

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE PESQUISAS HIDRAULICAS

I. P. H.

CURSO DE PÓS-GRADUAÇÃO EM  
RECURSOS HÍDRICOS E SANEAMENTO

MOMENTOS L NA REGIONALIZAÇÃO DE VALORES  
CARACTERÍSTICOS DE VAZÃO

Luis Edgar Montenegro Terrazas

Dissertação de Mestrado

ORIENTADOR: Prof. Robin T. Clarke

DEZEMBRO / 1989

Porto Alegre - Rio Grande do Sul

BRASIL

## AGRADECIMENTOS

Expresso meus agradecimentos:

Ao Professor Robin T. Clarke, pela orientação e incentivo durante a elaboração desta dissertação.

Ao Professor Carlos E. M. Tucci, pelo apoio proporcionado durante meus estudos de mestrado.

A Coordenadoria de Aperfeiçoamento de Pessoal de nível Superior (CAPES) do Ministerio de Cultura e Educação do Brasil, pela bolsa de estudo concedida.

A Auxiliadora dos Santos, pela paciente colaboração na elaboração do texto em português.

A meus pais Rosendo e Leonor e irmãs, Mirtha, Melvy e Gilva, pelo espírito solidário com que me acompanharam no transcorrer deste período.

A todas as pessoas que pelas mais diversas formas proporcionaram as condições para a realização deste trabalho.

# INDICE

	Pág.
<b>1.- <u>INTRODUÇÃO</u></b>	
1.1. Regionalização, uma necessidade em hidrologia -----	1
1.2. Objetivos do estudo realizado -----	2
<b>2.- <u>MOMENTOS PONDERADOS POR PROBABILIDADE E MOMENTOS L</u></b>	
2.1. Introdução -----	3
2.2. Definição dos MPP, dos momentos-L e estimativa amostral	
2.2.1. Momentos ponderados por probabilidade (MPP) -----	4
2.2.2. Momentos L -----	6
2.2.3. Estimativa dos MPP e momentos L -----	7
2.3. Caracterização de uma distribuição de probabilidade	
2.3.1. Momentos-L populacionais -----	8
2.3.2. Estimativa dos parâmetros da distribuição -----	9
<b>3.- <u>MOMENTOS L NOS METODOS DE REGIONALIZAÇÃO</u></b>	
3.1. Introdução -----	12
3.2. MPP no método índice Cheia -----	12
3.2.1. Regiões homogêneas -----	13
3.2.2. Escolha da distribuição -----	17
3.2.3. Independência -----	19
3.2.4. Avaliação do método índice Cheia com MPP -----	19
<b>4.- <u>MOMENTOS L NA REGIONALIZAÇÃO DA VAZÃO MÁXIMA E MÍNIMA NA BACIA DO RIO URUGUAI</u></b>	
4.1. Introdução -----	21
4.2. Descrição da região de aplicação -----	21
4.3. Dados disponíveis -----	24
4.4. Regionalização da vazão máxima	
4.4.1. Descrição do procedimento adotado -----	24
4.4.2. Distribuições de probabilidade utilizadas -----	28
4.4.3. Resultados	
a) Adoção da distribuição Gumbel	
Transferência da informação -----	33
Estimativa da vazão máxima $Y_r$ e intervalo de confiança	34

Análise da eficiência, por amostragem aleatória simples	37
Regressão ponderada -----	38
b) Adoção da distribuição Gamma	
Estimativa das curvas de frequência individuais -----	41
Transferência da informação -----	42
Estimativa da vazão máxima $\hat{Y}_T$ e intervalo de confiança	42
c) Adoção da distribuição GEV	
Estimativa das curvas de frequência individuais -----	44
Transferencia da informação -----	44
Estimativa da vazão máxima $\hat{Y}_T$ e intervalo de confiança	45
4.5. Regionalização da vazão mínima	
4.5.1. Distribuições de probabilidade adotadas -----	45
4.5.2. Caracterização da descarga mínima -----	46
4.5.3. Método de regionalização adotado -----	46
4.5.4. Resultados -----	47
5.- <u>CONCLUSÕES</u>	
5.1. Estimativa regional da vazão máxima com T anos de período de retorno -----	48
5.2. Estimativa regional da vazão mínima com T anos de período de retorno -----	50
6.- <u>BIBLIOGRAFIA</u> -----	52
<u>ANEXOS</u>	
I . Tab. de interpret. dos mom.-L e comparação com est. Máx. Veross.---	57
II . Informação sobre os postos usados na regionalização -----	58
III. Tabelas de estimativas pelo uso de momentos-L e momentos.-----	61
IV . Regressão ponderada -----	72
V . Variância de $\hat{Y}_T$ , caso da distribuição Gamma com momentos-L -----	74
VI . Características desejáveis de um "bom" estimador -----	76

## 1. INTRODUÇÃO

### 1.1 - Regionalização, uma necessidade em hidrologia

O planejamento e projeto de obras hidráulicas são atividades que no âmbito dos recursos hídricos utilizam valores representativos da resposta do sistema hidrológico para compatibilizá-lo com a atividade que se quer desenvolver.

A vazão máxima e a vazão mínima associadas, através de abordagens estatísticas, a uma frequência de ocorrência ou período de retorno  $T$  são valores tipicamente utilizados, por exemplo, para dimensionamento de vertedor e classificação de rios por seu potencial de diluição de resíduos respectivamente.

Uma situação muito comum na estimativa destes valores ( $\hat{Y}_T$ ) é a inexistência de registros fluviométricos no local de interesse. No entanto, se estes registros estiverem disponíveis em outros locais, técnicas de regionalização hidrológica podem ser usadas a fim de transferir a informação requerida dentro de uma região hidrologicamente homogênea.

A transferência da informação fluviométrica é possível através do uso de equações de regressão ajustadas entre algum índice (média, desvio padrão, coeficiente de assimetria ou quantil  $Y_T$  obtidos a partir das séries de vazão máxima ou mínima existentes), e variáveis fisiográficas e climáticas das bacias. Dependendo do tipo de método de regionalização utilizado, a estimativa destes índices no local sem posto fluviométrico nos conduzirá a estimativa do evento com  $T$  anos de período de retorno procurada.

