



## Análise da temperatura máxima por meio de distribuição de probabilidade: estudo na cidade de Rio Verde

### *Maximum temperature analysis through probability distribution: a study in the city of Rio Verde*

Marcos Vinicius de Oliveira Peres<sup>1</sup> 

Marcos Vinicius Bueno<sup>2</sup> 

Micaías Mozzer de Oliveira<sup>3</sup> 

John Arnold Martins Santos<sup>4</sup> 

**Resumo:** Este trabalho teve como objetivo ajustar algumas distribuições de probabilidade disponível na literatura (Normal, Log-Normal, Weibull, Gama, Gumbel) para a variável resposta temperatura máxima mensal, para a cidade de Rio Verde, Goiás. De acordo com os testes de aderência, a distribuição Weibull não se ajustou adequadamente apenas aos meses de maio, junho, julho e setembro, já a distribuição Normal foi adequada para os meses de janeiro e outubro e a Gama foi adequada apenas para o mês de outubro. Com estes resultados, é possível obter a probabilidade de uma determinada temperatura máxima para os meses indicados, tendo usabilidade em diversas áreas.

**Palavras-chave:** Dados climatológicos; distribuição Weibull; modelos probabilísticos.

**Abstract:** The purpose of this work was to adjust some probability distributions widely known in the literature (Normal, Log-Normal, Weibull, Gama, Gumbel) for the monthly maximum temperature response variable of the city of Rio Verde, Goiás. According to the good of fit tests, the Weibull distribution was not suitable only for the months of May, June, July, and September, while the Normal distribution was suitable for the months of January and October and the Gamma was the only suitable for the month of October. Considering the results, it is possible to obtain the probability of certain maximum temperature for the indicated months, having usability in several areas.

**Keywords:** Climatological data; probabilistic models; Weibull distribution.

<sup>1</sup>Doutor em Ciências, Universidade Estadual do Paraná - Campus de Paranavaí, mvperes1991@gmail.com.

<sup>2</sup>Mestrando em Bioestatística, Universidade Estadual de Maringá, viniciusbueno201@outlook.com

<sup>3</sup>Mestrando em Bioestatística, Universidade Estadual de Maringá, oliveiramicaias@outlook.com.br

<sup>4</sup>Graduado em Matemática, Universidade Estadual de Maringá, john2008\_arnold@hotmail.com

# 1 INTRODUÇÃO

Os especialistas em meteorologia e climatologia afirmam que a transferência de calor e massa entre a atmosfera e o oceano é dependente da temperatura da superfície do mar, exercendo enorme influência nas condições climáticas. De acordo com Karl, Knight e Christy (1994), a baixa amostragem de observações das temperaturas do ar em algumas regiões indica a utilidade no uso de informações de áreas oceânicas para a obtenção da variável meteorológica.

As construções e as implantações de atividades humanas em meio as áreas naturais, acabam gerando consequências a região. Por exemplo, em áreas urbanas, a redução da cobertura vegetal pode ocasionar alterações climáticas, processos erosivos e deslizamentos, o que podem impactar a qualidade ambiental (COSTA; RODRIGUES, 2015).

Com o crescimento populacional na cidade de Rio Verde - Goiás, foi possível observar nos últimos tempos um processamento rápido da ampliação da área urbana no qual resultou em perda de florestas. Segundo Monteiro (1990), o fato da redução da vegetação causar um desconforto térmico, ocupando ao solo sem possíveis planejamentos ocasiona problemas térmicos, com picos de temperaturas e quedas de umidade relativa do ar. A oscilação nos índices de temperatura e na umidade do ar, é favorável à problemas de saúde, impactando negativamente a qualidade de vida humana, segundo a Organização Mundial da Saúde - OMS.

Diante do exposto, este trabalho tem como objetivo sistematizar os dados climatológicos obtidos pelas estações automáticas do Instituto Nacional de Meteorologia (INMET) na cidade de Rio Verde do estado de Goiás, realizando análises críticas dos ajustes de diferentes distribuições de probabilidade de temperatura máxima.

