

## **Desenvolvimento de algoritmo para preenchimento de falhas de dados pluviométricos**

**Felipe Oliveira PAULINO<sup>1</sup>; Thomé S. ALMEIDA<sup>1</sup>**

### **RESUMO**

A presença de falhas em dados de estações meteorológicas, como ausência de valores de precipitação, além de comprometerem a continuidade de dados, implicam a exclusão do período em questão das análises climáticas e agrometeorológicas. Dessa forma, o presente trabalho teve como objetivo de elaboração um algoritmo para a identificação e preenchimento da falha de uma série histórica de precipitação através do método estatístico de ponderação regional, utilizando a linguagem de programação JAVA. O algoritmo apresentou-se de fácil uso, indicando a quantidade de dados faltantes, dados preenchidos e não preenchidos como informação para o usuário.

### **INTRODUÇÃO**

Uma série meteorológica de uma determinada região é essencial para o estudo climático da mesma e possibilita seu uso em modelos agrometeorológicos para monitoramento das culturas agrícolas. Além disso, para um estudo hidrológico de uma determinada região é necessária a análise dos dados de precipitação, em que as primeiras análises são o preenchimento de falhas da série e análise de homogeneidade (consistência dos dados) em relação às estações vizinhas próximas (BERTONI & TUCCI, 2007).

---

<sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – campus Passos. Passos/MG - E-mail: [felipeoliveirapaulino2@gmail.com](mailto:felipeoliveirapaulino2@gmail.com)

<sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – campus Passos. Passos/MG - E-mail: [thome.almeida@ifsuldeminas.edu.br](mailto:thome.almeida@ifsuldeminas.edu.br)

Para a realização da primeira etapa, que refere ao preenchimento de falhas, alguns métodos foram desenvolvidos. Dentre eles pode-se destacar a utilização do método da regressão linear e ponderação regional (BERTONI & TUCCI, 2007), regressão linear múltipla (SILVA, 2011) e também a aplicação de redes neurais artificiais (MAIER & DANDY, 2000; LUCIO *et al.*, 2006).

Atualmente é muito comum que esses métodos sejam aplicados por meio do uso de planilhas eletrônicas. Entretanto, esse processo se torna muito dispendioso em termos de tempo e esforço, além de estar sujeito a falhas humanas devido à dependência da constante intervenção do usuário. Posto nesses termos, a utilização de um sistema computacional surge como uma alternativa para facilitar e agilizar a análise e recuperação dos dados de uma estação meteorológica.

Neste trabalho foi desenvolvido um algoritmo para identificação e preenchimento de falhas de dados pluviométricos utilizando o método da ponderação regional aplicado ao preenchimento de dados diários.

## MATERIAL E MÉTODOS

Para aplicação do algoritmo foram utilizados os dados de precipitação diária das estações meteorológicas do INMET (Instituto Nacional de Meteorologia) das cidades de Passos, Bambuí, Machado e São Sebastião do Paraíso, todas estações próximas e no estado de Minas Gerais, aplicando o preenchimento para as falhas nos dados de Passos.

Para o preenchimento de falhas foi implementada a metodologia da ponderação regional (Eq.1).

$$Y = \frac{1}{3} \left[ \frac{x_1}{x_{m1}} + \frac{x_2}{x_{m2}} + \frac{x_3}{x_{m3}} \right] \times Y_m \quad (1)$$

sendo, Y a precipitação diária estimada referente ao dia, mês ou ano faltante (mm),  $x_1$ ,  $x_2$  e  $x_3$  as precipitações correspondentes do dia, mês ou ano das estações vizinhas,  $x_{m1}$ ,  $x_{m2}$  e  $x_{m3}$  as precipitações médias das três estações vizinhas e  $Y_m$  a precipitação média da estação que deseja-se preencher.

Dentre as diversas linguagens de programação disponíveis para o desenvolvimento do sistema, optou-se pela utilização do Java, que implementa o paradigma de orientação a objetos, além de operar em máquina virtual. Com isso

garante-se que a aplicação possa ser executada em diversos sistemas operacionais, de forma a abranger uma maior gama de usuários.

Além disso, a linguagem dispõe de ferramentas para o desenvolvimento de uma interface gráfica do usuário (GUI) completa. A biblioteca Swing conta com uma grande variedade de componentes que podem amplamente explorados. A manuseabilidade desses componentes se torna especialmente simples com a utilização do GUIBuilder, ferramenta inerente a IDE Netbeans que permite uma interação dinâmica com elementos da interface gráfica.

