

APLICABILIDADE E AJUSTE DA DISTRIBUIÇÃO LOG-NORMAL A 3 PARÂMETROS EM ESTUDO DE PRECIPITAÇÃO MÁXIMA DIÁRIA ANUAL NA BACIA DO RIO VERDE

RESUMO

Estimativas probabilísticas de eventos extremos são de grande relevância para o planejamento das atividades humanas e também na gestão de recursos hídricos. Neste contexto, objetivou-se analisar o desempenho da distribuição de probabilidades, Log Normal a 3 parâmetros, os quais foram ajustados pelos Métodos dos Momentos (MM) e da Máxima Verossimilhança (MV), aplicadas sobre as séries históricas de precipitação diária máxima anual de 7 estações pluviométricas na bacia hidrográfica do Rio Verde, em Minas Gerais, obtidas do HIDROWEB/ANA com pelo menos 30 anos de observação. A bacia hidrográfica do rio Verde possui uma área de 6.891,4 km² e cerca de 450 mil habitantes. Para a verificação da melhor estimativa dos parâmetros dos métodos utilizados, foram aplicados os testes de aderência de Komolgorov Smirnov e Qui-quadrado. O teste de Qui-quadrado foi mais rígido que o de Kolmogorov-Smirnov, classificando como inadequadas 43% dos modelos. Tomando este teste como base, identificou-se maior adequação (88%) na estimativa pelo método MV. Em uma análise de precisão, a série histórica de Varginha ajustada pelo método MV, apresentou o menor valor, demonstrando melhor ajuste. Em estudos com precipitação máxima anual, nesta região, a distribuição Log-normal a 3 parâmetros estimada pelo método MV pode ser mais adequada.

Palavras-chave: Distribuição de Probabilidades. Métodos de estimativa dos Parâmetros. Testes de aderência. Precipitação Máxima Diária Anual. Séries Hidrológicas.

APPLICABILITY AND ADJUSTMENT OF THE LOG-NORMAL DISTRIBUTION TO 3 PARAMETERS IN THE STUDY OF ANNUAL DAILY MAXIMUM PRECIPITATION IN THE RIO VERDE BASIN

ABSTRACT

Probabilistic estimates of extreme events are of great relevance for the planning of human activities and also for the management of water resources. In this context, we aimed to analyze the performance of the probability distribution, Log Normal to 3 parameters, which were adjusted by the Moment (MM) and Maximum Likelihood (MV) methods applied to historical series of maximum annual daily precipitation of 7 rainfall stations in the Rio Verde watershed, in Minas Gerais, obtained from

HIDROWEB / ANA with at least 30 years of observation. The catchment area of the Verde River has an area of 6,891.4 km² and about 450 thousand inhabitants. In order to verify the best estimation of the method parameters used, the adhesion tests of Komolgorov Smirnov and Chi-square were applied. The Chi-square test was more rigid in Kolmogorov-Smirnov, classifying as inadequate 43% of the models. Taking this test as base, identify greater adequacy (88%) in the estimation by the MV method. In a precision analysis, a historical Varginha series adjusted by the MV method presented the lowest value, showing a better fit. In studies with maximum annual rainfall, in this region, a Log-normal distribution of 3 parameter estimates by the MV method may be more appropriate..

Keywords: Distribution of Probabilities. Methods of estimating Parameters. Adhesion tests. Maximum Daily Maximum Precipitation. Hydrological series.

Recebido em: 15/12/2017 - Aprovado em: 10/03/2018 - Disponibilizado em: 15/07/2018

1. INTRODUÇÃO

Estimativas probabilísticas de eventos extremos são de grande relevância para o planejamento das atividades humanas. O registro do comportamento das variáveis hidrológicas ao longo do tempo possibilita estas estimativas. Para tanto, foi desenvolvida a teoria de probabilidades, possibilitando a estimativa de eventos hidrológicos associados a diferentes frequências de ocorrência, como base em modelos probabilísticos, os quais são estruturados em parâmetros estimados pela inferência estatística (NAGHETTINI & PINTO, 2007).

Dentre as características de grande interesse das precipitações está o estudo da frequência da precipitação pluvial diária máxima, que é normalmente representada por uma Função de Probabilidade que melhor se ajuste a valores extremos, sendo do tipo assintótico (MELLO & SILVA, 2005).