## 2 MATERIAIS E MÉTODOS

O conjunto de dados analisados contém informações referente a temperatura máxima para o Município de Rio Verde, Goiás, com medições diárias entre 01/01/1980 e 31/12/2020. Os dados foram coletados por meio do site do Instituto Nacional de Meteorologia (INMET), referentes a estação de medição número 83.470.

Para a modelagem da temperatura máxima utilizou-se das distribuições de probabilidade Normal, Log-Normal, Weibull, Gama e Gumbel, qual a mais adequada para descrever a temperatura máxima em Rio Verde, Goiás, de acordo com cada mês.

Na literatura, diversos trabalhos podem ser encontrados usando técnicas para descrever a temperatura máxima de um local por meio de distribuições de probabilidade. Em Araújo et al. (2010), as distribuições Beta, Gama, Gumbel, LogNormal, Normal e Weibull são utilizadas

para tentar descrever temperatura diária do ar máxima e mínima em Iguatu, Ceará, e utilizando os testes de aderência de Kolmogorov-Smirnov e Qui-quadrado, relatam que as distribuições Normal e Log-normal demonstraram um melhor ajuste aos dados.

Torsen et al. (2015), utilizam as distribuições Johnson SB, Beta, Kumaraswamy e Pareto generalizada para verificar a adequação para descrever dados de temperatura máxima no Estado Adamawa, Nigéria. De acordo com os autores, com base nos testes de aderência de Kolmogorov-Smirnov, Anderson-Darling, a distribuição Johnson SB obteve um melhor ajuste entre as descritas.

Trabalhando com a mesma temática, Hossian, Abdulla e Rahman (2016) utilizam sete distribuições probabilísticas para tentar descrever a temperatura máxima em Dacca, capital de Bangladesh. É relatado pelos autores que a distribuição Logística enviesada generalizada foi a mais parcimoniosa, entre as testadas, para os dados utilizados.

A distribuição Normal, ou Gaussiana, possui grande usabilidade para descrever diversos fenômenos naturais. Sua função de densidade de probabilidade é descrita da seguinte forma (MORETTIN; BUSSAB, 2017):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}, \quad (1)$$

em que  $\mu \in \mathbb{R}$  é a média da distribuição e  $\sigma > 0$  o desvio padrão.

A distribuição Log-Normal também possui grande usabilidade para a modelagem de dados, em que, considera-se que o logaritmo natural da variável resposta segue uma distribuição normal. Sua função de densidade de probabilidade, vista em Filho, Parisi e Soares (2018), é dada por:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp \left\{ -\frac{(\ln(x) - \mu)^2}{2\sigma^2} \right\}, \quad (2)$$

em que  $x > 0$ ,  $\mu > 0$  e  $\sigma > 0$  são, respectivamente, a média e o desvio padrão do logaritmo dos dados.

A distribuição Weibull é amplamente conhecida na literatura e possui grande usabilidade na modelagem de dados nas mais diversas áreas das ciências. Sua função de densidade de probabilidade pode ser observada em Johnson, Kotz e Balakrishnan (1995):

$$f(x) = \frac{\alpha}{\beta} \left( \frac{x}{\beta} \right)^{\alpha-1} \exp \left\{ -\left( \frac{x}{\beta} \right)^\alpha \right\}, \quad (3)$$

em que  $x > 0$ ,  $\alpha > 0$  e  $\beta > 0$  são os parâmetros de forma e escala, respectivamente.

O modelo probabilístico Gama é amplamente conhecido literatura e usado para descrever fenômenos naturais, como em Silva e Sentelhas (2001) e Murta et al. (2005), de modo que, sua função de densidade de probabilidade é dada por:

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{(\alpha-1)} \exp\left\{-\frac{x}{\beta}\right\}, \quad (4)$$

em que  $x > 0$ ,  $\alpha > 0$  e  $\beta > 0$  são os parâmetros de forma e escala, respectivamente.

