

RECUPERAÇÃO DIGITAL DE DADOS METEOROLÓGICOS

**Giampaolo Queiroz PELLEGRINO¹, Jurandir ZULLO Junior¹,
Maria Helena MELLO², Hilton Silveira PINTO¹**

RESUMO

O objetivo deste trabalho é apresentar a viabilidade da utilização de técnicas de processamento digital na recuperação de dados registrados em fitas de papel por equipamentos meteorológicos convencionais. São descritos dois métodos baseados na utilização de mesas digitalizadoras e "scanners" considerando a formulação da correção da curvatura dos gráficos originais e avaliando sua eficiência em relação aos métodos tradicionais. A grande agilidade desses dois métodos permite recomendar sua utilização na recuperação de bancos de dados ainda no formato analógico, provavelmente a maioria dos dados por todo o Brasil.

PALAVRAS CHAVE: METEOGRAMAS, DIGITALIZAÇÃO, BANCO DE DADOS

INTRODUÇÃO

É perfeitamente clara para qualquer profissional da área agrícola a necessidade de dados meteorológicos e modelos que os representem e permitam o desenvolvimento de uma agricultura mais eficiente e produtiva. Para isso são necessárias boas séries históricas de dados que já existem, porém, em forma de meteogramas, que são as fitas de papel dos aparelhos registradores e que estão sendo deteriorados por traças ou más condições de condicionamento. Esses dados precisam ser digitalizados para que não se percam.

A digitalização corresponde à conversão dos dados coletados no formato analógico para o formato digital. Essa conversão, quando acontece, geralmente é feita através da digitação convencional via teclado, o que demanda extrema atenção e tempo na interpretação visual. Duas alternativas de digitalização podem facilitar e agilizar essa conversão: a mesa digitalizadora e o "scanner". A principal dificuldade nesses dois tipos de digitalização é que ambas as formas assumem eixos retilíneos, o que não ocorre nos meteogramas que têm o eixo y circular devido à

¹ Centro de Ensino e Pesquisa em Agricultura, CEPAGRI/UNICAMP. E-mail: giam@cpa.unicamp.br; jurandir@cpa.unicamp.br; hilton@turing.unicamp.br

² Centro de Ecofisiologia e Biofísica, IAC. E-mail: mhelena@cec.iac.br

alavanca do braço dos aparelhos registradores. Com isso, torna-se necessária uma transformação nos valores de x digitalizados para que eles representem corretamente os horários associados aos seus valores de y no eixo circular, que correspondem aos valores do elemento meteorológico em questão. O objetivo deste trabalho é portanto, realizar a formulação teórica para essa transformação ou correção e avaliar sua utilização na digitalização de meteogramas através de mesas digitalizadoras e de “scanners”.

MATERIAL E MÉTODOS

A metodologia utilizada é composta de duas etapas principais:

Determinação da correção da curvatura

No sistema cartesiano que tem origem no canto inferior esquerdo do meteograma, representado pela figura 1, pode-se perceber que quanto mais distante do valor de y central, maior a correção necessária no eixo x. Essa correção é tanto menor quanto maior o raio da curvatura, ou seja, a correção é função do raio e inversamente proporcional a ele. Portanto, o primeiro passo é determinar o raio de curvatura do aparelho. Com base na figura 1, o raio pode ser determinado pela relação trigonométrica do triângulo ΔABC :

$$\overline{AC}^2 = \overline{BC}^2 + \overline{AB}^2 \Rightarrow R^2 = (R - X_A)^2 + \left(\frac{Y_A}{2}\right)^2$$

onde, R é o raio de curvatura dado pelo comprimento do braço do aparelho;

X_A é o valor de X digitalizado para o ponto A do arco AA' ; e

Y_A é o valor de Y digitalizado para o ponto A do arco AA' .

Desenvolvendo, obtém-se:

$$R = \frac{X_A}{2} + \frac{Y_A}{8X_A}$$

A utilização do ponto A é conveniente pois ele já fornece o valor máximo de Y e simplifica o cálculo de R. O valor de Y_A também será usado no cálculo do valor da correção apresentado adiante.

