Rev. Bras. Agrometeorologia, v. 13, n.2, p.273-283, 2005
Recebido para publicação em 14/04/04. Aprovado em 08/08/05.

ISSN 0104-1347

Ajuste de modelos de distribuição de probabilidade a séries históricas de precipitação pluvial diária em Piracicaba-SP

Adjustment of models of probability distribution to the historical daily rainfall series in Piracicaba, Brazil

Durval Dourado Neto[1,6]; Janilson Pinheiro de Assis[2,6]; Luiz Carlos Timm[3]; Paulo Augusto Manfron[4,6]; Gerd Sparovek[5,6] ; Thomas Newton Martin[2,6]

Resumo: No presente trabalho foram analisadas cinco distribuições densidade de probabilidade (Gama, Exponencial, Weibull, Log-normal e Normal). O teste de qui-quadrado foi utilizado para verificar a aderência das probabilidades estimadas às freqüências observadas (5% de significância). Foram consideradas, para fins de análise, apenas cada série de precipitação pluvial diária de dias semelhantes (por exemplo, apenas os dados de 1 de janeiro, de 2 de janeiro,..., e de 31 de dezembro) separadamente em Piracicaba-SP. Como período chuvoso, consideraram-se os meses de janeiro, fevereiro, março, outubro, novembro e dezembro. Observou-se a superioridade do ajustamento da distribuição Gama e Weibull, com predominância da Gama em relação a Weibull. Verificou-se para algumas datas a aderência das distribuições Log-normal e Exponencial, as quais predominam nos meses de fevereiro, março, outubro e novembro para o modelo Exponencial, e março, novembro e dezembro para o modelo Log-normal. Verificou-se ainda que em algumas séries, nenhum dos modelos estudados se ajustou aos dados.

Palavras-chave: Chuva, função densidade de probabilidade, previsão.

Abstract: In the present study the adjustment of rainfall data series of Piracicaba, São Paulo State, Brazil to five probability density distributions (gamma, exponential, Weibull, Log-normal and Normal) was analyzed. The qui-square test (5 % of significance level) was used to verify the accuracy of the ajusted probabilities functions when compared with the observed frequencies. The daily rainfall series of similar days (for example, only January 1st, February 2nd, ..., December 31st) was considered for the following months: January, February, March, October, November, and December, which represents the rainy season. In general, better adjustments were observed for gamma and Weibull distributions, being the gamma distribution predominant. In some days of February, March, October, November and December the Log-normal and exponential distributions, fitted the data series well, predominantly the exponential one in February, March, October and November; and the Log-normal in March, November, and December. In some series, the studied models did not adjust well.

Key words: rainfall, probability density function, forecasting.

[1] Engenheiro Agrônomo, Professor Associado, Departamento de Produção Vegetal, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo. Caixa Postal 9. 13.418-900. Piracicaba, SP. dourado@esalq.usp.br
[2] Engenheiro Agrônomo, Aluno do Programa de Pós-Graduação em Fitotecnia, Departamento de Produção Vegetal, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo. Caixa Postal 9. 13.418-900. Piracicaba, SP.
[3] Engenheiro Agrônomo, Pesquisador, Centro de Energia Nuclear na Agricultura, Universidade de São Paulo. Caixa Postal 9. 13.418-900. Piracicaba, SP.
[4] Engenheiro Agrônomo, Professor Titular, Departamento de Fitotecnia, Universidade Federal de Santa Maria, 97105-900.Santa Maria, RS.
[5] Engenheiro Agrônomo, Professor Associado, Departamento de Solos e Nutrição de Plantas, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo. Caixa Postal 9. 13.418-900. Piracicaba, SP.
[6] Bolsista CNPq

Introdução

O estudo das distribuições de variáveis climáticas, ao longo do tempo, como um meio de compreender os fenômenos meteorológicos, determinando seus padrões de ocorrência e permitindo uma previsibilidade razoável do comportamento climático de uma região é uma ferramenta de grande valor para o planejamento e gestão de inúmeras atividades agropecuárias e humanas.

O aproveitamento otimizado dos recursos hídricos requer o conhecimento de técnicas de planejamento que se baseiam na estimativa das probabilidades associadas a determinadas variáveis hidrológicas, nesse caso a precipitação pluvial. Nessas atividades, o principal interesse é prever, com base em funções de densidade de probabilidade aplicadas aos dados observados, as precipitações capazes de ocorrer em uma certa região.

O uso de funções densidade de probabilidade está diretamente ligado à natureza dos dados a que elas se relacionam, onde algumas têm boa capacidade de estimação para pequeno número de dados, outras requerem grande série de observações. Devido ao número de parâmetros de sua equação, algumas podem assumir diferentes formas, enquadrando-se em um número maior de situações, ou seja, são mais flexíveis. Desde que respeitado o aspecto da representatividade dos dados, as estimativas dos seus parâmetros, para uma determinada região, podem ser estabelecidas como de uso geral, sem prejuízo da precisão na estimação da probabilidade (CATALUNHA et al., 2002).

Informações sobre precipitação pluvial, elemento meteorológico disponível na maior parte das estações meteorológicas, é útil tanto no planejamento agrícola em curto prazo (práticas agronômicas cuja umidade do solo e/ou do ar são condicionantes) como em longo prazo (definições das regiões e épocas mais adequadas para semeadura de culturas), conforme descrevem FONTANA & ALMEIDA (2002). Nesse contexto, destaca-se a importância de se avaliar e quantificar a variabilidade do número de dias com precipitação pluvial, utilizando uma análise estatística adequada.