A estimativa dos parâmetros de cada distribuição de probabilidades pode ser realizada a partir de uma amostra das observações, onde são feitas inferências estatísticas sobre a população, por diferentes métodos. A comparação dos modelos pode ser conduzida com base em testes de aderência não paramétricos, os quais podem informar a adequabilidade do ajuste de cada modelo à série histórica de dados observados (NAGHETTINI & PINTO, 2007; CASELLA & BERGER, 1990).

Há diversas distribuições de probabilidade para valores extremos, como as chuvas máximas diárias anuais. Sendo recomendadas, de acordo com Naghettini & Pinto (2007), distribuições de Gumbel, a Generalizada de Valores Extremos (GEV), Fréchet, Weibull, Log-Normal a 2 Parâmetros e Log-Normal a 2 Parâmetros e Pearson tipo III. Sendo, que a estimativa dos parâmetros destas distribuições estatísticas podem ser encontrados por meio dos Métodos dos Momentos (MM),

Máxima Verossimilhança (MV) e Método dos Momentos L (ML).

Os métodos de ajuste dos parâmetros das distribuições de probabilidades, dentre eles o dos momentos, da máxima verossimilhança, podem conduzir a resultados diferentes, sendo, segundo NAGHETTINI & PINTO (2007), o método dos momentos (MM) o mais simples para a estimação dos parâmetros e menos eficiente do que os estimadores de máxima verossimilhança (MV), particularmente para distribuições com três parâmetros. O método da máxima verossimilhança (MV) é considerado como o método maximiza a plausibilidade da distribuição ser representada pelos parâmetros estimados. (NAGHETTINI & PINTO, 2007).

Diversos estudos têm utilizado várias distribuições de probabilidade para valores extremos, sendo destacado como as mais eficientes, devido aos melhores ajustes, a distribuição de Gumbel e a Generalizada de Valores Extremos (GEV). A distribuição Log-Normal a 3 P também sendo utilizada, devido principalmente à sua simplicidade e rapidez. A Log – Normal a 3 P apresenta um coeficiente de assimetria com um valor não fixo e sempre maior do que zero, o que, de acordo com Naghettini & Pinto (2007) fazem da distribuição log-normal uma forma paramétrica que pode se adequar muito bem à modelação de vazões e alturas de chuva máximas (ou médias) mensais, trimestrais ou anuais.

Observa-se na literatura, vasta utilização da distribuição Log-Normal a 3 parâmetros na estimativa de precipitação provável, a fim de considerar a chuva em projetos de irrigação e evitar o superdimensionamento.

Back (2001) destaca a distribuição Log-Normal a 3 parâmetros como a mais adequada para estimativas de precipitação máxima diária de Santa Catarina quando se tratam de séries de dados com baixa assimetria e curtose. O mesmo autor também considera o método MV para estimativa dos parâmetros mais adequado que o MM para esta distribuição.

Junqueira Júnior et al. (2007) caracterizaram a precipitação provável na região do município de Madre de Deus, MG, comparando quatro diferentes modelos de distribuição de frequência (Gama, Normal, Log-normal 2 e 3 parâmetros), concluíram pelo teste de Kolmogorov-Smirnov que os modelos de probabilidade adequaram-se melhor ao período chuvoso, sendo a distribuição Log-normal 3 parâmetros a mais adequada para as séries históricas de período mensal e a distribuição Gama para os períodos quinzenal e decendial.

Rodrigues et al. (2013) avaliou as distribuições gama, log-normal e Weibull na modelagem de precipitação mensal na cidade de Bento Gonçalves, Rio Grande do Sul e concluíram que as três distribuições apresentaram aderência, porém o modelo gama ajustou-se mais adequadamente às condições pluviométricas da cidade.

Caldeira et al. (2016) analisaram o ajuste das distribuições Log-Normal a 2 e 3 parâmetros e Gumbel, de séries pluviométricas de precipitação máxima diária anual de 342 estações no Rio Grande do Sul, concluindo que as três distribuições de probabilidade foram adequadas pelo teste de Qui-Quadrado, entretanto a distribuição Log-Normal a 3

parâmetros é a mais recomendada devido aos melhores ajustes dos parâmetros.

