

## Aplicação do Modelo Random Forest para a Predição da Produtividade do Arroz Tropical Irrigado<sup>(1)</sup>

Igor Kuivjogi Fernandes<sup>2</sup>, David Henriques da Matta<sup>3</sup>, Germano Costa-Neto<sup>4</sup> e Alexandre Bryan Heinemann<sup>5</sup>

<sup>1</sup> Pesquisa financiada pela Embrapa Arroz e Feijão.

<sup>2</sup> Graduando em Estatística pela UFG, estagiário da Embrapa Arroz e Feijão, Santo Antônio de Goiás, GO

<sup>3</sup> Matemático e Estatístico, mestre em Matemática e Estatística, professor da Universidade Federal de Goiás, Goiânia, GO

<sup>4</sup> Engenheiro-agrônomo, doutorando em Genética pela Esalq/USP, Piracicaba, SP

<sup>5</sup> Engenheiro-agrônomo, doutor em Irrigação e Drenagem, pesquisador da Embrapa da Arroz e Feijão, Santo Antônio de Goiás, GO

**Resumo** - As previsões de produtividade agrícola, em especial do arroz irrigado, são fundamentais para o desenvolvimento das políticas agrícolas, notadamente no ambiente tropical. A possibilidade do uso de alternativas que têm a capacidade de prever o desempenho das culturas em resposta ao clima constituem estratégias que possibilitam grande avanço para a tomada de decisões na agricultura tropical. Avaliamos um método de aprendizado de máquina, denominado Random Forest (RF), por sua capacidade de prever as respostas da produção agrícola ao clima em escalas globais e regionais em arroz irrigado tropical. Foram utilizados dados de experimentos (231) provenientes do Programa de Melhoramento de Arroz Irrigado da Embrapa Arroz e Feijão, de 1995 a 2019, que fazem parte dos ensaios de valor de cultivo e uso (VCU) e que possuem data de semeadura. Esses experimentos foram ligados a variáveis climáticas relacionadas às temperaturas máximas e mínimas do ar, precipitação pluvial, radiação solar global e umidade relativa do ar, produzindo 53 variáveis relacionadas com o clima. Além disso, foram utilizadas variáveis geográficas (altitude, latitude e longitude) e variáveis relacionadas aos fenômenos El Niño, La Niña e Neutro. A acurácia do modelo RF para prever a produtividade foi de 65%. A razão da baixa acurácia está relacionada à pequena massa de dados disponíveis (231 experimentos). As variáveis que apresentaram maior capacidade de discriminar a produtividade do arroz irrigado na região tropical foram precipitação pluvial acumulada no período vegetativo (variável contínua) e UF (variável categórica).