O ajuste de modelos probabilísticos aos dados diários de chuva além de fornecer um resumo sucinto destes dados, representa uma técnica eficiente para a análise dessas informações. Cada distribuição de freqüência apresenta uma certa forma, a qual pode ser aproximada via utilização de função densidade de probabilidade com os parâmetros extraídos da amostra em estudo. A utilização ou não de uma distribuição reside na capacidade da mesma em estimar os dados observados, com base em seus parâmetros e essa capacidade é medida com a aplicação de testes de aderência (ALMEIDA, 1995). Estudos de ajuste de distribuição densidade de probabilidade com dados diários em séries de dias semelhantes (por exemplo, apenas o dia 1 de cada mês, é muito raro na literatura especializada), apresentando uma lacuna nesse tipo de investigação, por sinal muito importante em estudos de rendimento de culturas agrícolas.

Segundo MORETTIN & BUSSAB (2003), a Normal é uma das mais importantes variáveis aleatórias contínuas, cuja distribuição é chamada distribuição Normal ou Gaussiana, a qual serve como modelo de distribuição para muitos problemas da vida real, mas também aparece em muitas investigações teóricas, pois sua importância em análise matemática resulta do fato de que muitas técnicas estatísticas, como análise de variância e de regressão além de alguns testes de hipóteses, assumem ou exigem a Normalidade dos dados.

A distribuição Weibull é utilizada em análise hidrológica para eventos extremos. Contudo, sua utilização em séries históricas de variáveis climáticas ainda é pouco conhecida. Segundo CATALUNHA et al. (2002), essa distribuição foi proposta primeiramente por Fisher e Tippet em 1928, tendo sido desenvolvida independentemente por Walodi Weibull, em 1939. O trabalho de Weibull se destacou e a distribuição passou a ser chamada pelo seu nome (JOHNSON & KOTZ, 1970).

O presente trabalho tem por objetivo avaliar o ajuste das funções densidade de probabilidade Gama, Exponencial, Weibull, Normal e Log-normal aos dados diários de precipitação pluvial, em Piracicaba-SP.

Material e Métodos

Os dados utilizados no presente estudo foram fornecidos pelo Setor de Agrometeorologia do

Departamento de Ciências Exatas da Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, em Piracicaba, Estado de São Paulo. A estação agrometeorológica está situada à latitude de 22°42'30"S, à longitude de 47°38'30"W, e com altitude de 546 m (VILLA NOVA, 2003).

O clima da região é do tipo Cwa, segundo a classificação de Köppen: tropical úmido com chuvas de verão e seca no inverno, caracterizado por um total de chuvas no mês mais seco de 26 mm e do mês mais chuvoso de 217 mm, por uma temperatura média do mês mais quente de 24,6°C.

As observações utilizadas neste trabalho se referem às precipitações pluviais (mm) diárias de dias considerados semelhantes (por exemplo: 1 de janeiro, 2 de janeiro,..., 31 de dezembro) referentes aos meses de janeiro, fevereiro, março, outubro, novembro e dezembro. As séries históricas de chuva abrangeram o período de 1 de janeiro de 1917 a 31 de dezembro de 2002 num total de 86 anos.

Os dados foram analisados individualmente em cada dia dos meses de janeiro, fevereiro, março, outubro, novembro e dezembro para cada ano observado da série histórica estudada, sendo ajustadas cinco distribuições densidade de probabilidade, utilizando o sistema SAS (SAS INSTITUTE, 1996).

Foi utilizado um número de anos superior ao mínimo de 30, bem como foi utilizado o "run test" para verificar a distribuição aleatória dos dados, conforme preconizado pela Organização Meteorológica Mundial (THOM, 1966). Foram consideradas, para fins de ajuste às diferentes distribuições referentes à análise da precipitação pluvial diária, apenas os dias com chuva (dias com precipitação igual ou superior a 0,1 mm). Para verificar o grau de aderência das distribuições empíricas aos modelos das distribuições Gama, Exponencial, Weibull, Log-normal e Normal, foi aplicado o teste de qui-quadrado (nível de significância de 5 %).

Foi aplicado o seguinte modelo, conforme utilizado por FONSECA & ALBUQUERQUE (1978) para a região de Pelotas-RS:

$$G(X) = Po + (1 - Po)\, F(X) \qquad (1)$$

em que Po se refere à probabilidade de ocorrência de valores nulos (zeros) de precipitação pluvial diária; e F(X) à distribuição cumulativa de probabilidade, onde os parâmetros são estimados na ausência de zeros, conforme preconizaram ASSIS et al. (1996).

A variável aleatória *X*, que toma todos os valores reais, tem uma distribuição Normal (ou Gaussiana) se sua função densidade de probabilidade for da forma:,

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

em que μ se refere à média das observações na série de dados, σ ao desvio padrão das observações na série de dados; e α ao nível de significância.