Miranda et al. (2017) utilizou-se as distribuições de probabilidade de Gumbel para máximos, Fréchet, Gama, Log-Normal 2 parâmetros e Log-Normal a 3 parâmetros visando o estudo do comportamento de modelos probabilísticos na previsão da precipitação máxima diária anual, bem como modelos de chuvas intensas para o município de Divinópolis – MG, sendo concluído que as distribuições que melhor se ajustaram aos dados foi a de Gumbel, seguida pela Log-Normal 3P e 2P.

Marques et al. (2017), utilizaram a distribuição Log-Normal a 3P pelo Método dos Momentos (MM) e Máxima Verossimilhança (MV) em séries históricas de precipitação diária máxima anual de 7 estações pluviométricas em regiões com média a alta intensidade de ocorrência de chuvas intensas e concluíram que não se pode identificar qual o melhor método de estimativa dos parâmetros, porém o teste de aderência de qui-quadrado é o mais restritivo com devendo o mesmo ser tomado como base para os estudos, principalmente quando associados ao risco do potencial erosivo de uma chuva.

Neste contexto, objetivou-se, analisar o ajuste da distribuição de probabilidade Log-Normal a 3P, utilizando para estimativa dos parâmetros, os métodos dos momentos (MM) e Máxima Verossimilhança (MV), em sete séries históricas de precipitação máxima diária anual, da bacia hidrográfica do Rio Verde (UPGRH GD4), a fim de identificar qual método melhor se aplica a esta distribuição para esta região.

2. MATERIAL E MÉTODOS

A bacia hidrográfica do rio Verde corresponde a 4,25% da área total da Bacia Hidrográfica do rio Grande, constituindo a Unidade de Planejamento e Gestão dos Recursos Hídricos UPGRH GD4, na região Sul de Minas Gerais. Possui área de drenagem de 6.891,4 km², onde estão inseridos 31 municípios, com população total de cerca de 450 mil habitantes.

O clima da região é do tipo mesotérmico brando e úmido, com 3 meses secos e temperatura média anual de 18 a 19°C, ou seja, do tipo Cwa (classificação de Koppen) em sua grande maioria, podendo ocorrer Cwb na região de cabeceira. O alto curso do rio Verde localiza-se próximo ao município de São Lourenço, onde se predominam encostas de grande declividade e solos rasos. A região de Varginha corresponde ao baixo rio Verde, com altitudes variando entre 900 m e 1.000 m (CBH GRANDE, 2012).

Foram utilizados dados de precipitação diária máxima anual para a constituição das séries históricas. Estes dados foram obtidos junto à Agência Nacional de Águas (ANA/HIDROWEB) de sete estações pluviométricas distribuídas pela bacia hidrográfica sendo elas: Aiuruoca, Caxambu, Cristina, Cruzeiro, São Lourenço, Usina de Varginha (Varginha) e Usina do Chicão (Campanha) com diferentes períodos, com pelo menos 30 anos de dados observados conforme a Tabela 1 e Figura 1.

Figura 1 – Distribuição das estações pluviométricas trabalhadas na região da Sub-bacia do Rio Verde.



Tabela 1 - Estações pluviométricas obtidas

Nome da estação	Código	Município	Número de dados
Aiuruoca	2144018	Aiuruoca	60
Caxambu	2144003	Caxambu	60
Cristina	2245065	Cristina	66
Cruzeiro	2244003	Cruzeiro	58
São Lourenço	2245107	São Lourenço	38
Usina de Varginha	2145018	Varginha	30
Usina do Chicão	2145009	Campanha	57

Para cada série histórica de dados de chuvas extremas foram ajustados os parâmetros da distribuição Log-normal a 3 parâmetros, sua Função de Densidade de Probabilidade (FDP) é apresentada na Equação 1.

$$FDP: f(x) = \frac{1}{(x - \beta) \cdot \sigma_n \cdot \sqrt{2} \cdot \pi} \cdot e^{-0,5 \left(\frac{\ln(x - \beta) - \mu_n}{\sigma_n} \right)^2} \quad (1)$$

Para estimativa dos parâmetros desta FDP, foram aplicados os métodos dos momentos (MM) e máxima verossimilhança (MV). Para detalhes sobre as citadas metodologias,

recomenda-se Naghettini & Pinto (2007), Casella & Berger (1990) e Haan (2002).

Com os parâmetros estimados ajustados pelos métodos MM e MV, a fim de se observar qual o melhor ajuste de probabilidade, representando adequadamente o conjunto de dados, aplicou-se os testes de aderência de Komolgorov Smirnov (ΔF) (eq2) e Qui-quadrado (λ^2) (eq. 3), a 5% de significância.