A distribuição Gumbel também possui grande usabilidade em fenômenos naturais, sendo utilizada para descrever temperatura máxima em Filho, Matzenauer e Trindade (2005) e Ferreira e Liska (2019). Sua função de densidade de probabilidade é dada por:

$$f(x) = \frac{1}{\beta} \exp\{- (z + \exp\{z\})\}, \quad (5)$$

em que  $x > 0$ ,  $z = (x - \alpha)/\beta$ ,  $\alpha > 0$  e  $\beta > 0$  são os parâmetros de forma e escala, respectivamente.

Para as encontrar as estimativas dos parâmetros foi utilizado o método da máxima verossimilhança, por meio do pacote *fitdistrplus* (DELIGNETTE-MULLER; DUTANG, 2015), para o ambiente estatístico R (R Core Team, 2021), e para a implementação da distribuição Gumbel, foi utilizado o pacote *actuar* (DUTANG; GOULET; PIGEON, 2008). O método da máxima verossimilhança consiste em encontrar as estimativas dos parâmetros que maximizam a função de verossimilhança, definida por:  $L(\theta, x_1, \dots, x_n) = \prod_{i=1}^n f(x_i|\theta)$ , onde  $x_1, \dots, x_n$  é uma amostra com  $n$  observações independentes e identicamente distribuídas. Para uma melhor compreensão do método da máxima verossimilhança, ver Ramachandran e Tsokos (2009).

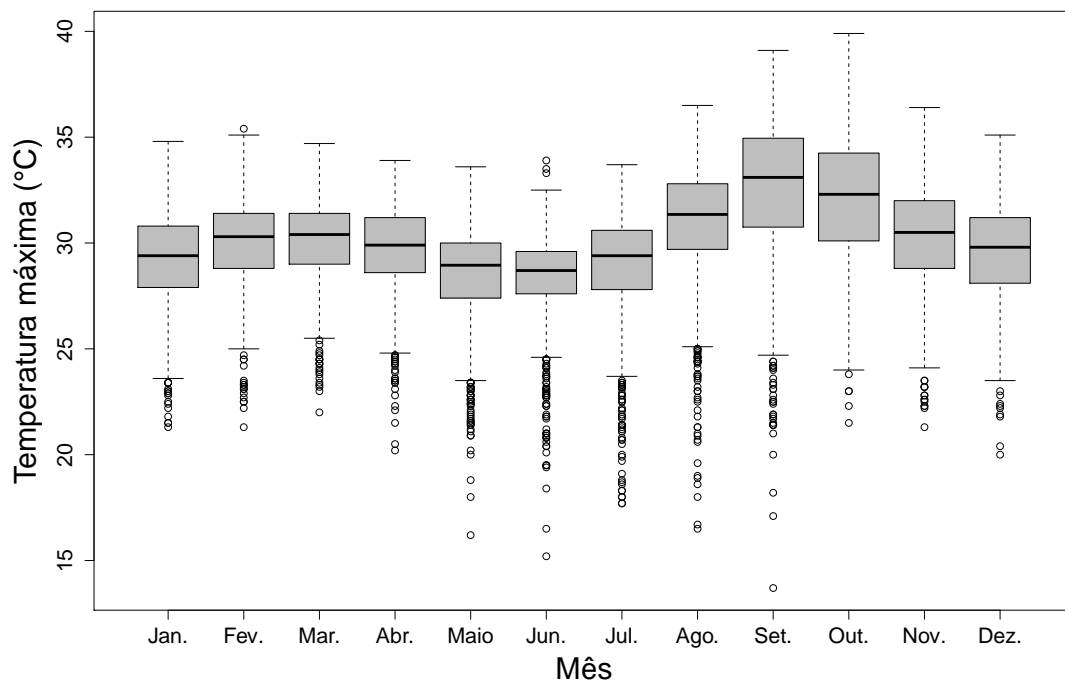
Para a verificação da aderência da distribuição aos dados, foram utilizados os testes de Kolmogorov-Smirnov, Anderson-Darling e de Cramér-von Mises, ao nível de 5% de significância e com auxílio do pacote *gofTest* (FARAWAY et al., 2019). Em meses que observou-se mais de um ajuste, o critério de informação de Akaike (AIC) foi utilizado para selecionar a distribuição mais parcimoniosa.

