

# PROPOSTAS BASEADAS EM ALGORITMOS GENÉTICOS PARA OTIMIZAÇÃO DE PARÂMETROS DE CURVAS INTENSIDADE-DURAÇÃO FREQUÊNCIA (IDF)

ZANDRA ALMEIDA DA CUNHA<sup>1</sup>; MARCELLE MARTINS VARGAS<sup>2</sup>; MAÍRA MARTIM DE MOURA<sup>3</sup>; TAMARA LEITZKE CALDEIRA<sup>4</sup>; SAMUEL BESKOW<sup>5</sup>

<sup>1</sup>Universidade Federal de Pelotas (UFPel) – zandraacunha@gmail.com

<sup>2</sup>Universidade Federal de Pelotas (UFPel) – marcellevarg@gmail.com

<sup>3</sup>Universidade Federal de Pelotas (UFPel) – martimdemoura@gmail.com

<sup>4</sup>Universidade Federal de Pelotas (UFPel) – tamaraleitzkecaldeira@ufpel.edu.br

<sup>5</sup>Universidade Federal de Pelotas (UFPel) – samuelbeskow@gmail.com

## 1. INTRODUÇÃO

A curva Intensidade-Duração-Frequência (IDF) é uma das ferramentas mais utilizadas e de grande relevância para a gestão dos recursos hídricos, visto que o estudo de precipitações extremas permite a estimativa das vazões de projeto para dimensionamento de obras de engenharia. As informações necessárias para a determinação das relações IDF são dados provenientes de monitoramento hidrológico (ALEMAW; CHAOKA, 2016). No entanto, as informações provenientes do monitoramento existente são na grande maioria das vezes disponibilizadas em escala temporal diária, tornando-se necessário lançar mão de métodos para discretizar essas precipitações em intervalos de tempo subdiários.

Um dos métodos tradicionalmente utilizado para a obtenção da precipitação em intervalos de tempo subdiários é o método da desagregação proposto por CETESB (1979). Este método traz relações para conversão das precipitações máximas diárias observadas em intervalos de diferentes durações, sendo necessário o ajuste de funções de distribuições de probabilidades (FDP) para representar a frequência dos valores observados de forma satisfatória.

Após uma FDP ser ajustada à série histórica de precipitação máxima diária anual (P<sub>MDA</sub>), a acurácia dos parâmetros da IDF é dependente do método de otimização utilizado. O método comumente aplicado é o de regressão múltipla não linear associada ao método de mínimos quadrados, o qual apresenta bons resultados, no entanto, apresenta dificuldades associadas à inicialização dos parâmetros da IDF e à experiência necessária por parte do usuário. Neste sentido, a utilização de algoritmos de otimização que busquem ótimos globais minimiza as desvantagens supracitadas acerca do método tradicional supracitado. Dentre os algoritmos disponíveis na literatura, o *Non-dominated Sorting Genetic Algorithm II* (NSGA-II) é o comumente utilizado para otimização multiobjetivo (AUGER et al., 2016), enquanto que, o Nelder-Mead (NM) é frequentemente utilizado para resolver problemas de otimização irrestrita (NELDER; MEAD, 1965), e o Shuffled Complex Evolution (SCE-UA) é mais utilizado para abordagens determinísticas e probabilísticas (DUAN et al., 1994).

Frente ao exposto, este estudo propõe uma abordagem inicial para a elaboração de uma ferramenta computacional que objetiva otimizar de forma automática os parâmetros da curva IDF, descartando a dependência de conhecimento acerca dos limites dos parâmetros por parte do usuário.

## 2. METODOLOGIA

Foram utilizadas séries de precipitação de 112 postos pluviométricos de responsabilidade da Agência Nacional de Águas (ANA), selecionados de maneira a representar a heterogeneidade da precipitação das cinco regiões brasileiras e que atendessem aos critérios de falhas pré-estabelecidos, e.g. 31 dias de falhas e um mínimo de 10 anos de dados. As séries de  $P_{MDA}$  foram constituídas com o auxílio do software *System of Hydrological Data Acquisition and Analysis* (SYHDA) (VARGAS et al., 2019).