Na obtenção do ΔF , calculou-se o erro absoluto máximo entre a frequência observada e a probabilidade estimada. O valor de ΔF tabelado é encontrado em função da significância de 5%

(α) e do número de dados da série histórica (n). Para valores menores que os tabelados, o modelo foi classificado como adequado (Equação 2).

$$|\Delta F|_{\text{calculado máx}} \leq |\Delta F|_{\text{tabela}(n,\alpha)} \quad (2)$$

O Teste do Qui-quadrado (λ^2) a hipótese é testada fazendo-se a comparação entre as frequências observadas e as frequências teóricas, em cada classe de frequência da amostra, com a variável aleatória do λ^2 , dada por CAMPOS (1979) citado por CATALUNHA et al. (2002). Para o cálculo de λ^2 , foram geradas classes de frequência para a série de dados das quais foram extraídas frequências teóricas com cada distribuição por cada método de ajuste de parâmetros. O somatório das diferenças quadráticas (Equação 3) resulta no valor de λ^2 calculado, para valores menores que o λ^2 tabelado (em função do grau de significância e do grau de liberdade), o modelo foi classificado como adequado.

$$\lambda^2_{\text{calculado}} = \sum_{i=1}^n \frac{(f_{\text{obsi}} - f_{\text{teoricoi}})^2}{f_{\text{teoricoi}}} \quad (3)$$

Em que n é o número de classes, F_{obsi} e f_{teoricoi} são respectivamente as frequências observada e teórica na classe i .

O λ^2 tabelado foi obtido com base no número de classes de frequência de cada série menos o número de parâmetros menos um, fornecendo os graus de liberdade, e significância de 5%. Se $\lambda^2_{\text{calc}} < \lambda^2_{\text{tab}}$ conclui-se que a distribuição de probabilidade é adequada à série estudada.

3. RESULTADOS E DISCUSSÃO

A distribuição Log-normal a 3 parâmetros foi ajustada para cada uma das sete estações pluviométricas. A título de exemplo, a Figura 2 apresenta os ajustes pelos métodos MM e MV para a série histórica de Aiuruoca.

Observa-se na referida figura que a estimativa pelo MM superestima valores médios e máximos enquanto que por MV, ocorre superestimativa de valores mínimos e bom ajuste a valores médios e máximos. Resultados semelhantes foram observados nas demais seis séries históricas.

A fim de se analisar estatisticamente os ajustes, identificando o melhor método de estimativa de parâmetros para esta distribuição na bacia do Rio Verde, apresentam-se na Tabela 2 os resultados dos testes de aderência.

Observa-se que o teste de Qui-quadrado se destaca por ser mais rígido que o de Kolmogorov-Smirnov, sinalizando um maior rigor do mesmo na constatação do ajuste da distribuição de probabilidades aos dados observados, uma vez que classificou como inadequadas 43% dos modelos, enquanto que o ΔF , apenas 7%.

Tomando como base o teste de λ^2 , identifica-se maior adequação (88%) na estimativa pelo método MV em comparação com o MM (58%). Em uma análise da precisão do teste de λ^2 , observou-se que para a série histórica de Varginha ajustada pelo método MV, o teste apresentou o menor valor, em comparação com as demais, demonstrando melhor ajuste.

Figura 2- Ajustes da distribuição Log-normal a 3 parâmetros à série histórica de precipitação máxima diária anual de Aiuruoca.