### 3 RESULTADOS E DISCUSSÕES

Inicialmente, foram removidas as observações perdidas para a variável em estudo, de modo que o conjunto final dos dados contou com 10.008 observações, que foram agrupadas de acordo com cada mês. Na Figura 1, pode-se observar os gráficos boxplot para temperatura máxima para cada mês. Na Tabela 1, observa-se um resumo estatístico para a variável em

estudo.

**Figura 1:** Boxplot da temperatura máxima diária para cada mês.



**Fonte:** Os autores (2022).

**Tabela 1:** Resumo descritivo para temperatura máxima diária agrupada mensalmente.

Mês	Média	Mediana	Desvio Padrão
Janeiro	29,19	29,40	2,31
Fevereiro	30,03	30,30	2,13
Março	30,07	30,40	1,99
Abril	29,67	29,90	2,11
Maio	28,42	28,95	2,39
Junho	28,28	28,70	2,39
Julho	28,82	29,40	2,66
Agosto	30,92	31,35	2,92
Setembro	32,42	33,10	3,58
Outubro	32,15	32,30	3,14
Novembro	30,21	30,50	2,49
Dezembro	29,56	29,80	2,33

**Fonte:** Os autores (2022).

Por meio do ambiente estatístico R e do pacote *fitdistrplus*, foram estimados os parâmetros das distribuições testadas para cada mês. Para verificar a qualidade do ajuste das distribuições

ajustadas para cada mês, foi observado o p-valor para os testes de Anderson–Darling (Tabela 2), Kolmogorov-Smirnov (Tabela 3) e Cramér–von Mises(Tabela 4).

**Tabela 2:** P-valores dos testes de aderência de Anderson–Darling.

Mês	Distribuições				
	Normal	Log-Normal	Weibull	Gama	Gumbel
Janeiro	0,04939	0,00167	0,09259(*)	0,00584	<0,00001
Fevereiro	0,00176	<0,00001	0,23797(*)	0,00014	<0,00001
Março	0,00010	<0,00001	0,09451(*)	<0,00001	<0,00001
Abril	0,00030	<0,00001	0,69979(*)	<0,00001	<0,00001
Maió	<0,00001	<0,00001	0,00346	<0,00001	<0,00001
Junho	<0,00001	<0,00001	0,00036	<0,00001	<0,00001
Julho	<0,00001	<0,00001	0,00806	<0,00001	<0,00001
Agosto	<0,00001	<0,00001	0,05006(*)	<0,00001	<0,00001
Setembro	<0,00001	<0,00001	0,02058	< 0,00001	<0,00001
Outubro	0,41895(*)	0,01975	0,04606	0,06348(*)	<0,00001
Novembro	0,00477	<0,00001	0,72635(*)	0,00037	<0,00001
Dezembro	0,02245	0,00059	0,21740(*)	0,00227	<0,00001

(\*): aderência a distribuição indicada, ao nível de 5% de significância.

**Fonte:** Os autores (2022).

De acordo com a Tabela 2, a distribuição Weibull foi adequada para expressar a temperatura máxima em Rio Verde apenas nos meses de maio, junho, julho e setembro, enquanto as distribuições Normal e Gama foram adequadas apenas para outubro. Considerando os resultados do teste de Kolmogorov-Smirnov (Tabela 3), a distribuição Normal foi adequada para o mês de outubro, enquanto a Weibull foi adequada para os meses de fevereiro, março, abril, outubro, novembro e dezembro. Por fim, os resultados do teste de Cramér–von Mises (Tabela 4 ) indicam que a distribuição Normal apresentou um ajuste adequado aos meses de janeiro e outubro, a Weibull apresentou um ajuste adequado para os meses de janeiro, fevereiro, março, abril, julho, outubro e novembro, enquanto a Gama foi adequada para o mês de outubro.

