÍ, Ijuí, RS
Júlio Daronco Berlezi⁷
DEAg/UNIJUÍ, Ijuí, RS
Ester Mafalda Matter⁸
DEAg/UNIJUÍ, Ijuí, RS

Resumo. Modelos de regressão linear múltipla podem simular com eficiência as alterações da concentração de nutrientes em grãos da aveia. O objetivo do trabalho é definir as variáveis potenciais para construção de regressões lineares múltiplas voltadas a simulação de variáveis ligadas a qualidade nutricional de grãos da aveia em função de doses de nitrogênio e de elementos meteorológicos. O delineamento experimental foi de blocos casualizados com quatro repetições, seguindo um esquema fatorial 4 x 2 para os fatores doses de nitrogênio, e cultivares, respectivamente, nos anos agrícolas de 2011 a 2016. Foi utilizado a técnica de *Stepwise* para selecionar as variáveis independentes que influenciam de forma significativa na variável dependente, para assim compor os modelos de regressão linear múltipla. As variáveis selecionadas para compor os modelos se mostraram eficientes para simular a concentração de proteína, fibra e amido contidos nos grãos de aveia. Os modelos encontrados mostraram simulações adequadas para gerar prognósticos de qualidade química dos grãos.

Palavras-chave. Matemática Aplicada, Seleção de variáveis, Qualidade Nutricional, *Avena sativa*

¹BOLSISTA CAPES/BRASIL, odenisalessi@hotmail.com

²rdmantai@yahoo.com.br

³jagsfaem@yahoo.com.br

⁴rlinassibarta@gmail.com

⁵vpansera@hotmail.com

⁶maryshelei@yahoo.com.br

⁷julio.berlezi28@hotmail.com

⁸estermafaldamatter@gmail.com

1 Introdução

Modelos matemáticos de regressão linear múltipla são utilizados em estudos de produções eficientes, com potencial de explicar por meio de uma equação dinâmica a relação e a magnitude do efeito de variáveis independentes sobre uma variável dependente, desenvolvendo modelos de simulação que auxiliam na previsibilidade do comportamento dos sistemas [1, 4]. A eficiência destes modelos matemáticos está associada a escolha das variáveis independentes mais expressivas sobre as variáveis dependentes [10, 11]. A escolha das variáveis ocorre com a utilização de técnicas de seleção, que acabam qualificando a concepção dos modelos de regressão linear múltipla. A técnica *Stepwise* é um dos métodos mais utilizados de seleção de variáveis, pois o método seleciona de forma iterativa as variáveis que mais influenciam o conjunto de saída, excluindo possíveis redundâncias [2, 7]. O procedimento de seleção ou exclusão de variáveis pelo modelo é baseado em um algoritmo que verifica as estatísticas F parciais, testando a importância de cada uma com a variável dependente [7].

Considerado um cereal de alto teor nutricional, os grãos de aveia são ricos em proteínas, possuem altos teores de fibras alimentares, lipídio, minerais e grande quantidade de amido [6, 13]. Entretanto, a composição química e nutricional da aveia sofre influências das condições meteorológicas e das práticas de manejo [5, 12]. Dentre estas práticas culturais, a adubação nitrogenada se destaca, visto que, além de possuir efeito sobre a produtividade de grãos, também atua sobre a expressão do teor de proteína e fibra dos grãos [8, 9]. Além disto, as variações meteorológicas, influenciam no conteúdo de proteína, fibra, amido e óleo da cariopse de aveia. [3, 14].

Nesse contexto, modelos de regressão linear múltipla podem simular com eficiência as alterações da qualidade nutricional de grãos da aveia levando em consideração o fornecimento de nitrogênio e a não linearidade dos elementos meteorológicos, gerando prognósticos na promoção de uma produção eficiente com maior qualidade à alimentação humana e ampla aceitação comercial. Desta forma, o objetivo do trabalho é definir as variáveis potenciais para o desenvolvimento de regressão linear múltipla voltadas a simulação de variáveis ligadas a qualidade nutricional de grãos da aveia em função de doses de adubação nitrogenada e elementos meteorológicos.