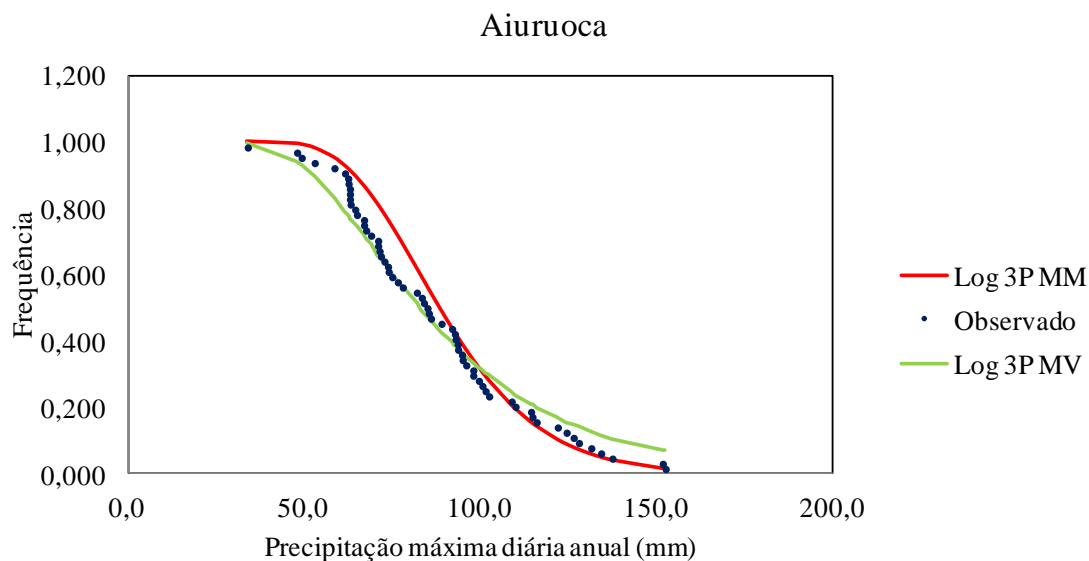


Tabela 2 - Adequabilidade da distribuição Log-Normal a 3 parâmetros e precisão de λ^2 para as séries históricas de precipitação máxima diária anual, na bacia do Rio Verde, Sul de Minas Gerais.

Série Histórica	Método	KS	λ^2	λ^2 calculado
Aiuruoca	MM	Adequado	Inadequado	–
	MV	Adequado	Adequado	8,36
Camapanha	MM	Adequado	Inadequado	–
	MV	Adequado	Adequado	7,378
Caxambu	MM	Adequado	Inadequado	–
	MV	Adequado	Adequado	4,219
Cristina	MM	Adequado	Adequado	4,595
	MV	Adequado	Adequado	6,002
Cruzeiro	MM	Adequado	Inadequado	–
	MV	Adequado	Inadequado	–
São Lourenço	MM	Adequado	Adequado	7,09
	MV	Inadequado	Adequado	5,225
Varginha	MM	Adequado	Inadequado	–
	MV	Adequado	Adequado	2,839

*KS= Komolgorov-Smirnov

Vários estudos de métodos de ajuste dos parâmetros corroboram como os obtidos no presente estudo indicando o método da Máxima Verossimilhança como o mais adequado (SILVA & CLARKE (2004); MELLO & SILVA (2005), HARTMANN et al. (2011); Caldeira et al.

(2016); MARQUES et al. (2014); FRANCO et. al. (2014); OLIVEIRA et al., (2014).

Ressalta-se que a estação de São Lourenço não foi observado ajuste nem para o método dos momentos e nem para o método da Máxima verossimilhança pelos dois testes de aderência.

4. CONCLUSÃO

Para sete séries históricas de precipitação máxima anual distribuídas pela bacia do Rio Verde (UPGRH GD4), ajustadas pela distribuição de probabilidades Log-normal a 3 parâmetros, observou-se melhor ajuste pelo método MV, em comparação ao MM. O teste de Qui-quadrado foi considerado o mais restritivo e a série histórica de Varginha obteve o melhor ajuste. Em estudos com precipitação máxima anual, nesta região, pela distribuição Log-normal a 3 parâmetros, a utilização do método MV pode ser a mais adequada.