A Figura 2 apresenta uma comparação gráfica entre as curvas de cada distribuição abordada para os meses que tiveram mais de uma distribuição ajustada aos dados, já na Figura 3 pode-se observar essa mesma comparação para os meses que tiveram apenas uma distribuição ajustada.

As distribuições Log-Normal e Gumbel não apresentaram ajuste aos dados em nenhum mês, de acordo os resultados apresentados nas Tabelas 2, 3 e 4.

**Tabela 3:** P-valores dos testes de aderência de Kolmogorov-Smirnov.

Mês	Distribuições				
	Normal	Log-Normal	Weibull	Gama	Gumbel
Janeiro	0,03905	0,00171	0,01774	0,00544	< 0,00001
Fevereiro	0,00072	< 0,00001	0,15908(*)	< 0,00001	< 0,00001
Março	0,00028	< 0,00001	0,07489(*)	< 0,00001	< 0,00001
Abril	< 0,00001	< 0,00001	0,466652(*)	< 0,00001	< 0,00001
Mai	< 0,00001	< 0,00001	0,00048	< 0,00001	< 0,00001
Junho	< 0,00001	< 0,00001	< 0,00001	< 0,00001	< 0,00001
Julho	< 0,00001	< 0,00001	0,02895	< 0,00001	< 0,00001
Agosto	< 0,00001	< 0,00001	0,01522	< 0,00001	< 0,00001
Setembro	< 0,00001	< 0,00001	0,03694	< 0,00001	< 0,00001
Outubro	0,25691(*)	0,01443	0,10455(*)	0,04401	< 0,00001
Novembro	0,00265	< 0,000013	0,68614(*)	0,00024	< 0,00001
Dezembro	0,00737	0,00015	0,22346(*)	0,00061	< 0,00001

(\*): aderência a distribuição indicada, ao nível de 5% de significância.

**Fonte:** Os autores (2022).

**Tabela 4:** P-valores dos testes de aderência de Cramér-von Mises.

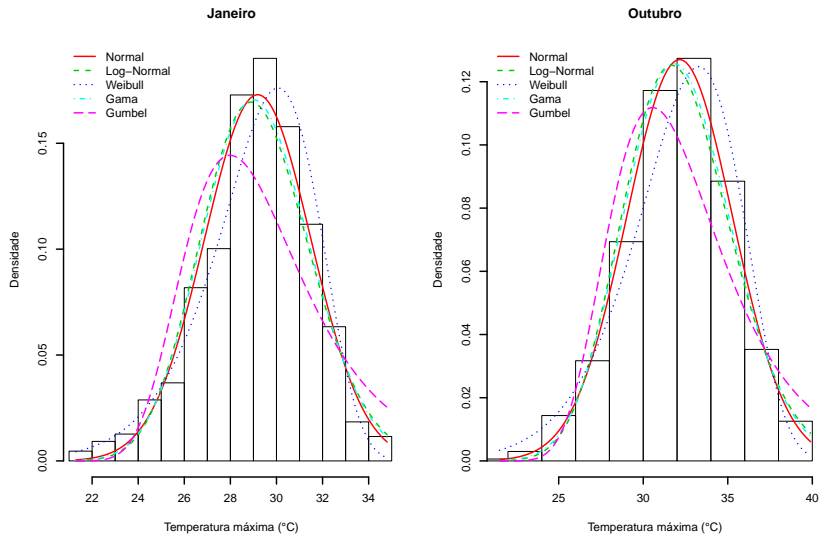
Mês	Distribuições				
	Normal	Log-Normal	Weibull	Gama	Gumbel
Janeiro	0,07250(*)	0,00464	0,22346(*)	0,01275	< 0,00001
Fevereiro	0,00504	0,00019	0,28676(*)	0,00063	< 0,00001
Março	0,00308	0,00013	0,14929(*)	0,00040	< 0,00001
Abril	0,00174	< 0,00001	0,62163(*)	0,00015	< 0,00001
Mai	< 0,00001	< 0,00001	0,00623	< 0,00001	< 0,00001
Junho	< 0,00001	< 0,00001	0,00129	< 0,00001	< 0,00001
Julho	< 0,00001	< 0,00001	0,03494	< 0,00001	< 0,00001
Agosto	< 0,00001	< 0,00001	0,09112(*)	< 0,00001	< 0,00001
Setembro	< 0,00001	< 0,00001	0,04394	< 0,00001	< 0,00001
Outubro	0,39717(*)	0,03120	0,09848(*)	0,08125(*)	< 0,00001
Novembro	0,00922	0,00031	0,7720(*)	0,00107	< 0,00001
Dezembro	0,03522	0,00181	0,30049	0,00541	< 0,00001