A probabilidade da variável aleatória *X* assumir valores entre *a* e *b* quando a mesma possua distribuição Normal com média μ e variância σ^2 [$X \cap N(\mu, \sigma^2)$] foi estimada (MEYER, 1969) por:

$$F(x) = P(a \le X \le b) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{\frac{(x-\mu)^2}{2\sigma^2}}\, dx = 1 - \alpha \qquad (3)$$

Essa equação não pode ser resolvida analiticamente sem o uso de integração aproximada. Por esse motivo, utilizou-se a transformação $z=(x-\mu)/\sigma$, considerando que essa variável aleatória Z tem distribuição Normal padrão com média nula e variância unitária [$Z \cap N(0, 1)$]. A variável Z é chamada reduzida, e a equação (1) pode ser reescrita na seguinte forma:

$$f(z) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2} z^2\right], \ (-\infty < Z < +\infty) \qquad (4)$$

A função densidade de probabilidade da distribuição Log-normal [f(x)] é representada pela seguinte equação (MIRSHAWKA, 1971).

(5)

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

em que x se refere ao valor da variável aleatória; μ à média dos logaritmos da variável aleatória X; σ ao desvio padrão dos logaritmos da variável aleatória X; e a ao limite inferior da amostra. Na Distribuição Log-normal, os logaritmos das variáveis aleatórias são Normalmente distribuídos. Neste trabalho foi utilizada a distribuição Log-normal com três parâmetros.

Para encontrar a probabilidade de que uma variável aleatória X, tendo distribuição Log-normal, assuma valores entre a e b ($a \leq X \leq b$), foi usada:

$$P(a\leq X\leq b)=\int_a^b \frac{1}{(x-a)\sigma\sqrt{2\pi}}\exp\left[-\frac{[\ln(x-a)-\mu]^2}{2\sigma^2}\right]dx=1-\alpha \quad (6)$$

em que a pode ser zero, quando se considera a distribuição Log-normal com dois parâmetros, ou um valor mínimo da série, quando se considera Log-normal com três parâmetros; b pode ser o valor da variável aleatória, quando se considera a probabilidade cumulativa de ocorrência daquele valor (HASTINGS & PEACOCK, 1975).

Segundo BOTELHO & MORAIS (1999), uma variável aleatória contínua X $(X>0)$ se distribui segundo uma distribuição Gama de parâmetros α ($\alpha>0$) e β ($\beta>0$) se sua função de densidade de probabilidade é:

$$f(x)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}}\ ,\ (X>0) \quad (7)$$

$$f(x)=0\ ;\ (X\leq 0)$$

em que $P(X\leq x)$ se refere à probabilidade de ocorrer um valor $X\leq x$ (probabilidade de ocorrer uma quantidade de precipitação pluvial igual ou inferior a x); X se refere à variável aleatória contínua que representa os valores das precipitações; α ao parâmetro de forma da variável aleatória X; β ao parâmetro de escala da variável aleatória X (mm); e ao número de Neper (2,718...); x à quantidade de chuva (mm) e $\Gamma(\alpha)$ à função Gama incompleta, assim definida:

$$\Gamma(\alpha)=\int_0^{\infty} e^{-x}\cdot x^{\alpha-1}dx \quad (9)$$

Fazendo-se uso da série de Taylor, obtém-se:

$$F(t)=\frac{t^{\hat{\alpha}}}{\hat{\alpha}\,\Gamma(\hat{\alpha})e^{t}}\left[1+\frac{t}{\hat{\alpha}+1}+\frac{t^2}{(\hat{\alpha}+1)(\hat{\alpha}+2)}+\frac{t^3}{(\hat{\alpha}+1)(\hat{\alpha}+2)(\hat{\alpha}+3)}+...\right] \quad (10)$$

em que $F(t)$ permite o cálculo aproximado da probabilidade de ocorrência de um valor menor ou igual a t, onde:

$$x=t\cdot\hat{\beta} \quad (11)$$

As estimativas dos parâmetros α e β foram efetuadas pelo método da máxima verossimilhança (ASSIS et al., 1996), produzindo estimativas eficientes dos parâmetros estatísticos. THOM (1958), usando este método derivou as equações para estimativas dos parâmetros da distribuição Gama através da resolução da seguinte equação quadrática:

$$12A\ \alpha^2-6\alpha-1=0 \quad (12)$$

$$\hat{\alpha}=\frac{1}{4A}\left(1+\sqrt{1+\frac{4A}{3}}\right) \quad (13)$$

A estimativa para o parâmetro de escala foi feita por:

$$\hat{\beta}=\frac{\bar{x}}{\hat{\alpha}} \quad (14)$$

sendo que:

$$A=\ln\bar{x}-x_g \quad (15)$$

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \quad (16)$$

$$x_g=\frac{1}{n}\sum_{i=1}^{n}\ln(x_i) \quad (17)$$

Rev. Bras. Agrometeorologia, v. 13, n.2, p.273-283, 2005

em que $\bar{x}$ se refere à média aritmética das observações, mm; x_g à média geométrica das observações, mm; x_i ao i-ésimo valor da precipitação pluvial (mm) diária; e, n ao número de observações.

A distribuição Exponencial é geralmente aplicada a dados com forte assimetria (forma de "J" invertido). Sua função densidade de probabilidade é descrita conforme a seguinte equação (KITE, 1978):

$$f(x) = \lambda e^{-\lambda X}\ ;\ (X > 0) \qquad (18)$$
$$f(x) = 0\ ;\ (X \leq 0)$$

Sua função de distribuição acumulada é do tipo:

$$F(x) = P(X \geq 0) = \int_0^\infty f(x)dx = 1 - e^{-\lambda x} = 1 - \alpha \qquad (19)$$

O único parâmetro da distribuição (λ) é estimado pelo inverso da média.