Dado a possibilidade de se ter que trabalhar com um grande volume de dados o algoritmo foi desenvolvido de maneira a conseguir a máxima eficiência e praticidade possível. Pensando nisso, optou-se pela utilização de estruturas estáticas para armazenar os dados e facilitar a consulta às informações lidas e processadas.

Como a leitura dos dados é feita somente uma única vez, ao passo que as consultas são frequentes, foram utilizados TreeMaps para o armazenamento das informações. Como consistem em estruturas *key-value*, as TreeMaps garantem maior rapidez na recuperação das precipitações diárias, além de realizar a inserção dos dados de forma ordenada, ao contrário de outras estruturas indexadas como HashMaps.

## **RESULTADOS E DISCUSSÃO**

A versão mais recente do software conta com três janelas principais. A primeira janela (Figura 1) solicita o carregamento (feito por meio do componente `JFileChooser`) do arquivo texto contendo os dados da estação a ser corrigida. Ao clicar em avançar, os dados são armazenados em uma `TreeMap`, cuja chave consiste na data do valor precipitado (em formato `aaaaddmm`).

Em seguida é aberta outra janela (Figura 2), onde são carregados, independentemente, os dados das três estações auxiliares. Ao clicar em avançar, os dados são armazenados em `TreeMaps`, tal como descrito na etapa anterior. Nesse ponto é feita a verificação dos dados faltantes da estação a ser corrigida e, caso valores precipitados no mesmo dia tenham sido devidamente aferidos nas outras estações, é aplicado o método de ponderação regional, caso contrário o dado permanecerá com o valor "null".

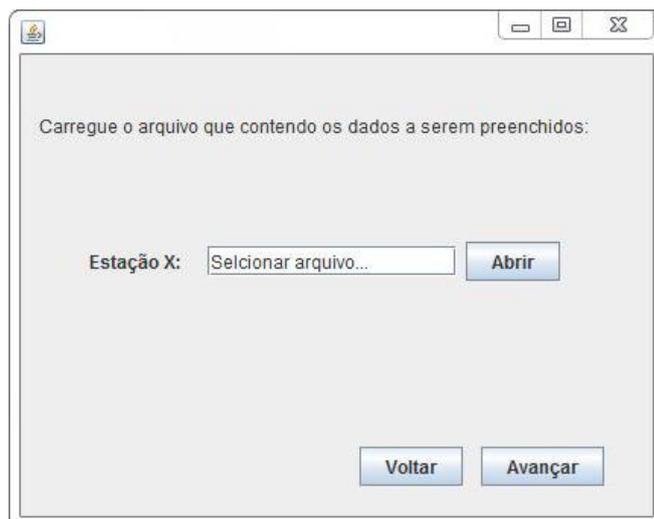


Figura 1: Tela para carregamento da estação a ser corrigida.

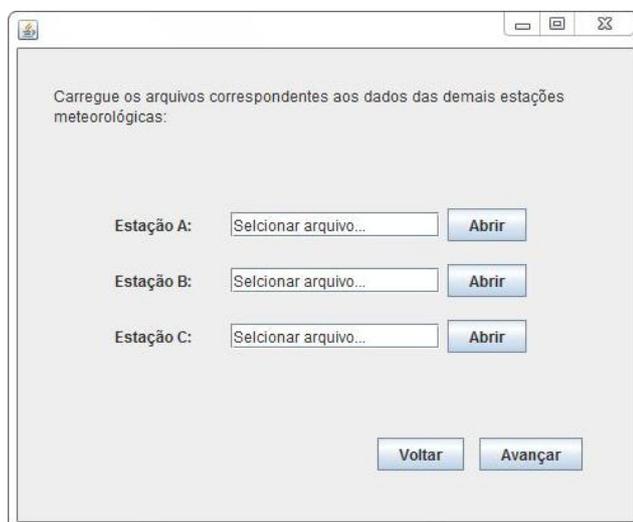


Figura 2: Tela na qual serão carregados os arquivos das estações auxiliares

No que se refere à estação de Passos, 2910 dados foram analisados – do período correspondido entre o dia 18 de julho de 2006 até o dia 05 de julho de 2014. Dentre esses, haviam 231 dados faltantes, dos quais 188 dados foram recuperados utilizando-se das aferições das estações de Machado, Bambuí e São Sebastião do Paraíso. A Figura 3 mostra a tela contendo os resultados.

Ou seja, para a estação de Passos, cerca 81,4% dos dados faltantes foram recuperados. O resultado foi, portanto, significativo para aplicação em possíveis análises meteorológicas.