(\*): aderência a distribuição indicada, ao nível de 5% de significância.

**Fonte:** Os autores (2022).

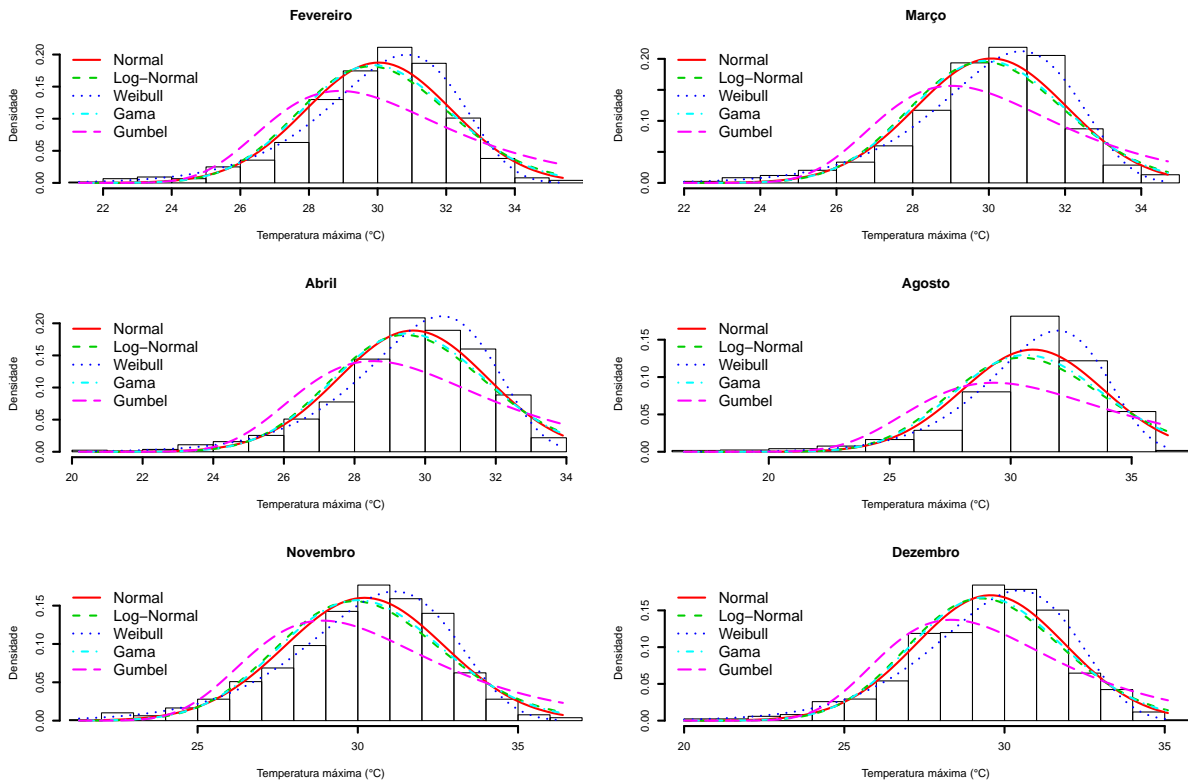
Considerando os meses em que houveram ajustes considerados adequados, a Tabela 5 apresenta as estimativas dos parâmetros das distribuições ajustadas, bem como o intervalo de confiança, e o AIC de cada ajuste. Com base nos valores do AIC, observa-se que a distribuição Weibull apresentou um melhor ajuste para o mês de janeiro, comparada a Normal, e a distribuição Normal apresentou um melhor ajuste para o mês de outubro, comparada a Weibull e Gama.