2 Metodologia

O trabalho foi desenvolvido nos anos agrícola de 2011 a 2016 na área experimental do Instituto Regional de Desenvolvimento Rural (IRDeR) pertencente ao Departamento de Estudos Agrários (DEAg) da UNIJUÍ. O delineamento experimental adotado foi de blocos casualizados com quatro repetições, seguindo um esquema fatorial 4 x 2 para os fatores doses de nitrogênio, e cultivares, respectivamente. Os níveis do fator dose de nitrogênio foram assim determinados: 0 kg ha⁻¹, 30 kg ha⁻¹, 60 kg ha⁻¹ e 120 kg ha⁻¹; e utilizadas duas cultivares de aveia, sendo elas: Barbarasul e Brisasul. A adubação nitrogenada sucedeu em uma única aplicação na condição de terceira folha expandida. O momento de colheita de grãos ocorreu aproximadamente aos 120 dias de período de cultivo, em que os grãos apresentavam umidade próxima de 15%. A colheita dos experimentos ocorreu de forma manual, pelo corte das três linhas centrais de cada parcela, que após trilhadas com colheitadeira estacionária, foram direcionadas ao laboratório para correção da umidade de grãos para 13% e posterior pesagem para estimativa da produtividade de grãos, convertida para a unidade de um hectare. Os dados das variáveis ligadas a qualidade nutricional de grãos da aveia, sendo elas, a proteína total, a fibra total e o amido, foram quantificadas através de amostras de 300 grãos não descascados, provenientes de cada parcela útil. Estas amostras foram encaminhadas para análises químicas realizadas através da espectrofotometria do infravermelho proximal - NIR (Near infrared Reflectance), no laboratório de bromatologia da UNIJUÍ. As variáveis meteorológicas como a temperatura mínima, média e máxima e precipitação pluviométrica, foram obtidas através da

estação meteorológica automatizada, instalada a 200 metros do experimento. A variável soma térmica foi obtida pelo somatório da diferença entre a temperatura média diária e a temperatura basal inferior ao desenvolvimento da cultura da aveia (4°C), no ciclo de cultivo. Foi utilizado a técnica *Stepwise* como critério de seleção das variáveis mais ajustadas aos modelos de regressão linear múltipla, excluindo possíveis redundâncias. Verificada a distribuição normal das variáveis dependentes, proteína total, fibra total e amido, foram desenvolvidos três modelos de regressão linear múltipla, representado pelo modelo geral:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \tag{1}$$

em que: X_p é a p-ésima variável observada, β_p é o coeficiente associado à p-ésima variável e ε é o erro. Todas as técnicas estatísticas foram aplicadas utilizando o *software* livre GENES. Para realização da simulações com os modelos de regressão linear múltipla obtidos, foram utilizados valores médios dos dados meteorológicos ocorridos para cada dose de nitrogênio. Estes dados foram aplicados aos modelos a fim de simular a variável de interesse e validar a sua eficiência através do erro absoluto dado pela diferença entre o valor observado em campo e o valor simulado pelo modelo.

3 Resultados e discussão

A Tabela 1 traz os valores médios reais da proteína, fibra total e amido dos grãos de aveia obtidos por bioexperimentação, por condição de dose de adubação nitrogenada em cobertura junto com as informações meteorológicas em diferentes anos de cultivo da aveia.

Tabela 1: Valores médios da proteína total, fibra total e do amido, junto à dados meteorológicos da temperatura, soma térmica e precipitação pluviométrica

Ano	Dose de N (kg ha ⁻¹)				Amplitude Térmica		Temperatura Média (°C)			Σ ST	Σ PP
	0	30	60	120	Min	Máx	T _{min}	T _{med}	T _{max}	(Graus _{dia})	(mm)
Proteína Total (g kg ⁻¹)											
2011	105	107	112	119	-3,5	34,0	10,1	16,6	22,4	1529	767
2012	103	106	110	120	-2,8	31,6	11,5	16,9	22,2	1671	698
2013	98	102	103	107	-3,9	33,0	8,7	15,0	21,2	1383	510
2014	99	101	103	106	-2,0	36,0	11,9	17,4	22,9	1674	844
2015	108	109	110	112	-0,2	31,7	12,3	17,2	22,0	1672	651
2016	95	97	102	115	-0,4	34,7	9,9	16,7	23,5	1590	625
Geral	101	104	107	113	-2,1	33,5	10,7	16,6	22,4	1597	683
Fibra Total (g kg ⁻¹)											
2011	129	123	122	121	-3,5	34,0	10,1	16,3	22,4	1592	767
2012	137	136	133	132	-2,8	31,6	11,5	16,9	22,2	1671	698
2013	119	118	117	115	-3,9	33,0	8,7	15,0	21,2	1383	510
2014	133	132	130	129	-2,0	36,0	11,9	17,4	22,9	1674	844
2015	134	130	128	128	-0,2	31,7	12,3	17,2	22,0	1672	651
2016	128	126	124	123	-0,4	34,7	9,9	16,7	23,5	1590	625
Geral	130	128	126	125	-2,1	33,5	10,7	16,6	22,4	1597	683
Amido (g kg ⁻¹)											
2011	443	444	446	448	-3,5	34,0	10,1	16,3	22,4	1592	767
2012	458	456	450	461	-2,8	31,6	11,5	16,9	22,2	1671	698
2013	431	434	434	435	-3,9	33,0	8,7	15,0	21,2	1383	510
2014	403	403	406	406	-2,0	36,0	11,9	17,4	22,9	1674	844
2015	445	449	453	454	-0,2	31,7	12,3	17,2	22,0	1672	651
2016	409	412	422	420	-0,4	34,7	9,9	16,7	23,5	1590	625
Geral	432	433	435	437	-2,1	33,5	10,7	16,6	22,4	1597	683