Detalhes dos tipos de métodos de regionalização para estimativa de vazões extremas associadas a uma frequência de ocorrência foram descritos por Kite (1978); como exemplo podemos citar o "método índice-cheia" proposto inicialmente por Darlymple (1960) e de uso muito difundido.

Em termos gerais ressaltamos que é comum neste tipo de métodos utilizar estatísticas como a média, desvio padrão e coeficiente de assimetria na

tentativa de descrever as características de um conjunto de dados observados, como as séries anuais de vazão extrema. É também comum, ao se ajustar uma distribuição de probabilidades paramétrica a este conjunto de dados, estimar os parâmetros pela solução de equações dos momentos da amostra com aqueles da distribuição ajustada.

Hosking (1986), a partir de momentos ponderados por probabilidade, definiu quantidades denominadas momentos L, mostrando que são um meio alternativo de caracterizar um conjunto de dados observados e que também podem ser usados para definir uma distribuição de probabilidades a partir de uma amostra aleatória de dados. O autor ainda mostrou que nestas aplicações os momentos L podem ser usados com vantagens em relação ao uso de momentos (comuns). Torna-se interessante portanto, inserir os momentos L em algum método de regionalização e quantificar o valor do seu uso na estimativa da vazão máxima ou mínima com T anos de período de retorno em um local sem posto fluviométrico.

## 1.2 - Objetivos do estudo

Os trabalhos desenvolvidos neste estudo foram direcionados aos seguintes objetivos:

a) Adotar uma metodologia de regionalização e investigar se o uso dos momentos-L propostos por Hosking proporcionam vantagens em relação aos procedimentos tradicionais.

(b) Analisar aspectos fundamentais dos métodos mais utilizados na regionalização da vazão máxima e mínima anual, função da frequência.

(c) Apresentar o estágio de desenvolvimento atual nos aspectos mais salientes dos métodos de regionalização considerados.

## 2. MOMENTOS PONDERADOS POR PROBABILIDADE E MOMENTOS L

### 2.1 - Introdução

A aplicação hidrológica de momentos ponderados por probabilidade (denominados MPP e definidos a seguir), teve início com Greenwood et al. (1979), principalmente como um auxílio na estimativa dos parâmetros da distribuição Wakeby que Houghton (1978) introduzira na análise de frequência de cheias, observando que esta distribuição pode explicar a condição de "separação". (A condição de separação foi notada por Matalas et al. (1975) como sendo a característica que distingue as séries de cheia regional de séries obtidas com experimentos de Monte Carlo de distribuições comumente usadas em hidrologia).

Um exemplo do uso da distribuição de Wakeby é encontrado nos trabalhos de Damazio et al. (1983) que definiram cenários, considerados representativos das vazões máximas de rios brasileiros, para os que através da distribuição Wakeby geraram registros sintéticos aos que ajustaram diversas distribuições a fim de escolher um estimador robusto para a cheia de projeto de vertedores de grandes barragens. Isto é, procuraram uma distribuição e um método de estimativa que fossem capazes de estimar eventos extremos, qualquer que seja a população do mundo real, sem perdas de eficiência desastrosas.

A distribuição de Wakeby na sua forma inversa tem a seguinte representação matemática:

$$Y_T = m + a [1 - (1 - F)^b] - c [1 - (1 - F)^{-d}]$$

onde  $Y_T$  é o quantil associado com dada probabilidade de não excedência ( $F = F(y) = P[Y \leq y]$ ) da variável aleatória  $Y$ . Os parâmetros  $m$ ,  $a$ ,  $b$ ,  $c$ ,  $d$  seriam difíceis de ser estimados pelo método de momentos (convencionais) já que embora os momentos podem ser representados função dos parâmetros, é difícil obter relações inversas. A estimativa dos parâmetros pelo método de máxima verossimilhança também não seria simples; no entanto os parâmetros da distribuição de Wakeby podem ser estimados através do método de momentos ponderados por probabilidade pela solução de equações muito simples.

Hosking (1986, 1989, 1990) aumentou a utilidade dos MPP ao derivar destes, um conjunto de quantidades denominadas momentos L que possibilitam:

- a) O resumo de uma amostra de dados através de valores, que como os momentos (convencionais) de uma amostra, são interpretados como medidas de tendência central, dispersão, assimetria e kurtosis; denotados respectivamente por  $l_1$ ,  $l_2$ ,  $l_3/l_2$  e  $l_4/l_2$ .
- b) Descrever o comportamento da variável aleatória Y associada a uma distribuição de probabilidade acumulada ou a uma função de densidade. São denotados como  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3/\lambda_2$  e  $\lambda_4/\lambda_2$ , sendo respectivamente interpretados como medidas de tendência central, dispersão, assimetria e kurtosis da distribuição de probabilidade.
- c) Ajustar certa distribuição de probabilidade paramétrica resolvendo as equações obtidas pela consideração de que os momentos L da amostra são boas estimativas dos momentos L populacionais. (No Anexo VI são descritos alguns critérios para julgar um estimador e conseqüentemente a estimativa)

Neste capítulo, partindo da definição dos MPP, são apresentados detalhes de cada um dos aspectos acima mencionados. São também consideradas as potencialidades do uso de momentos L frente a práticas mais usuais da estatística em hidrologia.

## 2.2 - Definição dos MPP, dos momentos-L e estimativa amostral

### 2.2.1. Momentos ponderados por probabilidade (MPP)

Os momentos ponderados por probabilidade de uma variável aleatória Y; tal como descarga máxima anual com função de distribuição  $F(y)=P(Y \leq y)$ , foram definidos por Greenwood et al. (1979) como sendo as quantidades:

$$\alpha_{p,r,s} = E[Y^p (F(Y))^r (1-F(Y))^s] \quad (2.1)$$

onde  $p,r,s$  são números inteiros e  $E[.]$  denota o valor esperado. As quantidades  $\alpha_{p,0,0}$  ( $p = 1, 2, \dots$ ) são os momentos em relação a origem comumente usados. Os momentos  $\alpha_{p,r,s}$  poderiam no entanto, ser preferíveis para estimativa dos parâmetros da distribuição de Y, já que o expoente de Y

na expressão para  $\alpha_{1,r}$ , e significa que as relações entre parâmetros e momentos frequentemente são de uma forma mais simples que para o caso dos momentos comuns.