5. REFERÊNCIAS

- ANA - Agência Nacional das Águas. Sistema de Informações Hidrológicas. Disponível em: <http://hidroweb.ana.gov.br/>. Acesso em 10/04/2012.
- BACK, A. J. Seleção de distribuição de probabilidade para chuvas diárias extremas do estado de Santa Catarina, **Revista Brasileira de Meteorologia**, 2001. v. 16, n.2, 211p.
- CASELLA, G., BERGER, R. L. **Statistical Inference**. California: Wadsworth and Brooks/Cole, Pacific Grove, 1990. 650 p.
- CBH Grande - Comitê da bacia hidrográfica do rio grande. Disponível em: <http://www.grande.cbh.gov.br/GD4>. Acesso em 16/04/2017.
- CALDEIRA, T. L, BESKOW, S. MELLO, C. R. FARIA, L. C. SOUZA, M, R. GUEDES, H. A. S. Modelagem probabilística de eventos de precipitação extrema no estado do Rio Grande do Sul. **Agriambi**, Campina Grande, v.19, n. 3, p. 197-203, 2016.
- CATALUNHA, M. J. et al. Aplicação de cinco funções densidade de probabilidade a series de precipitação pluvial no Estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, Santa Maria, v. 10, n. 01, p. 153-162, 2002.
- HAAN, C. T. **Statistical methods in hydrology**. 2.ed. Ames: The Iowa State University, 2002. 377p.
- HATMANN, M.; MOALA, F. A.; MENDONÇA, M. A. Estudo das precipitações máximas anuais em Presidente Prudente. **Revista Brasileira de Meteorologia**, São José dos Campos, v.26, n.4, 561 – 568, 2011.
- JUNQUEIRA JÚNIOR, J. A. GOMES, N. M. MELLO, A. R. SILVA, A. M. Precipitação provável para a região de Madre de Deus, Alto rio Grande: modelos de probabilidades e valores característicos Lavras: **Ciência e Agrotecnologia**, 2007. v. 31, n.3.
- MARQUES, R. F. P. V.; MELLO, C. R., SILVA, A. M. ; FRANCO, C. S. ; OLIVEIRA, A. S. . Performance of the probability distribution models applied to heavy rainfall daily events. **Ciência e Agrotecnologia**, v. 38, p. 335-342, 2014.
- MARQUES, R. F. P. V.; MELLO, C. R. ; FRANCO, C. S. ; OLIVEIRA, A. S. ; ALCANTRA, E. . Métodos de estimativas dos parâmetros da distribuição Log-Normal a 3 parâmetros para a precipitação máxima diária em municípios com médio a alto potencial erosivo. **Sustentare**, v. 1, p. 108-117, 2017.
- MELLO, C. R.; SILVA, A. M. Métodos Estimadores dos Parâmetros da Distribuição de Gumbel e sua Influência em Estudos Hidrológicos de Projeto. **Irriga**, Botucatu, v. 10, n, 4, p. 318-334, 2005.
- MIRANDA, C. T. S., THEBALDI, M. S., ROCHA, G. M. R, B. Precipitação máxima diária anual e estimativa da equação de chuvas intensas do município de Divinópolis, MG, Brasil,. **Revista Scientia Agraria**, Curitiba v. 18, n. 4, p. 09-16, 2017
- NAGHETTINI, M.; PINTO, E. J. A.: **Hidrologia estatística**. Belo Horizonte: CPRM, 2007. 552p.
- OLIVEIRA, A. S.; Mello, C. R.; FRANCO, C. S.; MARQUES, R. F. P. V.; SILVA, A. M. . Aplicabilidade da distribuição GEV ao estudo da precipitação máxima diária anual na região sul de Minas Gerais. **Revista Agrogeoambiental**, v. 6, p. 31-44, 2014.
- RODRIGUES, J. A.; FILHO J. S.; CHAVES, L. M.. Funções densidade de probabilidade para a estimativa de precipitação mensal. **Semina: Ciências Exatas e Tecnológicas**, Londrina, v. 34, n. 1, p. 03-08, 2013

SILVA, B. C.; CLARKE, R. T. Análise estatística de chuvas intensas na Bacia do Rio São Francisco. **Revista Brasileira de Meteorologia**, São José dos Campos v. 19, n. 3, p. 265-272, 2004.

Camila Silva Franco

Engenheira Ambiental – UNESP, Doutora e Recursos Hídricos em Sistemas Agrícolas – UFLA. Professora adjunta do Departamento de Engenharia – UFLA..

Rosângela Francisca de Paula Vitor Marques

Engenheira Florestal – UFLA, Mestre e Doutora e Recursos Hídricos em Sistemas Agrícolas – UFLA. Professora do curso de Mestrado em Sustentabilidade em Recursos Hídricos – UninCor

Luiz Fernando Coutinho de Oliveira

Engenheiro Agrícola – UFLA, Mestre em Engenharia Agrícola – UFLA, Doutor em Engenharia Agrícola – UFV. Professor titular do Departamento de Engenharia – UFLA.

Antônio Marciano da Silva

Engenheiro Agrônomo – UFLA, Mestre em Recursos Hídricos e Saneamento – UFRGS, Doutor em Engenharia Civil – USP. Professor emérito do Departamento de Engenharia – UFLA.