O procedimento de ajuste dos parâmetros para as curvas IDF consistiu em três etapas (KOUTSOYIANNIS et al., 1998): i) ajuste de uma FDP; ii) cálculo das intensidades de precipitação para um conjunto de tempo de retorno (TR), utilizando as FDPs da etapa i; e iii) obtenção das curvas IDF.

AS FDPs Log Normal de 2 parâmetros (LN-2P), Gumbel para Máximos e Generalizada de Valores Extremos (GEV) foram empregadas neste estudo e avaliadas de acordo com o teste de aderência de Anderson-Darling (AD) (ANDERSON; DARLING, 1954) ao nível de 5% de significância. Para estimar os parâmetros das FDPs foi utilizado o Método da Máxima Verossimilhança (MMV). O processamento das metodologias foi realizado em ambiente MATLAB (MATRIX LABORATORY, 2018), utilizando funções nativas.

Os TRs utilizados foram de 2, 5, 10, 20, 50 e 100 anos e durações (D) de 5, 10, 15, 20, 30, 60, 360 e 1440 minutos. A desagregação da precipitação diária tomou como base as constantes de desagregação propostas por CETESB (1979). Os parâmetros da IDF foram otimizados tendo como base a Equação 1.

$$i = a \cdot TR^b \cdot (c + D)^d \quad (01)$$

Sendo:  $i$  a intensidade média ( $\text{mm.h}^{-1}$ ), TR o tempo de retorno (anos), D a duração (min) e  $a$ ,  $b$ ,  $c$  e  $d$  os parâmetros da IDF.

Para a otimização dos parâmetros da IDF foi utilizado como função objetivo o coeficiente de Nash-Sutcliffe ( $C_{NS}$ ) (NASH; SUTCLIFFE, 1970), seguindo classificação de Motovilov et al. (1999). O desempenho foi computado a partir da estatística raiz quadrada do erro quadrático médio (RMSE). A otimização da função objetivo  $C_{NS}$  foi realizada de acordo com ilustrado na Figura 1, por meio da utilização no ambiente do MATLAB de: i) regressão não-linear de Levenberg Marquardt em sua função nativa *lsqnonlin*; ii) algoritmo NSGA-II, em sua versão na função nativa *gamultiobj*; iii) algoritmo NM (NELDER; MEAD, 1965); e iv) algoritmo SCE-UA (DUAN et al., 1994).

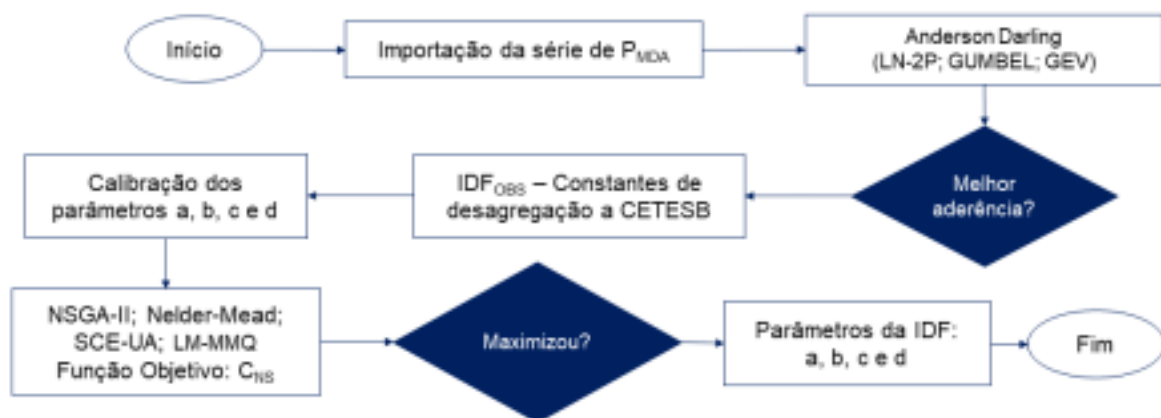


Figura 1 – Fluxograma compreendendo todas as etapas realizadas para a calibração dos parâmetros da IDF.