Os momentos ponderados por probabilidade  $\alpha_{1,0}$  e  $\alpha_{1,r,0}$  que por facilidade de notação serão distinguidos como  $\alpha_r$  e  $\beta_r$  respectivamente são definidos a partir da equação (2.1) pelas seguintes relações:

$$\alpha_r = E[ Y(1 - F(Y))^r ] \quad r=0,1,2,\dots \quad (2.2)$$

$$\beta_r = E[ Y(F(Y))^r ] \quad r=0,1,2,\dots \quad (2.3)$$

Observamos que  $r\alpha_{r-1} = E[Y_{1:r}]$  e  $r\beta_{r-1} = E[Y_{r:r}]$  são os valores esperados de estatísticas de ordem extrema definidos como:

$$E[Y_{j:r}] = \frac{r!}{(j-1)!(r-j)!} \int Y (F(Y))^{j-1} (1-F(Y))^{r-j} dF(Y)$$

onde  $j$  é um número inteiro maior que zero.

Utilizando o teorema binomial para expandir a expressão dentro de (.) da equação (2.2), ou alternativamente, escrevendo  $F(y)=1-(1-F(y))$  na equação (2.3), e depois utilizando o teorema binomial, pode ser visto que  $\alpha_r$  e  $\beta_r$  são funções um do outro:

$$\alpha_r = \sum_{k=0}^r (-1)^k \binom{r}{k} \beta_k \quad \beta_r = \sum_{k=0}^r (-1)^k \binom{r}{k} \alpha_k \quad (2.4)$$

onde  $\binom{r}{k}$  são combinações de  $r$  elementos de  $k$  em  $k$ . Em particular mostra-se que:

$$\begin{aligned} \alpha_0 &= \beta_0 & \beta_0 &= \alpha_0 \\ \alpha_1 &= \beta_0 - \beta_1 & \beta_1 &= \alpha_0 - \alpha_1 \\ \alpha_2 &= \beta_0 - 2\beta_1 + \beta_2 & \beta_2 &= \alpha_0 - 2\alpha_1 + \alpha_2 \\ \alpha_3 &= \beta_0 - 3\beta_1 + 3\beta_2 - \beta_3 & \beta_3 &= \alpha_0 - 3\alpha_1 + 3\alpha_2 - \alpha_3 \end{aligned}$$

Um resultado muito importante é que qualquer um dos dois conjuntos de MPP ( $\alpha_r; r=0,1,\dots$ ) ou ( $\beta_r; r=0,1,\dots$ ) é suficiente para caracterizar uma

distribuição de probabilidade. É, portanto, natural utilizar os primeiros  $\alpha_r$  (ou  $\beta_r$ ) para resumir as características principais de uma distribuição, e utilizá-los também para estimar os parâmetros da distribuição e, portanto, para regionalizar informação hidrológica.

### 2.2.2. Momentos L

Os MPP caracterizam uma distribuição, mas ao não serem facilmente interpretados, Hosking (1986) usou combinações lineares de  $\alpha_r$  e  $\beta_r$  da forma:

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^r (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \alpha_k = \sum_{k=0}^r (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \beta_k \quad r = 0, 1, 2, \dots \quad (2.5)$$

mostrando que  $\lambda_1$ ,  $\lambda_2$  e  $\lambda_3$  são suficientes para proporcionar um resumo descritivo simples de uma distribuição de probabilidade em termos de medidas de localização, escala e forma. As quantidades  $\lambda_{r+1}$  são os momentos L; sendo escritos de modo mais conveniente como:

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E[Y_{r-k:r}] \quad r = 1, 2, 3, \dots \quad (2.6)$$

onde  $Y_{k:n}$  é a estatística de k-ésima ordem de uma amostra aleatória de tamanho n, tirada da população, com função de distribuição  $F(y)$ . A equação (2.6) enfatiza que os momentos L são funções lineares do valor esperado de estatísticas de ordem; de fato os primeiros momentos L são:

$$\begin{aligned} \lambda_1 &= E[Y] \\ \lambda_2 &= 1/2 E[ Y_{2:2} - Y_{1:2} ] \\ \lambda_3 &= 1/3 E[ Y_{3:3} - 2Y_{2:3} + Y_{1:3} ] \\ \lambda_4 &= 1/4 E[ Y_{4:4} - 3Y_{3:4} + 3Y_{2:4} - Y_{1:4} ] \end{aligned} \quad (2.7)$$

ou

$$\begin{aligned}
\lambda_1 &= \alpha_0 & &= \beta_0 \\
\lambda_2 &= \alpha_0 - 2\alpha_1 & &= 2\beta_1 - \beta_0 \\
\lambda_3 &= \alpha_0 - 6\alpha_1 + 6\alpha_2 & &= 6\beta_2 - 6\beta_1 + \beta_0 \\
\lambda_4 &= \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 & &= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0
\end{aligned}
\tag{2.8}$$

O valor  $\lambda_1$ , que é igual à média da distribuição, é uma medida de localização;  $\lambda_2$  é uma medida de dispersão da distribuição, e pode ser considerada uma medida de escala; as relações de momento  $L$   $\tau_3 = \lambda_3 / \lambda_2$  e  $\tau_4 = \lambda_4 / \lambda_2$  são medidas de assimetria e kurtosis da distribuição.