**Figura 2:** Histograma e curvas de densidade de probabilidade acumulada, de todas as distribuições consideradas, para os meses com mais uma distribuição ajustada.



**Fonte:** Os autores (2022).

**Figura 3:** Histograma e curvas de densidade de probabilidade acumulada, de todas as distribuições consideradas, para os meses com apenas uma distribuição ajustada.



**Fonte:** Os autores (2022).

**Tabela 5:** Estimativas dos parâmetros dos meses com distribuições de probabilidade ajustadas

Mês	Normal			Weibull			Gama		
	$\mu$ / IC	$\sigma$ / IC	AIC	$\alpha$ / IC	$\beta$ / IC	AIC	$\alpha$ / IC	$\beta$ / IC	AIC
Jan.	29,186 (29,033; 29,340)	2,306 (2,198; 2,415)	3.918,081	14,436 (13,713; 15,160)	30,219 (30,072; 30,366)	3.908,979	-	-	-
Fev.	-	-	-	16,791 (15,889; 17,692)	30,966 (30,828; 31,104)	3.261,594	-	-	-
Mar.	-	-	-	17,842 (16,929; 18,756)	30,952 (30,828; 31,076)	3.261,594	-	-	-
Abr.	-	-	-	17,517 (16,582; 18,453)	30,579 (30,453; 30,704)	3.476,927	-	-	-
Ago.	-	-	-	14,122 (13,373; 14,874)	32,115 (31,955; 32,275)	4.073,217	-	-	-
Out.	32,15,3 (31,940; 32,366)	3,141 (2,990; 3,291)	4.290,029	11,333 (10,760; 11,907)	33,557 (33,344; 33,769)	4.314,250	30,549 (30,312; 30,786)	3,289 (3,133; 3,445)	4.481,680
Nov.	-	-	-	14,308 (13,545; 15,070)	31,309 (31,147; 31,470)	3.627,348	-	-	-
Dez.	-	-	-	14,706 (13,955; 15,456)	30;603 (31,147; 31,470)	3.83,938	-	-	-

**Fonte:** Os autores (2022).

## 4 CONSIDERAÇÕES FINAIS

Com base nos testes de aderência, a distribuição Weibull foi adequada para os meses de janeiro, fevereiro, março, abril, agosto, outubro, novembro e dezembro. A distribuição Normal foi adequada para os meses de janeiro e outubro e a distribuição Gama apenas para o mês de outubro. As distribuições Log-Normal e Gumpel não se mostraram adequadas para descrever a temperatura máxima de Rio Verde. A partir da análise dos dados climáticos, observa-se que as distribuições teóricas de probabilidade que melhor se ajustaram em descrever a variável temperatura máxima, foram Weibull para o mês de janeiro, em comparação com a Normal, e a distribuição Normal para o mês de outubro, em comparação com a Weibull e Gama.