T_{min} = temperatura média mínima; T_{med} = temperatura média; T_{max} = temperatura média máxima; Min=temperatura mínima do ciclo; Máx= temperatura máxima do ciclo; Σ ST = soma térmica;Σ PP = precipitação pluviométrica.

A partir dos dados apresentados na Tabela 1 é possível observar a influência das doses de nitrogênio em aumentar ou reduzir os níveis dos compostos químicos dos grãos. Percebe-se que os teores de proteína total aumentam a medida que há um incremento na dose de adubação nitrogenada. O comportamento oposto é observado para os teores de fibra total, havendo redução da concentração de fibras com o aumento nas doses de nitrogênio. As condições meteorológicas também alteraram a qualidade nutricional dos grãos, sendo o teor de proteína mais reduzido em anos em que houve a ocorrência de temperaturas mais amenas ou mais elevadas, seguido de reduzida ou alta precipitação pluviométrica. Desta forma, este comportamento indica que o ambiente adequado para a variável proteína requer temperaturas amenas e volume de chuvas em quantidade suficiente para favorecer suprimento de umidade no solo.

É evidenciado na Tabela 1, que os menores teores de fibra dos grãos são encontrados em condições ambientais de baixa temperatura e reduzida precipitação pluviométrica. Este fato pode ser explicado no sentido de que temperaturas mais baixas e menor precipitação, desde que bem distribuídas, incrementam o número de grãos, condição que indica que o particionamento de energia na elaboração das fibras pelo maior número de grãos pode reduzir a concentração do nutriente.

O maior teor de amido nos grãos de aveia foi encontrado em condições com temperaturas elevadas e chuvas amenas, sendo que, o excesso de chuvas mostrou prejudicar a sua quantidade. Contudo, apesar do comportamento da concentração de amido não ter sido crescente em função das doses do nitrogênio em todos os anos considerados, o seu comportamento geral demonstrou uma tendência de incremento da concentração total pelo aumento do fornecimento do nitrogênio.

Na Tabela 2 é apresentada a seleção das variáveis pela técnica *Stepwise* e os respectivos modelos de regressão linear múltipla estabelecidos para cada componente nutricional, independente de ano agrícola.

Tabela 2: Modelo de regressão linear múltipla da proteína total, fibra total e amido, em função de variáveis selecionadas pelo modelo *Stepwise*, independente de ano agrícola.

Fonte de Variação	Quadrado médio Modelo <i>Stepwise</i>	Regressão Múltipla
(2011 + 2012 + 2013 + 2014 + 2015 + 2016)		
Regressão	5461*	
DN	3700*	
\bar{T}_{min}	472,5*	
\bar{T}_{max}	767,2*	$PT = 124,36 + 0,10DN - 3,77\bar{T}_{min} - 4,92\bar{T}_{max} + 0,08\Sigma ST$
\bar{T}_{med}	ns	
ΣST	521,5*	
ΣPP	ns	
Regressão	5847*	
DN	704,7*	
\bar{T}_{min}	95,2*	
\bar{T}_{max}	83,2*	$FT = 70,62 - 0,04DN + 0,47\bar{T}_{min} - 0,73\bar{T}_{max} + 0,05\Sigma ST$
\bar{T}_{med}	ns	
ΣST	521,5*	
ΣPP	ns	
Regressão	44348*	
DN	1059,8*	
\bar{T}_{min}	7664,6*	
\bar{T}_{max}	17538,6*	$AM = 642,39 + 0,05DN - 17,12\bar{T}_{min} - 24,58\bar{T}_{max} + 0,37\Sigma ST - 0,10\Sigma PP$
\bar{T}_{med}	ns	
ΣST	9167,2*	
ΣPP	8916,9*	

DN = Doses de Nitrogênio ($kg\ ha^{-1}$); PT = Proteína total ($g\ kg^{-1}$); FT = Fibra Total ($g\ kg^{-1}$); AM = Amido ($g\ kg^{-1}$); \bar{T}_{min} = temperatura média mínima; \bar{T}_{med} = temperatura média; \bar{T}_{max} = temperatura média máxima; ΣST = soma térmica; ΣPP = precipitação pluviométrica; * = significativo a 5% de probabilidade de erro pelo teste F; ns= não significativo.