### 2.2.3. Estimativa dos MPP e momentos L

Considerando que dispomos de uma amostra aleatória ordenada  $y_1 \leq y_2 \leq y_3 \leq \dots \leq y_n$  (com número de dados,  $n$ , maior que a ordem,  $r$ , do MPP desejado), obtida da distribuição; o  $y_i$  poderia ser por exemplo, um conjunto ordenado de descargas máximas anuais. A estimativa de  $\alpha_r$  e  $\beta_r$  pode ser obtida pelas estatísticas:

$$a_r = n^{-1} \sum_{i=1}^n \binom{n-i}{r} y_i / \binom{n-1}{r} \quad r=0,1,2,\dots, n-1 \tag{2.9}$$

$$b_r = n^{-1} \sum_{i=1}^n \binom{i-1}{r} y_i / \binom{n-1}{r} \quad r=0,1,2,\dots, n-1 \tag{2.10}$$

onde, por convenção,  $\binom{k}{j} = 0$  se  $k < j$ . Landwehr et al. (1979) mostra que  $a_r$  é um estimador não-tendencioso de  $\alpha_r$  e  $b_r$  é um estimador não tendencioso de  $\beta_r$ . Substituindo  $a_r$  e  $b_r$  em lugar de  $\alpha_r$  e  $\beta_r$  nas equações (2.8) obtemos para  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ , e  $\lambda_4$  estimadores não tendenciosos  $l_1$ ,  $l_2$ ,  $l_3$  e  $l_4$ . Estimadores naturais das razões de momentos L são então  $l_3/l_2$  e  $l_4/l_2$ ; contudo, embora seja verdade que  $l_1$ ,  $l_2$ ,  $l_3$  e  $l_4$  são estimadores não-tendenciosos de  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  e  $\lambda_4$ , não é verdade que  $l_3/l_2$  e  $l_4/l_2$  sejam estimadores não tendenciosos de  $\lambda_3/\lambda_2$  e  $\lambda_4/\lambda_2$ , embora sejam estimadores consistentes dessas quantidades.

Ressaltamos dos resultados de Hosking (1986) que os momentos L obtidos a partir de uma amostra (por exemplo uma série de vazão máxima anual) podem ser usados de maneira similar aos momentos amostrais (convencionais) assim:

a) Resumem as propriedades básicas de um conjunto de dados. A tabela 1.1 do Anexo I, ilustra a correspondência de interpretação com relação aos momentos amostrais (convencionais).

b) Estimam as propriedades de uma distribuição de probabilidade com base nos dados obtidos por amostragem.

c) Podem ser usados para estimar os parâmetros da distribuição escolhida.

nestas aplicações os momentos L seriam frequentemente preferidos aos momentos convencionais já que sendo funções lineares dos dados, eles são menos sensíveis que os momentos convencionais aos erros de amostragem (este erro pode ser atribuído ao escasso número de registros de vazão) ou a erros de medição nos valores de dados extremos; por isto, se esperaria também que as estimativas das características ou parâmetros da distribuição de probabilidade ajustada sejam mais precisas e robustas.

## 2.3 - Caraterização de uma distribuição de probabilidade

### 2.3.1. Momentos L populacionais

Usando a definição do valor esperado de estatísticas de ordem nas equações (2.7) se obtém:

$$\begin{aligned}\lambda_1 &= \int_0^1 Y(F) dF \\ \lambda_2 &= \int_0^1 Y(F) (2F-1) dF \\ \lambda_3 &= \int_0^1 Y(F) (6F^2-6F+1) dF \\ \lambda_4 &= \int_0^1 Y(F) (20F^3-30F^2+12F-1) dF\end{aligned}\tag{2.11}$$

onde  $Y(F)$  é a função inversa da distribuição  $F=F(y)$ . A solução das integrais nos permitirá obter as expressões para os momentos L populacionais, dada uma determinada função de distribuição atribuída a variável aleatória  $Y$ .

Hosking (1990), apresenta os momentos L de várias distribuições de uso corrente em hidrologia. No capítulo 3 são apresentadas as expressões dos

momentos L para as distribuições usadas neste trabalho, sendo estas as distribuições Gumbel, Gamma e Valores Extremos Generalizada (GEV). Ressaltamos que embora as integrais acima são apresentadas para o caso em que a distribuição possa ser expressa na sua forma inversa, soluções também foram encontradas para o caso em que a distribuição não possa ser expressa na sua forma inversa. Um caso de interesse particular é encontrado nos trabalhos de Song e Ding (1988), Ding et al. (1989 a, b) para a distribuição Pearson III.

### 2.3.2. Estimativa dos parâmetros da distribuição

Analogamente ao método usual de momentos, a estimativa dos  $p$  parâmetros de uma distribuição de probabilidades pode ser realizada através do "método de momentos L". Para o efeito os primeiros  $p$  momentos L amostrais (ou equivalentemente os  $p$  primeiros MPP) são igualados as correspondentes quantidades populacionais.

Se por exemplo for considerada apropriada a distribuição Gumbel para representar a descarga máxima anual de uma bacia determinada, faremos  $l_1 = \lambda_1$  e  $l_2 = \lambda_2$  então

$$\begin{aligned} l_1 &= \hat{\theta} + \gamma \hat{\phi} \\ l_2 &= \hat{\phi} \ln 2 \end{aligned}$$

onde  $\gamma$  é a constante de Euler e igual a 0.5772... . Deste sistema de equações pode se obter ainda

$$\begin{aligned} \hat{\phi} &= l_2 / \ln 2 \\ \hat{\theta} &= l_1 - \gamma \hat{\phi} \end{aligned}$$

que são as expressões de estimativa pelo método de momentos L dos parâmetros, respectivamente de escala e localização, da distribuição de Gumbel. Expressões deste tipo são dadas por Hosking (1990) para o caso em que as distribuições desejadas são: Gamma, Valor Extremo Generalizado, Wakeby e outras que são menos usadas para a representação de variáveis hidrológicas. Ding et al. (1989 a, b), proporcionam relações dos parâmetros da distribuição Person Tipo III com os MPP.

Resultados encontrados na literatura nos permitem comparar alguns aspectos dos métodos de estimativa dos parâmetros de uma distribuição de

probabilidades quando se usa momentos-L com os métodos mais usuais de momentos e o método assintoticamente ótimo de máxima verossimilhança.

Embora quando modelamos as propriedades dos valores extremos de processos físicos, raramente ocorre que as séries de dados disponíveis sejam suficientemente longas, é interessante analisar as propriedades assintóticas de uma técnica estatística nova por dois motivos fundamentais. Primeiro, nos permite verificar a integridade da técnica no sentido em que, quando uma grande amostra estiver disponível, o método novo não será totalmente ineficiente comparado com um já estabelecido como o método assintoticamente ótimo de máxima verossimilhança. Segundo, a teoria assintótica (isto é, supondo que o comportamento das estimativas segue o Teorema do Limite Central) pode também proporcionar uma aproximação adequada para algum aspecto da distribuição da estatística para amostras de tamanho médio. Por exemplo Hosking et al. (1985 b) verificou que a variância dos estimadores dos parâmetros e dos quantiles via MPP da distribuição de Valores Extremos Generalizada é bem aproximada pela teoria assintótica para amostras com 50 elementos ou mais.