### Referências

- ARAÚJO, E. M. et al. Análise da aderência de distribuições de probabilidade aos dados de temperatura máxima e mínima do ar em Iguatu-CE. *Revista Caatinga*, v. 23, n. 3, p. 104–109, 2010. 2
- COSTA, Y. T.; RODRIGUES, S. C. Relação entre cobertura vegetal e erosão em parcelas representativas de Cerrado. *Revista Geográfica Acadêmica*, v. 9, n. 2, p. 61–75, 2015. 2
- DELIGNETTE-MULLER, M. L.; DUTANG, C. fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software*, v. 64, n. 4, p. 1–34, 2015. 4
- DUTANG, C.; GOULET, V.; PIGEON, M. actuar: An R Package for Actuarial Science. *Journal of Statistical Software*, v. 25, n. 7, p. 38, 2008. Disponível em: <<http://www.jstatsoft.org/v25/i07>>. 4
- FARAWAY, J. et al. *goftest: Classical Goodness-of-Fit Tests for Univariate Distributions*. [S.l.], 2019. R package version 1.2-2. Disponível em: <<https://CRAN.R-project.org/package=goftest>>. 4
- FERREIRA, R. V. de C.; LISKA, G. R. Análise probabilística da temperatura máxima em Uruguaiana, RS. *Revista Brasileira de Agricultura Irrigada*, Instituto de Pesquisa e Inovação na Agricultura Irrigada-INOVAGRI, v. 13, n. 3, p. 3390–3401, 2019. 4
- FILHO, A. C.; MATZENAUER, R.; TRINDADE, J. K. da. Ajustes de funções de distribuição de probabilidade à temperatura máxima no Estado do Rio Grande do Sul. *Pesquisa Agropecuária Brasileira*, v. 40, 2005. 4
- FILHO, C. de D. R.; PARISI, S. A.; SOARES, W. de A. Avaliação da distribuição estatística e elaboração de modelo de regressão múltipla linear da velocidade média do vento. *Journal of Environmental Analysis and Progress*, p. 24–36, 2018. 3
- HOSSIAN, M. M.; ABDULLA, F.; RAHMAN, M. H. Selecting the probability distribution of monthly maximum temperature of Dhaka (capital city) in Bangladesh. *Jahangirnagar University Journal of Statistical Studies*, v. 33, p. 33–45, 2016. 3

JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. *Continuous Univariate Distributions, volume 2*. New York: John Wiley & Sons, 1995. 3

KARL, T. R.; KNIGHT, R. W.; CHRISTY, J. R. Global and hemispheric temperature trends: Uncertainties related to inadequate spatial sampling. *Journal of Climate*, v. 7, n. 7, p. 1144–1163, 1994. 2

MONTEIRO, C. A. de F. Por um suporte teórico e prático para estimular estudos geográficos de clima urbano no Brasil. *Geosul*, v. 5, n. 9, p. 7–19, 1990. 2

MORETTIN, P. A.; BUSSAB, W. O. *Estatística básica*. São Paulo: Saraiva, 2017. 3

MURTA, R. M. et al. Precipitação pluvial mensal em níveis de probabilidade pela distribuição gama para duas localidades do sudoeste da Bahia. *Ciência e Agrotecnologia*, SciELO Brasil, v. 29, p. 988–994, 2005. 4

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>. 4

RAMACHANDRAN, K. M.; TSOKOS, C. P. *Mathematical statistics with applications in R*. [S.l.]: Academic Press, 2009. 4

SILVA, J. d.; SENTELHAS, P. C. Diferença de temperatura mínima do ar medida no abrigo e na relva e probabilidade de sua ocorrência em eventos de geada no Estado de Santa Catarina. *Revista Brasileira de Agrometeorologia*, v. 9, n. 1, p. 9–15, 2001. 4

TORSEN, E. et al. Probability distribution of maximum temperature in Adamawa State, Nigeria. *Journal of Mathematics (IOSR-JM)*, v. 11, n. 4, p. 1–6, 2015. 3

**Enviado em** : 15 dez. 2021.

**Aceito em** : 02 jul. 2022.

**Editores responsáveis** : Bianca Neves Machado.