Sua função densidade de probabilidade é apresentada de diversas formas, sendo comum em alguns trabalhos apresentá-la de acordo com a seguinte equação:

$$f(x) = \frac{\gamma}{\beta}\left(\frac{x-\alpha}{\beta}\right)^{\gamma-1} \exp\left[-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right]\ ;\ (\text{X e}\geq\alpha) \qquad (20)$$
$$f(x) = 0\ ;\ (X < \alpha)$$

em que X se refere à variável aleatória precipitação pluvial, a (a e" 0), b (b > 0) e g (g > 0) aos parâmetros empíricos da distribuição. Dessa forma, a distribuição é chamada de Weibull com três parâmetros. A sua função de distribuição acumulada é:

$$F(x) = P(X \geq 0) = \int_0^\infty f(x)dx = 1 - \exp\left[-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right] = 1 - \alpha \qquad (21)$$

O parâmetro a é chamado de parâmetro de posição, pois controla a posição da curva sobre o eixo da abscissa. O parâmetro de escala b controla as dimensões que a curva assume, dada uma forma constante. O parâmetro g é chamado parâmetro de forma. Algumas propriedades da Weibull são apresentadas por JOHNSON & KOTZ (1970).

Diversos procedimentos de estimar os parâmetros da distribuição Weibull foram desenvolvidos. O principal método de ajuste dessa distribuição é o da máxima verossimilhança (COUTO, 1980), o qual consiste em determinar os valores de γ e β pela suas equações fundamentais. Nota-se, entretanto que β é função de γ e que a solução desse tipo de sistema é obtido resolvendo a equação por um processo iterativo, sendo que neste trabalho foi utilizado o método da bissecção, sendo β encontrado a partir de g aproximado, utilizando o seguinte procedimento:

$$\frac{\sum_{i=1}^{n} x_i^{\hat{\gamma}} \ln(x_i)}{\sum_{i=1}^{n} x_i} - \frac{\sum_{i=1}^{n} \ln(x_i)}{n} - \frac{1}{\hat{\gamma}} = 0 \qquad (22)$$

$$\hat{\beta} = \left(\frac{\sum_{i-1}^{n} \ln\left(x^{\hat{\gamma}}{}_i\right)}{n}\right)^{\frac{1}{\hat{\gamma}}} \qquad (23)$$

Ao se ajustar uma série de dados a uma distribuição densidade de probabilidade, trabalha-se com a hipótese de que a distribuição pode representar adequadamente aquele conjunto de dados. Um critério de comprovar essa hipótese é através de alguns testes não paramétricos. No teste de aderência do qui-quadrado (χ^2), a hipótese de nulidade admite que a distribuição seja aquela especificada com os seus parâmetros estimados com base nos dados amostrais. A hipótese é testada fazendo-se a comparação entre as freqüências observadas e as freqüências teóricas ou esperadas, em cada classe de freqüência dos dados, através da estatística teste χ^2 (CAMPOS, 1983) dada por:

$$\chi^2 = \sum_{i=1}^{k} \left[\frac{(Fo_i - Fe_i)^2}{Fe_i}\right] \qquad (24)$$

em que k se refere ao número de classes; Fo_i à freqüência observada e Fe_i à freqüência esperada sob a hipótese Ho, de acordo com a distribuição que está sendo testada. Os valores críticos ou tabelados de $\div^2$ para alguns níveis de significância á são descritos por tabelas próprias.

Para verificar o grau de aderência das distribuições empíricas (distribuições Gama, Exponencial, Weibull, Log-normal e Normal) aos dados diários, foi utilizado o teste de qui-quadrado, sendo que neste trabalho adotou-se o nível de significância de 5% ($\alpha = 0{,}05$), em virtude de trabalhar-se com dados diários, os quais por natureza apresentam uma alta variabilidade, valores perdidos, além da presença de dados discrepantes. Deve-se salientar que as funções foram escolhidas dentre algumas comumente utilizadas para este tipo de análise (THOM, 1966; HASTINGS & PEACOCK, 1975).

Para a estação chuvosa (outubro a março), caracterizada pela maior probabilidade de dias chuvosos, procedeu-se o ajuste de séries históricas de precipitação pluvial diária às distribuições densidade de probabilidade Normal, Log-normal, Exponencial, Gama e Weibull, em Piracicaba-SP, onde se verificou o número absoluto e relativo de dias com aderência, conforme o teste de qui-quadrado a 5% de significância, relativo aos meses de janeiro, fevereiro, março, outubro, novembro e dezembro.

Resultados e Discussão

A caracterização da ocorrência chuva diária em Piracicaba-SP, de 1 de janeiro a 31 de dezembro, em termos de probabilidade de chover e de não chover pode ser visualizada na Figura 1. Verifica-se através dessa que a probabilidade de ocorrência de valores nulos de precipitação é maior nos meses de junho a agosto. Porém, a probabilidade de não ocorrência de valores nulos é menor, em torno de 50%, nos meses de outubro a março, ficando de acordo com o período de maior precipitação no município de Piracicaba, SP.