Em termos gerais observa-se que as estimativas dos parâmetros via momentos-L são razoavelmente eficientes; no sentido em que as variâncias das estimativas em que se usam momentos-L (var MOL), são somente um pouco maiores do que as variâncias quando o método de máxima verossimilhança (var MV) é usado. Por exemplo, Hosking (1990) assinala que as eficiências assintóticas ( $100 \times (\text{var MV} / \text{var MOL})$ ) dos estimadores por momentos-L da localização e escala para a distribuição Normal são 100% e 97.8% respectivamente. A situação correspondente a distribuição Gumbel seria de 99.6% e 75.6%. Com relação a distribuições de três parâmetros Hosking et al. (1985 b), afirma que as eficiências assintóticas dos estimadores via momentos-L tendem a ser menores que as eficiências de distribuições de dois parâmetros, mas que ainda são razoavelmente altas. Por exemplo, os parâmetros da distribuição GEV são todos estimados com pelo menos 70% de eficiência quando o parâmetro de forma K satisfaz  $-0.2 \leq K \leq 0.2$ .

Tiago de Oliveira (1963) mostra que as estimativas dos parâmetros de localização e escala para a distribuição Gumbel obtidas pelo método de

momentos tem uma eficiência assintótica de 95% e 55% respectivamente, sendo portanto menos eficientes que os estimadores com momentos-L.

Para pequenas amostras de dados como as que em geral são encontradas quando tratamos com registros de vazão extrema anual os resultados acima expostos podem ter pouco significado. Landwehr et al. (1979), estudaram através de experimentos de Monte Carlo o desempenho do método de momentos-L para estimativa dos parâmetros da distribuição Gumbel. A análise dos seus resultados para amostras com menos de 50 elementos, colocando especial atenção nos extremos da distribuição ( $T=50, 100$  e  $1000$  anos) nos permite indicar que o método de momentos L apresenta desempenho equivalente para a medida de eficiência com relação ao método de momentos convencionais. A comparação correspondente quanto aos valores de tendenciosidade favorecem um pouco o método de momentos-L. Hosking et al. (1985 b) comparou a estimativa de máxima verossimilhança e o método de momentos-L para a distribuição GEV mostrando que os desvios padrões das estimativas de parâmetros e quantiles via momentos-L são menores que os desvios correspondentes ao uso do método de máxima verossimilhança. Alguns dos resultados de Hosking para  $K=0.2$  são reproduzidos na tabela I.2. Nesta pode observar-se a preferência pelo método dos momentos-L principalmente para períodos de retorno elevados e para amostras pequenas.

Quanto a aplicação prática dos métodos de momentos e momentos-L observa-se que os momentos-L facilitam o uso de algumas distribuições como a GEV, Wakeby e outras.

Hosking et al. (1985 b) também assinala que como os momentos L da amostra são menos tendenciosos que os momentos (comuns) amostrais, ao equacionar momentos amostrais com momentos populacionais, inferências mais seguras poderão se obter com o método dos momentos L.

### 3.- MOMENTOS L NOS MÉTODOS DE REGIONALIZAÇÃO

#### 3.1. Introdução

Este capítulo considera aspectos dos métodos de regionalização cujos propósitos específicos sejam o de proporcionar estimativas das distribuições de frequência da vazão máxima e mínima em locais sem registros fluviométricos

O método de regionalização Índice Cheia com ajuste da distribuição de frequência das cheias via MPP, foi um método bastante estudado nos últimos anos, pelo que será especialmente analisado.

Riggs (1973), assinala que o grau de sucesso de determinado método de regionalização depende: (a) da representatividade das curvas de frequência individuais que são a base da relação regional, (b) do tamanho relativo; nas curvas de frequência individual, da variação casual devido as ocorrências aleatórias do clima e a variação devida a diferenças nas características da bacia, (c) do grau de independência das amostras nos diversos postos de medição, (d) da qualidade da relação com as características da bacia. Portanto, é necessário estudar como o uso de momentos L em um método de regionalização influencia nos aspectos acima mencionados.

#### 3.2 -MPP no Método Índice Cheia

Dairymple (1960), apresentou o método Índice Cheia que se caracteriza por considerar que uma dada região teria uma única distribuição de probabilidades para as vazões máximas anuais divididas por um fator de escala, usualmente a média. Combinando esta curva adimensional com a estimativa do fator de adimensionalização obtido por regressão para um local sem registros de vazão, pode-se estimar neste local a vazão máxima anual com determinado período de retorno.

Posteriormente NERC(1975), ELETROBRAS-IPH (1983), Lanna et al. (1983), são alguns dos autores que usaram a orientação do método Índice Cheia na

elaboração de trabalhos a fim de proporcionar estimativas dos valores de vazão máxima anual ou vazão mínima anual com duração  $D$ , associada a determinado período de retorno; em locais onde os registros de vazão são inadequados ou inexistentes.

Wallis (1980), Greis e Wood (1981), sugeriram uma metodologia do tipo índice cheia em que se usa os momentos ponderados por probabilidade, apresentando o seguinte algoritmo:

1. Ordenar os dados de vazão máxima anual no vetor  $(y_1, y_2, \dots, y_n)$ ,  $y_1 \leq y_2 \leq \dots \leq y_n$
2. Calcular os momentos ponderados por probabilidade ( $\alpha_r, r=0, 1, \dots$ ) para cada um dos postos na região.
3. Padronizar os  $\alpha_r$  de cada posto fazendo  $\alpha_r = \alpha_r / \alpha_0$ . Notar que se os dados já foram divididos pela média esta etapa é dispensável.
4. Obter os valores regionais de  $\alpha_r$  calculando a média com os  $\alpha_r$ 's de cada posto.
5. Estimar os parâmetros  $\Theta_r$  da distribuição regional. No caso da distribuição GEV  $\Theta_r = (\Theta, \phi, K)$ , onde  $\Theta, \phi$  e  $K$  são os parâmetros de localização, escala e forma respectivamente.
6. Estimar quantiles regionais  $q_r$  ( $T=10, 50, 100, 200, \dots$ )
7. Obter as estimativas da vazão máxima  $Q_r = \bar{Q} \times q_r$  para um local sem registros. Onde  $\bar{Q}$  é a estimativa da média da vazão máxima anual obtida por regressão regional para a bacia sem registros de vazão.