Na simulação da proteína total e da fibra total, as variáveis indicadas para compor o modelo

de regressão linear múltipla foram a dose de nitrogênio, a temperatura mínima, a temperatura máxima e a soma térmica. As variáveis temperatura média e precipitação não demonstraram influência significativa na concentração final destes nutrientes nos grãos de aveia. Para a simulação do amido, as variáveis dose de nitrogênio, temperatura mínima, temperatura máxima, soma térmica e precipitação pluviométrica se mostraram indicadas para a construção da equação.

Analisando os modelos encontrados é explicitado a magnitude e sentido da influência de cada variável independente sobre as variáveis dependentes. Por exemplo, percebe-se que existe um incremento de concentração de proteína e amido em grãos de aveia a medida em que ocorre o aumento da dose da adubação nitrogenada e um decremento para a concentração de fibra, indo ao encontro das análises feitas para os dados da Tabela 1.

A Tabela 3 traz os resultados das simulações para o teor de proteína total, fibra total e amido dos grãos de aveia, a partir dos modelos desenvolvidos, nas distintas condições de uso de nitrogênio e independente do ano agrícola.

Tabela 3: Simulação da proteína total, fibra total e amido pela regressão linear múltipla em função da dose de nitrogênio e variáveis meteorológicas, independente de ano agrícola

Ano	DN (kg ha ⁻¹)	\bar{T}_{min} (°C)	\bar{T}_{max} (°C)	ΣST (mm)	V_O (g kg ⁻¹)	V_S (g kg ⁻¹)	Erro Absoluto (g kg ⁻¹)
$(PT = 124,36 + 0,10DN - 3,77\bar{T}_{min} - 4,92\bar{T}_{max} + 0,08\Sigma ST)$							
	0				102	102	0
	30				103	104	1
Geral	60	10,7	22,3	1597	107	108	1
	120				113	114	1
$(FT = 70,62 - 0,04DN + 0,47\bar{T}_{min} - 0,73\bar{T}_{max} + 0,05\Sigma ST)$							
	0				130	129	1
	30				128	128	0
Geral	60	10,7	22,3	1597	126	127	1
	120				125	124	1
$(AM = 642,39 + 0,05DN - 17,12\bar{T}_{min} - 24,58\bar{T}_{max} + 0,37\Sigma ST - 0,10\Sigma PP)$							
	0				432	434	2
	30				433	435	2
Geral	60	10,7	22,3	1597	435	437	2
	120				438	440	2

V_O = Valor Observado; V_S = Valor Simulado; DN = Dose de Nitrogênio; \bar{T}_{min} = temperatura média mínima; \bar{T}_{med} = temperatura média; \bar{T}_{max} = temperatura média máxima; ΣST = soma térmica; ΣPP = precipitação pluviométrica.

Observando os valores de erros absolutos encontrados, percebe-se que em alguns casos, os valores simulados foram iguais aos valores observados. Os demais valores encontrados através das simulações variam entre 1 e 2 g kg⁻¹ de erro absoluto, não apresentando grande diferença dos valores reais observados e indicando eficiência dos modelos encontrados. Desta forma, os resultados das simulações obtidas evidenciam que os modelos de regressão linear múltipla apresentam valores simulados muito próximos aos observados, indicando possibilidade de simulações destes compostos químicos com grande qualidade, envolvendo a complexidade das condições ambientais.

4 Conclusões

O uso das variáveis doses de nitrogênio, temperatura mínima e máxima, soma térmica e precipitação pluviométrica, se mostram eficientes para compor modelos de regressão linear múltipla para simulação da proteína total, da fibra total e do amido em grãos de aveia. O desenvolvimento

de modelos de regressão linear múltipla em função das doses de nitrogênio e elementos meteorológicos mostram simulações adequadas para a determinação da concentração de proteína, fibra e amido contidos nos grãos de aveia, gerando prognósticos da concentração destes nutrientes em grãos de aveia.

Agradecimentos

Ao CNPq, à CAPES, à FAPERGS e à UNIJUÍ pelos recursos financeiros, estrutura física e materiais necessários para realização da pesquisa.

