



## CALIBRAÇÃO E VALIDAÇÃO DE MODELO BIOGEOFÍSICO PARA PREVISÃO DE CRESCIMENTO E PRODUTIVIDADE AGRÍCOLA

**Palavras-chave:** Modelo biofísico, validação de modelo, previsão, Linguagem R.

**Autores/as:**

*Bianca Barbosa Schorles* [IMECC/Unicamp]

*Guilherme Vieira Nunes Ludwig (Orientador)* [IMECC/Unicamp]

*Santiago Vianna Cuadra (Co-orientador)* [Embrapa]

---

### 1 Introdução

Ao longo das próximas décadas, o cultivo agrícola irá aumentar devido à crescente demanda pela produção de alimentos, dado o crescimento populacional e do consumo, assim como do crescimento exponencial do cultivo das culturas agrícolas voltadas à produção de biocombustíveis (veja, entre outros, Schmitz et al., 2014; Austin et al., 2022). Essas tendências são particularmente aplicáveis ao Brasil (Valdes et al., 2016; de Miranda, 2018), que vem apresentando uma crescente expansão das áreas de lavouras temporárias e de pastagem plantadas nas últimas décadas-segundo dados oficiais do IBGE (Instituto Brasileiro de Geografia e Estatística).

Uma das principais formas de se avaliar os impactos das mudanças climáticas na produtividade agrícola é através do uso de modelos computacionais de produtividade agrícola. No processo de evolução do Zonamento Agrícola de Risco Climático, modelos baseados em processos biogeofísicos têm sido incorporados nas avaliações de risco, dentre eles o modelo AgS (Agricultural crop Simulator). Os zoneamentos agrícolas realizados pela Embrapa subsidiam as portarias do ZARC. Essas portarias definem o calendário de plantio de todos os contratos de seguro agrícola do Governo Federal. Só um desses instrumentos, o PROAGRO<sup>2</sup>, segurou na safra 2018/2019 um valor segurado total superior a R\$ 13 bilhões. O ZARC também serve de base técnica para o Garantia Safra e Programa de Subvenção ao Prêmio do Seguro Rural - está previsto o recurso de R\$ 1 bilhão para o PSR na próxima safra de 2020.

O AgS simula não apenas as componentes do balanço hídrico, mas também o acúmulo de biomassa e rendimento das culturas em função da disponibilidade hídrica, desenvolvimento fenológico em função da temperatura do ar e interceptação da radiação solar em escala diária. O ajuste do modelo requer um processo estatístico de calibração e validação. Para tornar esse processo mais eficiente e obter melhores ajustes a equipe do Laboratório de Modelagem Agroambiental da Embrapa Agricultura Digital está desenvolvendo um Workflow para automatizar a leitura dos dados observados, execução e calibração do modelo. Neste projeto, analisamos o cultivo de soja, representado pelo seguinte ciclo:

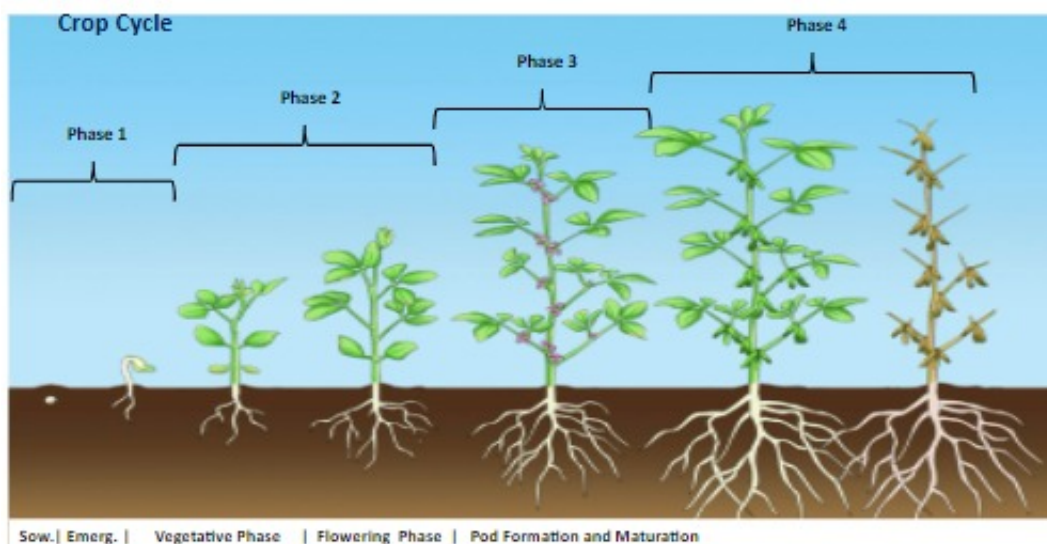


Figura 1: Crescimento e desenvolvimento da planta (fonte: Oliveira, M., 2023, comunicação pessoal)

## 2 Metodologia

A partir do banco de dados, análise da programação enviada pela Embrapa, reuniões e materiais de apoio, nota-se que o modelo AgS é um modelo de séries temporais com dados diários e foi construído para mapear o crescimento e desenvolvimento da agricultura de soja, utilizando variáveis como: biomassa, estresse hídrico, temperatura, gás carbônico, entre outras. Ademais, seus parâmetros são equações diferenciais (equação que apresenta como incógnita uma função que aparece na forma de derivada) que controlam as condições feitas no modelo para que a planta se desenvolva, sendo necessário agregar um erro aleatório dentro do ciclo para identificar determinada observação.