Tomando-se agora os triângulos ΔDEC e $\Delta D'EC$, sendo D qualquer ponto pertencente ao arco AA' , pode-se concluir que a correção valor de x digitalizado para qualquer ponto que tenha o mesmo valor de y que o ponto D é dada por:

$$\overline{DC}^2 = \overline{EC}^2 + \overline{DE}^2 \Rightarrow R^2 = (R - X_D)^2 + \left(Y_D - \frac{Y_A}{2}\right)^2 \Rightarrow$$

$$\Rightarrow (R - X_D)^2 = R^2 - \left(Y_D - \frac{Y_A}{2}\right)^2 \Rightarrow X_D = R - \sqrt{R^2 - \left(Y_D - \frac{Y_A}{2}\right)^2}$$

onde, Y_D é valor de Y digitalizado para o ponto D pertencente ao arco AA' . Como D e D' são simétricos em relação ao valor de Y central, tem-se que a correção para o ponto D é igual à do ponto D' e podem ser dadas pela equação:

$$X_D = X_{D'} = R - \sqrt{R^2 - \left|Y_D - \frac{Y_A}{2}\right|^2}$$

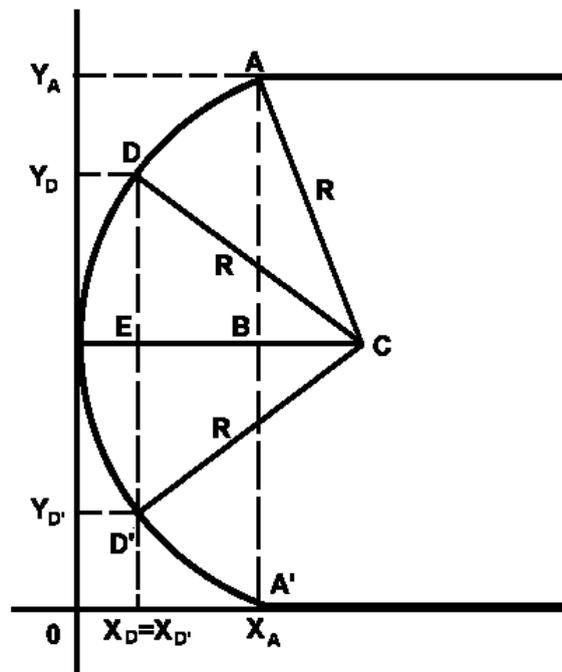


Figura 1- Esquema dos pontos para o cálculo da correção da curvatura

Para fazer a correção no valor de x de cada ponto digitalizado, utiliza-se esta última equação substituindo Y_D pelo seu valor y. Esse valor corrigido de x, e também o valor de y, já podem ser fornecidos nas unidades originais do meteograma na maioria dos "softwares" de digitalização. Porém, é comum que esteja em alguma unidade métrica que necessite de conversão. Isso exige a digitalização de um segundo ponto para se poder relacionar a variação métrica à variação em dia e hora, para o eixo x, e na unidade do elemento meteorológico observado, para o eixo y ($^{\circ}\text{C}$ para temperatura, por exemplo). Uma observação importante é que, se o valor de y não variar

linearmente é preciso estabelecer uma correção também para y. Esse é o caso da umidade relativa e deverá ser tratado em trabalho posterior envolvendo a digitalização de meteogramas mais complexos.

Digitalização dos Dados

-Através de mesa digitalizadora

A mesa digitalizadora fornece ao computador dados no formato cartesiano, ou seja, através de pares de coordenadas (x,y), normalmente já em horas e na unidade do elemento meteorológico de interesse.

-Através de "scanners"

O arquivo de saída gerado pelo scanner é uma imagem colorida onde o valor digital de cada ponto é uma composição de níveis de cinza nas três cores fundamentais (vermelho, verde e azul). Essas imagens digitais coloridas necessitam de um processamento adicional para que possam gerar os dados adequados à recuperação automática. O objetivo desse processamento é extrair da imagem apenas as informações referentes ao gráfico produzido pelo aparelho registrador, eliminando outras informações desnecessárias. Para que isso seja possível, é necessário haver uma nítida diferenciação entre a cor do gráfico e cor do "fundo" (papel e demais informações contidas nele). Isso pode ser conseguido separando-se a imagem original colorida em suas componentes fundamentais (vermelha, verde e azul), selecionando-se a seguir aquela cujos tons de cinza discriminam melhor o gráfico do seu "fundo". Como exemplo tem-se que no caso de um gráfico azul registrado em papel claro com linhas e textos avermelhados, a componente vermelha provavelmente será a mais adequada pois o gráfico tende a ficar escuro e o "fundo" mais claro. Analisando-se o histograma do canal selecionado, define-se um valor de corte ("threshold") para transformar a imagem de 256 para 2 níveis de cinza (imagem binária em preto e branco) correspondentes ao gráfico e ao fundo. Após a segmentação da imagem é recomendável afinar a linha obtida para melhorar a precisão dos resultados, utilizando algoritmos de geração de esqueletos. A imagem resultante desses processamentos deve então ser convertida para uma matriz cujas células contêm o valor 0 nos pontos correspondentes ao gráfico e o valor 1 para o fundo. Essa matriz, armazenada em um arquivo no formato texto, já está adequada para a recuperação automática. Observa-se que todos os processamentos descritos nessa etapa são normalmente encontrados na maioria dos programas de tratamento de imagens digitais disponíveis.