Para a distribuição Normal, ao observar os dados das Tabelas 1 e 2, nota-se que para as estimativas diárias a aderência à distribuição Normal é baixíssima, apenas em um dia de março verificou-se o ajustamento. Isso se deve ao fato de que quando ao analisar a distribuição de classes de freqüências de algumas séries diárias isoladamente, verifica-se que para esses valores há maior freqüência nas classes iniciais, reduzindo bruscamente a partir da segunda ou terceira classe. Essa queda não é acompanhada pelo modelo matemático da distribuição Normal (que é mais simétrico), como o são os modelos da distribuição Gama e Weibull (bastante assimétricas), superestimando os valores das classes seguintes. Resultados semelhantes foram obtidos por CATALUNHA et al. (2002). Quando se aplica o teste de hipótese, essas classes são reprovadas. Segundo CATALUNHA et al. (2002), o modelo Normal é mais apropriado para representar valores totais de chuva, sejam eles decendiais ou mensais.

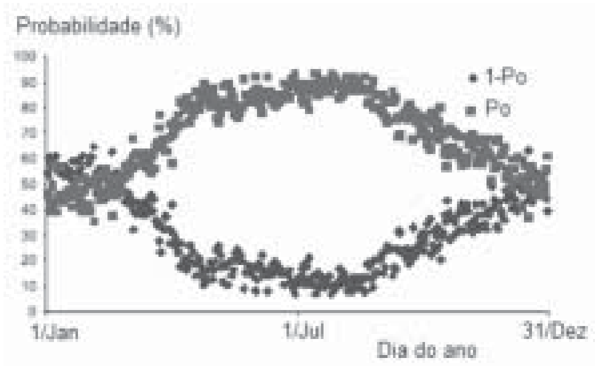

Figura 1. Probabilidade de ocorrência (Po) e de não ocorrência (1-Po) de valores nulos de precipitação pluvial diária em Piracicaba-SP (período de 1 de janeiro de 1917 a 31 de dezembro de 2002 num total de 86 anos).

Sendo assim, na simulação estocástica (o mesmo procedimento pode ser utilizado ‘n’ vezes: o usual é utilizar de 1.000 a 10.000 vezes) de chuva para a região de Piracicaba-SP, gera-se um número randômico, entre 0 (zero) e 1 (um), com distribuição uniforme. Estando o número gerado entre 0 (zero) e a probabilidade de não chover (Po/100), simular-se-á um dia seco (sem chuva). Estando o número gerado entre a probabilidade de não chover (Po/100) e 1 (um) (Figura 1), simular-se-á um dia com chuva, utilizando a distribuição de probabilidade com maior aderência naquele dia, com base na Equação 1.

Também para a distribuição Log-normal, apesar de ser bastante empregada em outras áreas de análise climatológica e hidrológica, nesse estudo ela não apresentou desempenho satisfatório na estimação das probabilidades, sendo em média apenas em algumas datas consideradas (Tabela 1), uma boa distribuição para estimação de dados nas condições e períodos estudados.

Na distribuição Exponencial os dados se concentram nas classes iniciais e diminuem a concentração nas finais. Pode-se questionar o fato de que, mesmo tendo naturalmente os dados, esta distribuição de freqüência não foi a que melhor estimou os valores observados em alguns períodos (março e novembro). Isso se deve à capacidade do modelo matemático em estimar cada classe de freqüência individualmente, quando ocorrem picos intermediários de freqüência.

A distribuição Log-normal foi testada para modelar as quantidades diárias de chuva em localidades da Jordânia, Nigéria, Botswana e Sri Lanka, tendo obtido resultados satisfatórios, conforme cita ALMEIDA (1995) nos trabalhos de Stern & Coe (1982). FARIA (1998) utilizou a distribuição Gama para estimar a precipitação dependente ao nível de 75% de probabilidade, obtendo boa aderência. Essa distribuição tem sido utilizada com freqüência na estimação da probabilidade de precipitação.

Considerando o erro relativo entre as probabilidades observadas e as estimadas, pode-se observar que as probabilidades estimadas no período com chuva são em grande parte superestimadas. Uma característica da distribuição Exponencial é a de que partindo de um valor da classe inicial, estimam-se outros valores proporcionalmente menores, formando uma curva em forma de "J" invertido. Ao analisar a distribuição de classes de freqüências de algumas datas isoladamente, nota-se que as classes intermediárias têm valores maiores que as classes iniciais, ou bem próxima desta, não formam uma curva em "J" invertido, mas sim um formato parecido com um "M". A capacidade da distribuição Exponencial é limitada em estimar essas lacunas intermediárias, superestimando-os, pois seu único parâmetro é a média, sendo assim ao se aplicar o teste de aderência, essas classes são reprovadas. Segundo CATALUNHA et al. (2002), esse modelo apresentou aderência mais freqüente em séries de dados de período secos, isso se deve segundo o autor ao fato de que nesses períodos as classes de freqüências não apresentam aqueles picos comentados anteriormente, tendo essas a forma de um "J" invertido suave, o que reduz o erro relativo na estimação, aumentando o número de aderências.