Referências

- [1] Abbad, G., Torres C. V. Regressão múltipla stepwise e hierárquica em Psicologia Organizacional: aplicações, problemas e soluções. *Estudos de Psicologia*, 7:19-29, 2002. DOI:10.1590/S1413-294X2002000300004.
- [2] Alves, M. F., Lotufo, A. D. P., Lopes, M. L. M. Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas. *Proceeding Series of the Brazilian Society of Applied and Computational Mathematics*, 11-6, 2013. <https://doi.org/10.5540/03.2013.001.01.0144>.
- [3] Beber, R. C., Francisco, A. de, Alves, A. C., Sá, R. M. de, Gliari, P. Caracterização química de genótipos brasileiros de aveia (*Avena sativa* L.). *Acta Científica Venezolana*, 53:202- 209, 2002.
- [4] Carvalho Junior, W. de, Calderano Filho, B., Chagas, C. da S., Bhering, S. B., Pereira, N. R., Pinheiro, H. S. K. Regressão linear múltipla e modelos random forest para estimar a densidade do solo em áreas montanhosas. *Pesquisa Agropecuária Brasileira*, 51:1428-1437, 2016. DOI:10.1590/S0100-204X2016000900041.
- [5] Crestani, M., Carvalho, F. I. F., Oliveira, A. C., Silva, J. A. G., Gutkoski, L. C., Sartori, J. F., Baretta, D., Luche, H. S., Tessmann, E. W., Paiva, R. P. Desempenho de cultivares de aveia branca quanto ao conteúdo de B-glicana no grãos conduzidas em diferentes ambientes. In: XXX Reunião da Comissão Brasileira de Pesquisa de Aveia, 2010, São Carlos-SP. 1:127-131. 2010.
- [6] Hawerth, M. C., Silva, J. A. G., Woyabb, L. G., Zimmer, C. M., Groli, E. L., Oliveira, A. C., Carvalho, F. I. F. Correlations among industrial traits in oat cultivars grown in different locations of Brazil. *Australian Journal of Crop Science*, 9:1182-1189, 2015.
- [7] Huda, S., Abawajy, J., Abdollahian, M., Islam, R., Yearwood, J. A fast malware feature selection approach using a hybrid of multi-linear and stepwise binary logistic regression. *Concurrency And Computation: Practice And Experience*, 29, 2016. <https://doi.org/10.1002/cpe.3912>.
- [8] Kolchinski, E. M., Schuch, L. O. B. Relações entre a adubação nitrogenada e a qualidade de grãos e de sementes em aveia branca. *Ciência Rural*, 34:379-383, 2004.
- [9] Lima, K. R., Carvalho, C. A. B., Azevedo, F. H. V., Malafaia, P. A. M. Sward structure and nutritive value of *Urochloa ruziziensis* under nitrogen and potassium fertilisation. *Revista Caatinga*, 30:220-229, 2017.

- [10] Marolli, A. Silva, J. A. G., Mantai, R. D., Brezolin, A. P., Gzergorczyk, M. E., Lambrecht, D. M. Oat yield through panicle components and growth regulator. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 21:261-266, 2017. DOI:10.1590/1807-1929/agriambi.v21n4p261-266.
- [11] Prunzel, J., Toebe, M., Lopes, A. B., Moreira, V. S. Multiple Linear Regression Models Applied to Urban Terrain Assessment - Municipality of Itaqui - RS Case. *Boletim de Ciências Geodésicas*, 22:651-664, 2016. DOI:10.1590/s1982-21702016000400037.
- [12] Scremin, O. B., Silva, J. A. G., de Mamann, A. T. W., Marolli, A., Mantai, R. D., Trautmann, A. P. B., Kraisig, A. R., Scremin A. H., Kruger, C. A. M. B., Dorenelles, E. F. Nitrogen and hydrogel combination in oat grains productivity. *International Journal of Development Research*, 7:13896-13903, 2017.
- [13] Singh, S., Kaur, M., Sogi, D. S. A comparative study on chemical composition, color, and functional characteristics of flours and protein concentrates from different oat cultivars. *International Journal of Recent Scientific Research*, 9:2666-26671, 2018. DOI: <http://dx.doi.org/10.24327/ijrsr.2018.0905.2100>.
- [14] Zute, S., Brunava, L., Jansone, I., Tamm, Y. Influence of farming system, nitrogen rates and meteorological conditions on Naked oat in Latvia. *Agriculture & Forestry*, 61:15-25, 2015.