Índice	Ano	Valor	
10	09	2012	0
11	09	2012	0
12	09	2012	0
13	09	2012	0
14	09	2012	0
15	09	2012	0
16	09	2012	0
17	09	2012	0
18	09	2012	0
19	09	2012	0
20	09	2012	1,3
21	09	2012	1,2
22	09	2012	21,6
23	09	2012	0
24	09	2012	39,6
25	09	2012	13
26	09	2012	null
27	09	2012	0
28	09	2012	0
29	09	2012	0
30	09	2012	0
01	10	2012	0

Total de dados analisados: 2910  
 Dados faltantes: 231  
 Dados recuperados: 188  
 Dados não recuperados: 43

Voltar   Gerar Arquivo   Fechar

Figura 3: Resultados dos dados para a cidade de Passos-MG

A Figura 4 mostra os valores de dados preenchidos pelo programa, a média histórica da estação e o valor máximo de precipitação encontrada durante todo o período. Observa-se que os dados preenchido ficaram em torno da média, não havendo preenchimentos com valores altos, fora representação dos dados de precipitação.

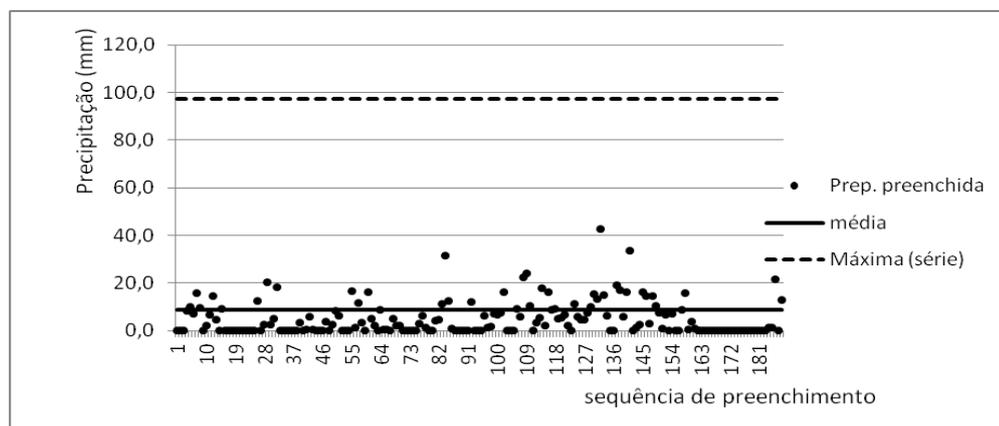


Figura 4: Dados preenchidos, média e valor máximo de precipitação para a série histórica da cidade de Passos.

## CONCLUSÕES

O algoritmo apresentou excelentes resultados em termos de processamento, não gastando mais que alguns segundos para recuperar todos os dados possíveis para a estação meteorológica da cidade Passos.

Outros dois algoritmos foram utilizados para a resolução do mesmo problema. Um deles utilizava apenas estruturas de dados lineares dinâmicas – ArrayList -,

sendo que o tempo de execução chegava a ultrapassar um minuto. O outro algoritmo utilizava a classe HashMap para armazenar os dados e, entretanto, necessitava da reordenação dos dados para que os dados pudessem ser apresentados ao final do processo. Posto nestes termos, a solução implementado TreeMaps se mostrou superior.

O software apresentado poderá ser aplicado, posteriormente, como módulo de um software de monitoramento agrometeorológico e análise climática. Além disso, futuramente serão implementados métodos adicionais de recuperação de dados, dentre os quais o método de regressão múltipla.

## REFERÊNCIAS

BERTONI J.C.; TUCCI, C. E. M. **Hidrologia: Ciência e aplicação**. Porto Alegre: UFRGS, 2007.

LUCIO, P. S.; CONDE, F. C.; CAVALCANTI, I. F. A.; RAMOS, A. M., CARDOSO, A. O. Reconstrução de séries meteorológicas via redes neurais artificiais. In: Anais XIV CBMET, Florianópolis — SC. 2006.

MAIER, H., DANDY, G. Neural networks for the predictions and forecasting of water resources variables: review of modeling issues and applications. **Environmental Modelling & Software** , v. 15, n. 1, p.101-124, 2000.

SILVA, V. S. V. **Estimativa de precipitação pontual em diferentes escalas para uso em modelo concentrado chuva-vazão**. 2011. 139 p. Dissertação (Mestrado). Programa Pós-graduação em Recursos Hídricos e Saneamento Ambiental. UFRGS. Porto Alegre.