Foi necessário estipular os limites inferior e superior dos parâmetros da IDF para a otimização com o NSGA-II e o SCE-UA, e os valores iniciais para o LM-MMQ e NM. Os valores limites foram obtidos na literatura, tomando como base trabalhos que obtiveram parâmetros da IDF para diferentes locais do Brasil. Dada a flutuação dos valores dos parâmetros ao longo do território brasileiro, a otimização dos mesmos foi realizada sob dois cenários: Cenário I, onde foram considerados os mesmos limites para todas as regiões do Brasil; e o Cenário II, onde os limites foram os encontrados para cada região.

### 3. RESULTADOS E DISCUSSÃO

Ao analisar o ajuste das séries hidrológicas às distribuições de probabilidade, de acordo com o teste AD, verificou-se que 85 séries de  $P_{MDA}$  (79%) se ajustaram à FDP GEV, 26 séries (23,2%) se ajustaram à LN-2P e apenas 1 série (0,9%) à Gumbel. Na Tabela 1, é possível observar as estatísticas de ajuste dos parâmetros médios obtidos para a IDF, por região, nos cenários I e II para cada método de otimização.

Tabela 1 – Estatísticas médias para os Cenários I e II de otimização dos parâmetros da IDF																				
Região/ Método	Cenário I				Cenário II				Região/ Método	Cenário I				Cenário II						
	Método																			
	CNS		RMSE		CNS		RMSE			CNS		RMSE		CNS		RMSE				
Sudeste	NDGA-II	0,98	13,25	0,94	34,76															
	NDGA-II	0,98	7,85	0,93	11,53															
	SCE-UA	1,00	7,65	0,95	33,11	SCE-UA	1,00	2,83	0,96	7,40	NM	0,94	21,03	0,90	27,18	NM	0,95	13,68	0,95	12,44
	LM-MMQ	0,00	107,93	0,00	107,89	LM-MMQ	0,00	58,27	0,00	58,22										
Nordeste	NDGA-II	0,98	7,77	-0,04	34,80	SuI	NDGA-II	0,97	8,18	1,00	3,63									
	SCE-UA	1,00	3,15	0,03	33,68	SCE-UA	1,00	2,61	1,00	2,62	NM	0,94	15,46	0,97	9,28	NM	0,94	13,31	0,95	11,18
	LM-MMQ	0,00	67,42	0,00	67,33	LM-MMQ	0,00	57,02	0,00	56,99										
Centro-Oeste	NDGA-II	0,97	8,85	0,98	4,89															

SCE-UA 1,00 2,47 0,98 4,54  
NM 0,95 11,89 0,99 5,60  
LM-MMQ 0,00 53,05 0,00 53,02

Com base nos resultados obtidos para o Cenário I, pode-se observar que o algoritmo SCE-UA resultou nas melhores estatísticas, visto que todos os valores de  $C_{NS}$  foram enquadrados como “bom” e o valor de RMSE médio entre as regiões foi igual a  $3,7\text{mm}\cdot\text{h}^{-1}$ . Para o NSGA-II e o NM, os valores das estatísticas também mostraram resultados satisfatórios, sendo os valores de  $C_{NS}$  enquadrados como “bom” e os valores médios de RMSE entre as regiões de  $7,3\text{mm}\cdot\text{h}^{-1}$  e  $9,2\text{mm}\cdot\text{h}^{-1}$ , respectivamente. Contudo, o LM-MMQ resultou nas piores estatísticas, tendo todos os valores de  $C_{NS}$  enquadrados como “insatisfatórios” e um valor alto de RMSE médio entre as regiões, igual a  $68,7\text{mm}\cdot\text{h}^{-1}$ .