Diversos aspectos do uso de algoritmos deste tipo para a regionalização da vazão máxima anual são considerados a seguir.

### 3.2.1. Regiões homogêneas

Se conseguíssemos identificar um conjunto de bacias cuja descarga de cheia anual tem distribuição de frequência idêntica, exceto pelo fator de escala e diferenças provenientes da aleatoriedade do fenômeno que está sendo representado, estaríamos identificando uma região homogênea no sentido de atender a suposição geral dos métodos do tipo índice- cheia.

O comportamento das curvas de distribuição de frequência da vazão em um rio é, no entanto, o resultado da combinação de um grande número de fatores físicos e climáticos, sendo reconhecida a dificuldade de relacionar estatísticas da vazão com estes fatores. Ainda seria necessário verificar se as diferenças na declividade das curvas de frequência moduladas individuais, são maiores que as que se possa atribuir somente ao acaso, uma verificação análoga poderia ser referida aos parâmetros de ordem maior que um das distribuições.

Na prática, tem surgido vários critérios que podem ser usados nos métodos do tipo índice Cheia e que perseguem o estabelecimento de regiões homogêneas.

Dalrymple (1960), depois de estabelecer regiões supostamente homogêneas recomendou um teste que compara a variabilidade das estimativas de cheia com 10 anos de período de retorno para cada posto na região, com a variabilidade esperada se somente o erro de amostragem fosse o responsável pelas diferenças entre as estimativas no posto.

Além de Dalrymple (1960), Cruff e Rantz (1965), Chong e Moore (1983) são alguns dos autores que usaram este teste, concluindo em cada caso que as regiões estudadas eram homogêneas. Wiltshire (1986 a) assinala que estas conclusões somente indicam que o teste não é sensível ou que uma grande variedade de séries de cheia e tipos de bacia seriam consistentes com o conceito de homogeneidade.

Uma das deficiências do teste está em ser aplicado somente no período de retorno  $T=10$  anos. Benson (1962), observou que na prática o teste não poderia ser aplicado para tempos de retorno maiores, por que vários registros individuais são muito curtos, não definindo adequadamente a curva de frequência nestes níveis. Também ressalta que em alguns estudos, ainda que a homogeneidade seja aparentemente estabelecida ao nível de 10 anos, as curvas individuais mostram amplas diferenças em níveis elevados.

NERC (1975) usou um critério de similaridade de características geográficas para dividir o Reino Unido em 11 regiões geográficas realizando

nestas a análise de frequência de  $Q/\bar{Q}$ . Nenhum teste estatístico de uniformidade hidrológica das regiões foi utilizado. Os mesmos autores, subdividiram as regiões seguindo o critério de se agrupar as bacias de acordo ao seu tamanho, houve dificuldade em se definir as curvas de frequência pelo menor número de postos disponíveis em cada grupo, sendo impossível detectar qualquer efeito consistente da área das bacias na curva regional. Observaram ainda que para uma bacia em particular, a geologia e o clima podem influir mais na declividade da curva de frequência que a área simplesmente, não sendo recomendável se estabelecer de imediato que bacias pequenas tem curva de frequência com maior declividade que bacias com maior área.

Um outro critério no estabelecimento de regiões homogêneas é o empregado correntemente pelo U. S. Geological Survey, denominado método dos resíduos. ELETROBRAS-IPH (1983), o usaram como auxílio para delinear regiões homogêneas na bacia do Alto Paraguai. A técnica envolve o uso dos resíduos da equação de regressão que relaciona os valores médios da característica da vazão em estudo (média das vazões de cheia, média das vazões mínimas com dada duração, vazão média de longo período) de cada posto fluviométrico, com características físicas e climáticas da bacia. Se assume que as tendências gerais destes resíduos refletem variações inerentes a resposta de cheia das regiões.

Definidos os postos que constituirão as novas regiões, deve-se estabelecer as regressões e as curvas de frequência para estas regiões. A condição básica é de que o coeficiente de determinação deve aumentar e o erro padrão da estimativa da regressão e das curvas de frequência devem diminuir. Caso contrário a sub-divisão não melhorou as estimativas.

Choquette (1988), usou o método dos resíduos dividindo o Estado de Kentucky nos Estados Unidos em sete regiões de cheia. Neste trabalho a variável dependente da equação de regressão foi a vazão de cheia com 50 anos de período de retorno ( $Q_{50}$ ), estimada depois de se ajustar as séries de vazão máxima em cada posto à distribuição log-Pearson Tipo-III. Os resíduos são o resultado de  $Q_{50} - \hat{Q}_{50}$ , onde  $\hat{Q}_{50}$  é o valor estimado pela equação de regressão.

Wiltshire (1986 a), assinala que a maioria dos estudos que usam a abordagem dos resíduos delimitam regiões que frequentemente coincidem com

limites geográficos e/ou hidrológicos; e que estas regiões provavelmente contém bacias com geomorfologia diversa, cujas características de frequência de cheia não são similares. Com a intenção de resolver esta deficiência Wiltshire (1986 b) definiu grupos de bacias de acordo com a análise de variáveis que caracterizem sua resposta em termos de cheias. Para tal fim foram usadas técnicas de agrupamento ("cluster") sobre a cheia média anual específica (QSP) e o coeficiente de variação (CV) de rios do Reino Unido, conseguindo-se definir na região 10 grupos de bacias.

Observa-se que as variáveis de agrupamento usadas por Wiltshire são bastante significativas assim a QSP é uma descrição espacial da média da cheia máxima anual na bacia. A medida de variabilidade da cheia de ano a ano (CV) é muito útil ao ser relacionada com a declividade da curva de frequência para cada bacia.