A recuperação automática, no caso da utilização de imagens digitais, foi feita neste trabalho através de um programa escrito em linguagem FORTRAN que calcula o valor do parâmetro atmosférico de interesse e o horário de aquisição a partir da matriz gerada no passo anterior e dos seguintes dados gerais de entrada: tamanho do braço do aparelho (mm), resolução da digitalização

(pontos por polegada), coordenadas do ponto inicial na imagem (linha e coluna), valores do ponto inicial (parâmetro e horário de aquisição), distância do eixo central do gráfico aos pontos extremos superior ou inferior (mm).

DISCUSSÃO E AVALIAÇÃO DOS MÉTODOS

Digitalização através de mesa digitalizadora

A configuração da mesa, quanto à resolução e o modo de envio dos dados, ajudam a manter a precisão nos dados digitais obtidos. Esse método exige apenas um técnico atencioso que tenha o devido cuidado de evitar ondulações ao posicionar o meteograma sobre a mesa e a atenção necessária para a manutenção da precisão da posição do cursor, seguindo fielmente a linha traçada pelo aparelho. Não é necessário entretanto, que esse técnico possua conhecimentos mais avançados sobre processamento de dados ou de imagens, sendo o principal problema para esse método de digitalização, a possibilidade de desvios dos valores digitalizados em relação à linha. Esses desvios normalmente são avaliados pelo principais "softwares" de digitalização pela Razão Média Quadrática (RMQ), que é dada por:

$$RMQ = \sqrt{\frac{\sum (x_o - xd)^2}{\sum x_o^2}}$$

onde, x_o é o valor real observado pelo aparelho fornecido como de pontos de controle; e xd é o valor fornecido pela mesa digitalizadora.

Digitalização através de "scanners"

Esse método também exige um posicionamento adequado do meteograma sobre o "scanner" pois um leve deslocamento pode provocar distorção em todos os pontos da linha digitalizada e não apenas em pontos específicos como no caso da mesa digitalizadora. Esse método evita que o operador possa provocar erros ao se desviar da linha observada pelo aparelho. Porém, por exigir o processamento da imagem para extrair a linha digitalizada, exige do técnico um conhecimento mais avançado e discernimento para escolher a cor e valor de corte adequado, ao menos uma vez para os gráficos de um mesmo aparelho, ou seja, exige uma análise da imagem e talvez alguns testes para obter o processamento mais adequado. Porém, uma vez definido basta digitalizar a imagem e operar o programa de extração da linha.

Além da facilidade e agilização, ou até da própria a possibilidade de utilização desse banco "morto" de dados meteorológicos, a grande vantagem desses dois métodos está na rapidez na

geração dos dados com a mesma qualidade das análises visuais dos meteorogramas. Testes foram feitos com dez pessoas de diferentes graus de experiência na leituras desses meteorogramas mostraram que uma interpretação completa de um termograma dura em média 40 minutos. Selecionando o grupo mais experiente, esse tempo cai para 25 minutos. A digitalização através da mesa leva em média de 5 a 10 minutos, incluindo-se o tempo para posicionamento do papel e operação do software utilizado. A digitalização da linha toda propriamente dita dura em torno de 2 ou 3 minutos. O método do "scanner" possui um tempo inicial relativamente alto de definição dos parâmetros para os gráficos de um determinado tipo de aparelho, podendo durar um dia todo. Após essa definição, o tempo de geração dos pares (x,y) digitalizados fica próximos aos da mesa digitalizadora. Portanto, os dois métodos permitem uma grande agilização na digitalização dos dados permitindo que uma maior quantidade de dados seja processada melhorando a eficiência da operação e gerando bancos de dados digitais, dificilmente degradáveis. Convém ressaltar que a análise da RMQ permite afirmar que, tomando a média geral dos dados analisados visualmente como verdade, os dois métodos apresentaram os mesmos índices de precisão das leituras individuais.

CONCLUSÕES

Embora o caráter desse estudo seja eminentemente de avaliação da viabilidade da utilização desses dois métodos de digitalização de meteorogramas, ainda que sem a obtenção exaustiva e operacional de dados, fica claro que eles permitem a agilização e até a viabilização da recuperação digital de bancos de dados abandonados com velocidades pelo menos 5 vezes maiores que os métodos convencionais via teclado. Ainda, sem perda na qualidade dos dados.