Assim, foi utilizado o método de superfície de resposta com o objetivo de analisar o resultado do modelo AGS, a partir das informações dos experimentos de cultivo de soja. O método consiste em uma técnica estatística utilizada para a modelagem e análise de problemas nos quais a variável resposta é influenciada por vários fatores, por exemplo: temperatura e estresse hídrico, cujo objetivo é a otimização dessa resposta, neste caso, a variável resposta é o erro (*'rmse'*) gerado pelo modelo.

O pacote computacional utilizado foi o RSM, presente na ferramenta R-Studio, que constrói a metodologia em dois passos: inicialmente, assume um modelo de regressão para identificar a relação entre as variáveis de interesse e após o mapeamento do comportamento da variável resposta em relação às outras variáveis selecionadas.

## 3 Resultados e Conclusões

Para a construção das análises foram feitos cenários com diferentes variáveis explicativas em relação ao erro (variável resposta), assim, quando selecionado uma determinada variável as outras foram fixadas para não interferir no comportamento da variável escolhida. Na visualização desses processos, foram feitos gráficos de superfície para avaliar o desempenho do modelo.

No exemplo abaixo, a comparação foi feita entre as variáveis: TSUM e HWAD que avaliam as condições da planta ao passar pelo modelo. Com isso, após a manipulação dos dados para a execução do processo, foi feita uma modelagem com o modelo de regressão entre elas. Os resultados dos coeficientes se encontram na tabela 1.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1096.1657	1.3451	814.95	0.0000
poly(TSUM, HWAD, degree = 2)1.0	32776.0706	527.2709	62.16	0.0000
poly(TSUM, HWAD, degree = 2)2.0	7068.1408	527.2709	13.41	0.0000
poly(TSUM, HWAD, degree = 2)0.1	0.0000	527.2709	0.00	1.0000
poly(TSUM, HWAD, degree = 2)1.1	0.0000	206690.2036	0.00	1.0000
poly(TSUM, HWAD, degree = 2)0.2	0.0000	527.2709	0.00	1.0000

Tabela 1: Modelo de Regressão: TSUM x HWAD

Com os coeficientes calculados, é possível associar as variáveis observadas e gerar a superfície de resposta que mostra o comportamento do erro conforme a variação de TSUM e HWAD em um contexto tridimensional.

### Superfície de resposta

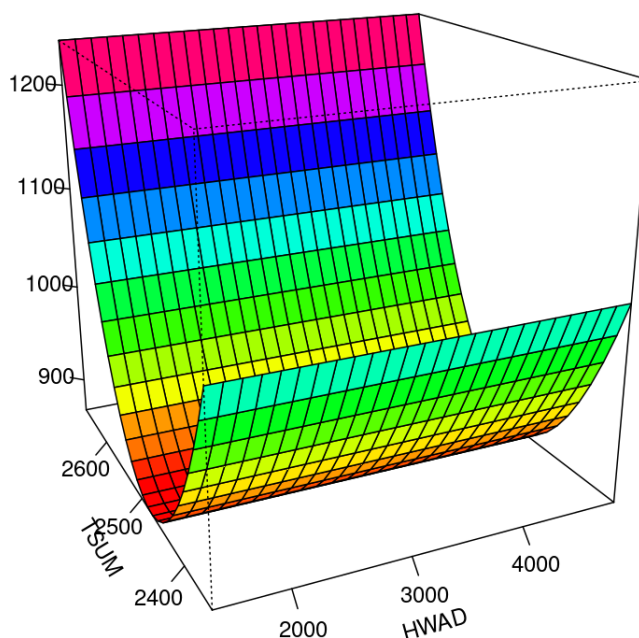


Figura 2: Superfície de resposta: TSUM x HWAD

A Figura 2 apresenta o gráfico do comportamento do erro gerado no modelo, assim, pode ser feita a validação e calibração a partir da distribuição do erro na superfície. Em princípio, quanto menor o erro gerado no modelo, melhor é representado o comportamento dos dados analisados. Observando que o menor erro se concentra no centro da distribuição da variável TSUM e aumenta nas extremidades dos valores. Concluindo que, a cada análise de cenário e as possíveis combinações de variáveis no modelo, é possível validar e calibrar o comportamento do modelo a partir dos erros gerados.

---

## 4 Referências

- K. Austin, J. Jones, and C. Clark. A review of domestic land use change attributable to us biofuel policy. *Renewable and Sustainable Energy Reviews*, 159:112181, 2022.
- Climate Action Reserve. Requirements and guidance for model calibration, validation, uncertainty, and verification for soil enrichment projects. Technical report, 2020. URL <https://www.climateactionreserve.org/wp-content/uploads/2020/10/SEP-Model-Requirements-and-Guidance-v1.0a.pdf>.
- N. Cressie. *Statistics for Spatial Data*, 2nd edition. Wiley, New York, 1993.
- E. E. de Miranda. Áreas cultivadas no Brasil e no mundo. *Agro ANALYSIS*, 38(2):25-27, 2018.
- P. J. Diggle and P. J. Ribeiro. *Model-based Geostatistics*. Springer, New York, 2007.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL <https://www.R-project.org/>.
- C. Schmitz, H. Van Meijl, P. Kyle, G. C. Nelson, S. Fujimori, A. Gurgel, P. Havlik, E. Heyhoe, D. M. d’Croz, A. Popp, et al. Land-use change trajectories up to 2050: insights from a global agro-economic model comparison. *Agricultural Economics*, 45(1):69-84, 2014.
- C. Valdes, K. Hjort, and R. Seeley. Brazil’s agricultural land use and trade: effects of changes in oil prices and ethanol demand. Technical report, 2016.
- CHUANG , Zhaoa. Título. *A SIMPLE crop model*, 2019. Disponível em: <https://www.sciencedirect.com/journal/european-journal-of-agronomy>.