Para o Cenário II, o comportamento das estatísticas variou de acordo com os algoritmos, no entanto, o LM-MMQ continuou apresentando desempenho insatisfatório para todas as regiões. De acordo com Campos et al. (2014), os resultados obtidos pelo LM-MMQ se devem ao fato de que, geralmente, a calibração com regressão não-linear necessita conhecimento do usuário acerca dos valores iniciais adequados para cada parâmetro, ao passo que, para outros algoritmos, como o SCE-UA, basta inserir um limite amplo para os parâmetros que o método encontrará uma solução representativa. Foram verificados os maiores valores médios de RMSE para as regiões Norte e Nordeste com a utilização dos algoritmos NSGA-II e SCE-UA. Para as demais regiões, os valores de RMSE foram menores, destacando melhor desempenho do algoritmo SCE-UA.

#### 4. CONCLUSÕES

Algoritmos de otimização que não são dependentes dos valores iniciais, e.g. NSGA-II e SCE-UA, são uma alternativa metodológica para a estimativa dos parâmetros da IDF. O algoritmo SCE-UA permitiu melhor otimização dos parâmetros para ambos os cenários analisados, enquanto que, o LM-MMQ não foi satisfatório. O algoritmo LM-MMQ pode ser utilizado com sucesso na otimização dos parâmetros da IDF, entretanto, os valores iniciais devem ser bem estabelecidos.

Visando à elaboração de uma ferramenta automática para otimização dos parâmetros da IDF, sem a necessidade do conhecimento dos valores iniciais por parte dos usuários, a utilização dos algoritmos LM-MMQ e NM não se mostram adequadas. Todavia, a utilização de algoritmos como o NSGA-II e o SCE-UA pode ser uma boa alternativa metodológica.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

ALEMAW, B. F.; CHAOKA, R. T. Regionalization of Rainfall Intensity-Duration Frequency (IDF) curves in Botswana. **Journal of Water Resource and Protection**, v. 8, n.12, pp. 1128-1144, 2016.

ANDERSON, T. W.; DARLING, D. A. A test of goodness of fit. **Journal of the American Statistical Association**, v. 49, n. 268, pp. 756-76, 1954.

AUGER, A.; BROCKHOFF, D.; HANSEN, N.; TUŠAR, D.; TUŠAR, T.; WAGNER, T. Benchmarking MATLAB's gamultiobj (NSGA-II) on the Bi-objective BBOB-2016

Test Suite. In GECCO 2016-Genetic and Evolutionary Computation Conference (pp. 1233-1239). ACM, 2016.

CAMPOS, A. R.; SANTOS, G. G.; SILVA, J. B. L.; IRENE FILHO, J.; LOURA, D. S. Equações de intensidade-duração-frequência de precipitações para o estado do Piauí. **Revista Ciência Agronômica**, v. 45, n.3, pp. 488-498, 2014.

CETESB – Companhia de Tecnologia de Saneamento Ambiental. Drenagem urbana – Manual de Projeto. 3 ed. São Paulo. 464 p, 1979.

DUAN, Q.; SOROOSHIAN, S.; GUPTA, V. K. Optimal use of the SCE-UA global optimization method for calibrating watershed models. **Journal of hydrology**, v. 158, n. 3-4, pp. 265-284, 1994.

MATLAB. The MathWorks, Inc., Natick, Massachusetts, United States, R2018b, 2018.

MOTOVILOV, Y. G.; GOTTSCHALK, L.; ENGELAND, K.; RODHE, A. Validation of a distributed hydrological model against spatial observations. **Agricultural and Forest Meteorology**, v. 98, pp. 257-277, 1999.

NASH, J. E.; SUTCLIFFE, J. V. River flow forecasting through conceptual models I: a discussion of principles. **Journal of Hydrology**, v. 10, n.3, pp. 282- 298, 1970.

NELDER, J.A.; MEAD, R. A Simplex method for function minimization. **The Computer Journal**, v. 7, pp. 308-313, 1965.

VARGAS, M. M.; BESKOW, S.; CALDEIRA, T. L.; CORRÊA, L. D. L.; CUNHA, Z. A. D. SYHDA–System of Hydrological Data Acquisition and Analysis. **Revista Brasileira de Recursos Hídricos** 24, 2019.