A função Gama de probabilidade possui dois parâmetros, o de forma (α) e o de escala (β) (MILLER & WEAVER, 1968). Para valores de maiores ou igual a 100, a distribuição Gama se aproxima da distribuição Normal (THOM, 1958). O parâmetro de escala (β) indica o grau de dispersão entre os dados de uma série estudada.

Um grande problema encontrado em trabalhos que envolvem a distribuição Gama é a estimação dos parâmetros α e β, devido à complexidade e extensão dos cálculos envolvidos. Vários métodos podem ser usados, como o método dos quadrados mínimos, o método dos momentos e o da máxima verossimilhança. Porém, todos possuem limitações, seja por problemas matemáticos ou por produzirem estimativas ineficientes. O método dos quadrados mínimos apresenta uma série de dificuldades quando aplicado à distribuição Gama, e não é recomendado. Os métodos da máxima verossimilhança e o dos momentos são os mais comumente utilizados, mas, segundo THOM (1958), deve-se preferir o da máxima verossimilhança devido às suas melhores propriedades. Algumas formas de estimar os parâmetros da distribuição Gama foram desenvolvidas, contribuindo, junto com a sua flexibilidade de formas, para sua utilização em diversas áreas (HAAN, 1977).

Os parâmetros da distribuição Gama tiveram valores mínimo, médio e máximo de 0,59, 0,70 e 6,09. Para as séries ajustadas nesse estudo, conforme o teste de qui-quadrado, a aderência dessa distribuição é predominante, conforme os valores contidos na Tabela 2. Sendo assim, esse modelo deve ser o recomendado para descrever a variabilidade temporal das precipitações diárias em Piracicaba-

SP. Resultados semelhantes foram obtidos por CATALUNHA et al. (2002). Sendo assim, verificou-se que no caso das estimativas diárias de probabilidade, detecta-se a superioridade do desempenho da distribuição Gama, em todas as datas estudadas.

BOTELHO & MORAIS (1999) estimaram os parâmetros da distribuição gama de dados pluviométricos de Lavras, MG, e concluíram que as estimativas dos parâmetros de forma, de modo geral são menores nos meses menos chuvosos, e maiores nos meses mais chuvosos e, além disso, aumentam de valor à medida que o tamanho do período de dias aumenta. Por outro lado às estimativas dos parâmetros de escala foram menores apenas nos períodos de 1 a 3 e 1 a 6 dias nos meses mais secos, já os maiores valores ocorreram nos meses mais chuvosos.

RIBEIRO & LUNARDI (1997) ajustaram a distribuição gama a dados diários de 34 anos (1961 a 1994). E de acordo com os autores a função de distribuição de probabilidade gama mostrou ser adequada para representar a chuva em todos os meses do ano. Os mesmos autores estimaram a participação pluviométrica mensal provável em Londrina, PR, nos níveis de 10, 20, 30, 40, 50, 70, 75, 80 e 90% de probabilidade, e verificaram que o valor médio de precipitação pluviométrica não deve ser utilizado como base na elaboração de projetos agrícolas, pois ocorre entre os níveis de 30 e 40% de probabilidade.

De acordo com os dados das Tabelas 1 e 2, verifica-se que a distribuição Weibull é uma das mais precisas para representar os valores médios diários de chuva em Piracicaba-SP, mostrando-se inclusive com desempenho bastante próximo da distribuição Gama. Resultados semelhantes foram obtidos por CATALUNHA et al. (2002), nas condições do Estado de Minas Gerais.

Os parâmetros da distribuição Weibull tiveram valores mínimo, médio e máximo de 0,60, 1,55 e 2,00 respectivamente para alfa, e de 5,09, 20,18 e 21,00 respectivamente para beta.

Sendo assim, verificou-se que, quando se obtém mais de um modelo aderido para uma determinada data estudada, deve-se optar por uma distribuição mais simples. Por exemplo, no dia 1 de fevereiro verificou-se que tanto a distribuição Log-normal, como a Exponencial, Gama e Weibull podem representar a distribuição de chuva nesse dia. No entanto, o pesquisador poderá usar um critério de parcimônia, escolhendo o modelo mais simples, versátil e flexível dentre aqueles ajustados, pois a qualidade das inferências são as mesmas.

Para atender as definições do teste de qui-quadrado, as classes estimadas com freqüência inferior a três devem ser somadas a outra classe mais próxima. Isso faz com que, segundo CATALUNHA et al. (2002), a somatória ocorrerá também nas classes de freqüências observadas, gerando um erro absoluto grande, que se somado aos anteriores, resulta em valores de qui-quadrado maiores que os tabelados, não aprovando a distribuição sob teste, quando a estimação não for boa. Isso faz com que esse teste não aprove a maioria ou qualquer distribuição, o que é um fato positivo dessa prova.

Os valores de qui-quadrado calculados são comparados com valores críticos ou tabelados obtidos de tabelas referenciadas pelo nível de significância e pelo grau de liberdade, o qual depende do número de parâmetros da distribuição, no caso igual a dois, e do número de classes (inerente aos dados). Analisando o fato de que quando ocorrem agrupamentos de classes para evitar o uso de freqüências menores que três ou cinco, este número reduz quando a distribuição subestima as classes finais, devido a este agrupamento mencionado anteriormente, e o grau de liberdade fica menor, reduzindo o valor tabelado para o qui-quadrado. Isso mostra que o valor crítico ou tabelado para o qui-quadrado depende da capacidade da distribuição em estimar as freqüências observadas; o mesmo não ocorre com o teste de Kolmogorov-Smirnov, o qual tem também o mesmo objetivo do teste de qui-quadrado.