Depois de definir os M grupos de bacias homogêneas em determinada região, a técnica usada para designar uma bacia sem registros a algum destes M grupos consiste em estabelecer, inicialmente, funções discriminante  $S_i$  ( $i=1,2,\dots,M$ ) cuja forma geral é:

$$S_i = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_A x_A$$

Rao (1973), mostra como calcular os parâmetros  $b_0, b_1, b_2, \dots, b_A$  para cada um dos grupos. As variáveis classificatórias neste caso são as características climato-fisiográficas da bacia  $x_1, x_2, \dots, x_A$ .

Definidas as M funções discriminantes, o "score" discriminante  $S_i$  ( $i=1, \dots, M$ ) é determinado para os dados fisiográficos e climáticos existentes na bacia (B) sem registros fluviométricos. O grupo cujo valor de  $S_i$  é o maior será o grupo ao qual pertence a bacia B; no entanto, qualquer grupo cujo "score" não é nulo reclama para si a bacia B com diversos graus de intensidade. Pode-se então estimar a probabilidade  $P_i$  da bacia sem registros fluviométricos B pertencer ao grupo i, como:

$$P_i = \frac{S_i}{\sum_{i=1}^M S_i}$$

A cheia com T anos de período de retorno  $Y_T$  poderá ser estimada para a bacia B pelo cálculo da média ponderada dos valores da cheia com T anos de período de retorno  $Y_{Ti}$ , estimados para cada grupo.

$$Y_T = \frac{\sum_{i=1}^M P_i Y_{Ti}}{\sum_{i=1}^M P_i} \quad \text{onde} \quad \sum_{i=1}^M P_i = 1$$

Observamos que os grupos para os que a probabilidade de conter a bacia B é zero, não contribuem na estimativa de  $Y_T$ , enquanto que outros contribuem em proporção ao valor de  $P_i$ . Esta aproximação também contorna a dificuldade reportada por Lanna (1983) de encontrar grupos cujas bacias tenham continuidade geográfica.

A fim de testar o critério de grupos homogêneos de bacias, Wiltshire o aplicou em uma bacia com características climato-fisiográficas supostas, mostrando que o comportamento da cheia  $Y_T$  é coerentemente associado a razões físicas.

Nageshwar et al. (1989), examinaram a média e mediana de variáveis hidrológicas importantes das bacias nos grupos e nas regiões delimitadas pelo método dos resíduos, verificando que estes parâmetros tem menor variabilidade nas bacias identificadas por técnicas de agrupamento.

### 3.2.2. Escolha da distribuição teórica

Embora a distribuição regional teórica escolhida por Greis e Wood (1981) tenha sido a Gumbel é evidente que outras funções de distribuição de frequência para a vazão máxima anual padronizada, caracterizadas através dos MPP, podem ser escolhidas.

Não existe, no entanto, uma teoria forte na qual se possa fundamentar a escolha da forma da distribuição de probabilidades para a vazão máxima ou mínima. Erros devido a incorreta especificação de uma distribuição de probabilidades podem ser aliviados pelo uso de um método robusto, ou seja, um método que proporcione estimativas regionais da cheia  $Q_T$  no local de interesse, com propriedades estatísticas de não tendenciosidade, e que exibam limites de confiança estreitos.

Lettenmaier et al. (1987), concluíram através de experimentos de Monte Carlo que as distribuições de dois parâmetros provenientes da família dos Valores Extremos, inseridas no método de regionalização tipo Índice Cheia com uso dos MPP, proporcionam bons resultados só quando a distribuição da população e a distribuição ajustada são semelhantes. Caso contrário, grande tendenciosidade resultará da incorreta especificação da distribuição (lembrando que a forma da distribuição da população não é conhecida na prática). Os mesmos autores verificaram que a distribuição de valores extremos generalizada (GEV), quando ajustada usando o método regional de MPP, tem se mostrado relativamente insensível às violações quanto a "verdadeira" distribuição de probabilidades da cheia máxima anual padronizada. Mostraram ainda que a distribuição GEV é relativamente insensível a uma modesta heterogeneidade regional quanto ao coeficiente de variação, e é bastante insensível à variação regional no coeficiente de assimetria.

Hosking et al. (1985 a), Hosking e Wallis (1988) indicam que a distribuição GEV junto ao método de regionalização com MPP pode ser a base de um método robusto de estimativa da vazão máxima.

A adoção da distribuição em estudos de frequência de cheias pode tornar-se um problema crítico, principalmente se os eventos que se quer estimar correspondem a períodos de retorno elevados. Matalas e Wallis (1973), Hosking et al. (1985) concluíram que para estes casos, as estimativas podem diferir desconcertantemente. Para tempos de retorno da ordem do tamanho dos registros, a discrepância entre os resultados práticos obtidos pela aplicação de diferentes distribuições, usualmente não é significativa.

### 3.2.3. Independência

Na análise de frequência regional, além de assumir que os dados de cheia máxima anual de diferentes anos são estatisticamente independentes, frequentemente também se assume que os dados de diferentes postos no mesmo ano são independentes.

Quando tratamos com valores de vazão máxima anual padronizada (comumente pela média, nos métodos do tipo Índice Cheia) a primeira suposição em geral é atendida. Aclaramos que a padronização pela média não introduz maior dependência entre os dados, no entanto outros efeitos indesejáveis produzidos por esta transformação nos registros, são apresentadas no próximo item.

Na prática é improvável que a suposição de independência entre os dados dos postos de uma região em um mesmo ano seja atendida. Hosking e Wallis (1988), estudaram os efeitos desta dependência para os quantiles de cheia  $Q_T$  estimados pelo método regional de MPP tipo Índice Cheia. Suas principais conclusões foram: Nenhuma tendenciosidade na estimativa de  $Q_T$  é modificada pela presença de dependência entre postos. A precisão de  $Q_T$  diminui quando a dependência entre postos está presente, no entanto este efeito na prática é menos importante que a tendenciosidade na estimativa de  $Q_T$  devido a heterogeneidade (distribuições de frequência diferentes na região).