Os resultados foram apresentados para os meses de janeiro, fevereiro, março, outubro, novembro e dezembro por serem os meses com maior probabilidade de ocorrência de precipitação pluvial (Tabela 1).

Utilizou-se o teste de qui-quadrado por ser considerado mais rigoroso do que o teste de Kolmogorov-Smirnov, onde as principais limitações

Rev. Bras. Agrometeorologia, v. 13, n.2, p.273-283, 2005

Tabela 1. Desempenho de seis métodos de estimativa da evapotranspiração de referência com base nos índices estatísticos eficiência de modelagem (EF), índice de concordância de Willmott (ICW), coeficiente de massa residual (CRM) e raiz quadrada do erro médio (RMSE).

Dia	Mês					
	Janeiro	Fevereiro	Março	Outubro	Novembro	Dezembro
1	E-G-W	L-E-G-W	G	L-E-G-W	*	L-G
2	G	G-W	G-W	E-G-W	E-G	L
3	L-E-G-W	L-G-W	L-G-W	*	*	L-E-G-W
4	G-W	L-G-W	L-E-G-W	G	L-G	E-G-W
5	L-E-G-W	E-G-W	*	L-E-G-W	L-E-G-W	L-E-G-W
6	L-E-G-W	L-E-G-W	N-E-G-W	E-G-W	L-E-G-W	L
7	L-E-G-W	L-E-G-W	E-G-W	E-G-W	L-E-G-W	L-E-G-W
8	G	L-G-W	E-G-W	L-E-G-W	E-G-W	L-E-G-W
9	L-E-G-W	L-E-G-W	L-G-W	L-G-W	E-G-W	G
10	G	E-G-W	L-E-G-W	E-G-W	L-E-G-W	G
11	*	E-G-W	L-E-G-W	G	L-G-W	E-G-W
12	*	E-G	L-E-G-W	G	L-E-G-W	L-G-W
13	L-E-G-W	L-E-G-W	L-G-W	G	G	L-E-G-W
14	L-E-G-W	L-E-G-W	L-E-g-W	L-E-G-W	E-G-W	L-E-G-W
15	E-G-W	E-G-W	L-E-g-W	L-E-G-W	L-E-G-W	L-E-G-W
16	*	G-W	L-E-g-W	E-G-W	G	E-G-W
17	G	L-E-G-W	L-E-g-W	E-G-W	G	E-G-W
18	*	E-G-W	E-G-W	E-G-W	L-E-G-W	L-E-G-W
19	L-E-G-W	E-G-W	L-E-G-W	L-E-G-W	L-E-G-W	L-E-G-W
20	L-G-W	E-G	E-G-W	L-E-G-W	L-G-W	L
21	*	G-W	E-G-W	G	L-E-G-W	L-G-W
22	G	L-G-W	L-E-G-W	L-E-G-W	L-E-G-W	*
23	L-E-G-W	L-G-W	L-E-G-W	L-E-G-W	L-E-G-W	L-G-W
24	G	L-G-W	L-G-W	L	L-E-G-W	L-G-W
25	L-E-G-W	L-G-W	L-E-G-W	L-E-G-W	L-E-G-W	E-G-W
26	E-G-W	G	L-E-G-W	G	*	E-G-W
27	*	E-G-W	L-E-G-W	E-G	E-G-W	*
28	E-G-W	*	L-E-G-W	L-E-G-W	L-E-G-W	G
29	L-E-G-W	E-G	L-E-G-W	L-E-G-W	E-G-W	L-G
30	L-E-G-W	-	L-E-G-W	L-E-G-W	L-E-G-W	L-E-G-W
31	G-W	-	N-L-E-G-W	L	-	L-E-G-W

* Nenhum dos cinco modelos se ajustou aos valores observados.

Tabela 2. Freqüências absolutas e percentuais do número de aderências em cada mês referentes às distribuições densidade de probabilidade Normal, Log-normal, Exponencial, Gama e Weibull, de acordo com o teste de qui-quadrado (5% de significância).

Distribuição	Mês					
	Janeiro	Fevereiro	Março	Outubro	Novembro	Dezembro
Gama	25(81%)	29(100%)	30(97%)	28(90%)	27(90%)	26(84%)
Weibull	19(61%)	25(86%)	29(94%)	20(65%)	24(80%)	22(71%)
Exponencial	16(52%)	18(62%)	23(74%)	20(65%)	21(70%)	17(55%)
Log-normal	13(42%)	15(52%)	22(71%)	16(52%)	19(63%)	21(68%)
Normal	0(0%)	0(0%)	1(3%)	0(0%)	0(0%)	0(0%)

para seu uso foram observadas, tais como: (i) só pode ser aplicado quando os dados forem agrupados, (ii) se houver classes que possuam valores menores que três ou cinco, esses devem ser agrupados em outras classes, sendo um fator limitante para uso em série de dados com poucas classes.

Verifica-se para as estimativas diárias (meses de janeiro, fevereiro, março, outubro, novembro e dezembro) da probabilidade de precipitação pluvial em Piracicaba-SP, pelos resultados obtidos com número relativo de dias do mês superior a 70%, que: (i) a distribuição Gama pode ser utilizada em todos os referidos meses; (ii) a Weibull em fevereiro, março, novembro e dezembro; (iii) a Exponencial em março e novembro; (iv) a Log-Normal em março e (v) não se recomenda a distribuição Normal no período estudado.

Conclusões

Conforme os resultados obtidos, conclui-se que, para Piracicaba-SP, recomenda-se a distribuição Gama para as estimativas diárias de probabilidade de precipitação pluvial (meses de janeiro, fevereiro, março, outubro, novembro e dezembro).

Referências Bibliográficas

ALMEIDA, R. M. B. **Características climatológicas do regime de chuva em Minas Gerais**, Viçosa: UFV, 1995, 64p. Dissertação (Mestrado em Engenharia Agrícola) - Curso de Pós-graduação em Engenharia Agrícola, Universidade Federal de Viçosa.

ASSIS, F. N. de; ARRUDA, H. V.; PAREIRA, A. R. Aplicações de estatística a climatologia: Teoria e prática, Pelotas: Ed. Universitária/UFPel, 1996. 161p.

BOTELHO, V. A.; MORAIS, A. R. de. Estimativas dos parâmetros da distribuição Gama de dados pluviométricos do município de Lavras, Estado de Minas Gerais. **Ciências e Agrotecnologia**, Lavras, v.23, n.3, p.697-706, julho/ setembro, 1999.

BOTELHO, V. A.; MORAIS, A. R. Estimativas dos parâmetros da distribuição gama de dados pluviométricos do município de Lavras, Estado de Minas Gerais. **Ciências e Agrotecnologia**, v.23, n.3, p.697-706, 1999.

CAMPOS, H. **Estatística não paramétrica**, 4ª ed. Piracicaba, ESALQ/USP, 349p. 1983.

CATALUNHA, M. J.; SEDIYAMA, G. C.; LEAL, B. G.; SOARES, C. P. B.; RIBEIRO, A. Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, v.10, n.1, p.153-162, 2002.

COUTO, H. T. Z. **Distribuições de diâmetro em plantações de *Pinus caribaea* Morelet**, Piracicaba: Escola Superior de Agricultura “Luiz de Queiroz”/ USP, 79p. Tese (Livre Docência), ESALQ, 1980.

FARIA, R. A. **Demanda de irrigação suplementar no Estado de Minas Gerais**, Viçosa, UFV, 1998, 75p. Dissertação (Mestrado em Meteorologia Agrícola) - Curso de Pós-graduação em Meteorologia Agrícola, Universidade Federal de Viçosa, 1998.

Rev. Bras. Agrometeorologia, v. 13, n.2, p.273-283, 2005

FONSECA, V.O.; ALBUQUERQUE, J.A.S. Estimativa dos parâmetros da distribuição gama de probabilidades para totais de precipitação em uma região de Pelotas, RS. **Pesquisa Agropecuária Brasileira**, v.13, n.3, p.47-58, 1978.

FONTANA, D. C.; ALMEIDA, T. S. de. Climatologia do número de dias com precipitação pluvial no Estado do Rio Grande do Sul. **Revista Brasileira de Agrometeorologia**, v.10, n.1, p.135-145, Santa Maria, 2002.

HANN, C. T. **Statistical methods in hydrology**, Ames: Iowa State University Press, 1977. 378p.

HASTINGS, N. A. J.; PEACOCK, J. B. **Statistical distributions: a handbook for students and practitioners**, Longon Butterworths, England, 1975. 129p.

JOHNSON, N. L.; KOTZ, S. **Distribution in statistics, continue univariate distribution**, New York: Hougton Mifflin, v.2, 1970. 328p.

KITE, G. W. Frequency and risk analysis in hydrology, **Water Resources Publications**, Fort Collins, v.3, 1978. 395p.

MEYER, P. L. **Probabilidade**: **aplicações à estatística**. Rio de Janeiro; Livro Técnico, 1969, 391p.

MILLER, M. E.; WEAVER, C. R. Monthly and annual precipitation probabilities for climatic divisions in Ohio. **Research bulletin**, n.1005, 1968. 11p.

MIRSHAWKA, V. **Estatística**, v.2, São Paulo: Nobel, 1971. 367p.

MORETTIN, P. A.; BUSSAB, W. O. **Estatística básica**. 5ª ed., São Paulo, Saraiva, 2003, 526p.

RIBEIRO, A. M. de A; LUNARDI, D.M.C. A precipitação mensal provável para Londrina – PR, através da função gama. **Revista Energia na Agricultura**, v.12, n.4, p.37-44, 1997.

SAS Institute. **QC SOFTWARE**: Usage and reference. In: Capability procedure. Cary: STATISTICAL ANALYSES SYSTEM INSTITUTE: release 6. 1 ed., v.1, 1996. 823p.

THOM, H. C. S. A note on the gamma distribution. Monthly Weather Review, Washington, v.86, p.117-122. 1958.

THOM, H. C. S. **Some methods of climatological analysis**. Roma, FAO, 1966. 50p. (FAO, Technical Notes. 81).

VILLA NOVA, N. A. **Dados meteorológicos do município de Piracicaba**. Piracicaba, ESALQ: Departamento de Ciência Exatas, 2003, 2p.