### 3.2.4. Avaliação do método Índice Cheia com MPP

Algumas deficiências no método de regionalização Índice Cheia são reconhecidas na literatura. Benson (1962), assinala que os valores  $Q/\bar{Q}$  para rios similares podem apresentar resultados muito diferentes, se o índice de cheia  $\bar{Q}$  em um dos rios não for típico por ter sido obtido a partir de um curto período de registros. Portanto, a curva de frequência neste rio será defeituosa. Outras dificuldades citadas por Benson poderiam ser superadas, através de critérios novos de estabelecimento de regiões homogêneas, como as referenciadas no item 3.2.1.

Stedinger (1983), mostra que a tentativa de obter quantiles  $Q_r$  a partir da curva de distribuição dos valores  $y_i = Q_i/\bar{Q}$ , pode proporcionar resultados muito distorcidos, já que a curva dos  $y_i$  descreve deficientemente a curva adimensional de frequência real.

#### 4.- MOMENTOS-L NA REGIONALIZAÇÃO DA VAZÃO MÁXIMA E MÍNIMA NA BACIA DO RIO URUGUAI

##### 4.1. Introdução

Assumindo que os valores de vazão máxima anual e mínima anual com duração  $d$  em diferentes postos da bacia do rio Uruguai possam ser tratados como valores aleatórios, é possível a caracterização das séries de cada posto de medição por seus primeiros momentos  $L_1, L_2$ , e as razões de momentos  $L_3/L_2$  e  $L_4/L_2$ .

Lembrando que estes momentos  $L$  e suas razões de momentos  $L$ , também permitem definir a relação da magnitude da variável aleatória com sua frequência de ocorrência, é de muito interesse transferir esta informação a locais sem registros fluviométricos. Também é de interesse analisar se as estimativas regionalizadas de vazão máxima ou mínima com período de retorno  $T$ , obtidas para estes locais através do uso de momentos  $L$ , resultam em estimativas com maior precisão, comparadas com outros métodos.

Neste capítulo, são descritos alguns detalhes da regionalização da vazão máxima e vazão mínima efetuada em uma região da bacia do rio Uruguai. Mostra-se também os resultados das diversas situações analisadas, a fim de obter conclusões sobre a conveniência do uso de momentos  $L$  em regionalização.

##### 4.2. Descrição da região de aplicação.

A área em que se investiga o valor do uso de momentos- $L$  na regionalização de características da vazão, corresponde a uma parte da bacia hidrográfica do rio Uruguai, e encontra-se compreendida entre  $26^{\circ} 30'$  e  $28^{\circ} 30'$  de latitude sul e entre  $49^{\circ} 30'$  e  $53^{\circ} 01'$  de longitude oeste, conforme indicado na figura II.1. Este trecho drena uma área fronteiriça entre os Estados de Santa Catarina e Rio Grande do Sul compreendendo aproximadamente  $52800 \text{ Km}^2$  de superfície.

A região do estudo faz limite ao nordeste e leste com a bacia do rio Itajaí e demais bacias menores da vertente oceânica da Serra Geral, ao oeste com a bacia do rio Uruguai subsequente ao posto fluviométrico Passo Caxambú localizado após a foz do rio Passo Fundo. O limite ao Sul é com as bacias dos rios Jacuí e Taquari.

A área de interesse do estudo é limitada a jusante pelo posto fluviométrico Passo Caxambú, citado anteriormente, e localizado no rio Uruguai. A montante deste posto e já dentro da região do estudo se distingue o principal formador do rio Uruguai, o rio Pelotas, que tem suas nascentes na Serra Geral, em altitudes da ordem de 1200 m acima do nível do mar, declina cerca de 840 m em seus 450 Km de extensão aproximada. Já o rio Canoas, principal afluente da margem direita do rio Pelotas e que o iguala em porte, apresenta em seus 500 Km de extensão dois trechos distintos: o superior, de aproximadamente 320 Km de extensão, com declividades bastante suaves e o inferior com gradiente bem acentuado (cerca de 2m/Km).

A aproximadamente 100 Km a jusante da foz do rio Canoas, o rio Pelotas faz confluência com o rio Peixe e passa a ser chamado de Uruguai. Este último, tem um gradiente de 0,5m/km.

Ordenados de montante para jusante, os principais afluentes do rio Pelotas são os rios Canoas, Forquilha e Ligeiro. Os principais afluentes do rio Canoas são os rios Marombas pela margem direita e o Caveiras pela margem esquerda. Quanto ao rio Uruguai propriamente dito, seus afluentes mais importantes são o rio Irani e o rio Passo Fundo.

ELETROSUL (1979), descreve e analisa diversos fatores de interesse hidrológico na região de estudo. A seguir reproduzimos a informação mais relevante em termos de regionalização de características de vazão extrema.

O rio Canoas a montante de Ponte Alta é um rio com vale largo e menor variação de relevo que vales de outros rios na região. Observou-se que as peculiaridades fisiográficas deste trecho produzem um maior abatimento das ondas de cheia por propagação em canal, ao comparar os hidrogramas de enchentes do Canoas com os correspondentes do Pelotas. Isto é, para igualdade

de volumes, apresentam tempos de base maiores e picos menores e, por outro lado, para igualdade de áreas, o tempo de concentração nesta bacia é também maior que na bacia do Pelotas.

Uma paisagem morfológica bem distinta ao do alto Canoas é mais geral na região do estudo, mostrando relevo suave nos topos de planalto, e forte ondulado, junto as drenagens principais, com diferentes níveis de dissecação.

O curso principal é notável pelo seu vale estreito e profundo. As precipitações, com escoamento favorecido pelo desmatamento em boa parte da bacia e pelos solos pouco profundos e em geral pouco permeáveis, afluem rapidamente até os rios. Deste modo, o tempo de concentração é muito reduzido e nos casos de precipitações intensas as cheias são respostas quase imediatas. Adicionalmente, o vale em "V" permite flutuações muito grandes do nível das águas no rio. Exemplo disso é a variação de 13m na cota média de Passo Caxambú entre os dias 18 e 19 de agosto de 1965.

O clima da bacia do rio Uruguai tem as características de zona de transição ou temperada. Os fatores genéticos preponderantes constituem: o anticiclone móvel polar da América do Sul e o anticiclone do Atlântico Sul. O fator estático de maior influência na bacia é sua orografia, que atua fundamentalmente como regulador do elemento temperatura.

O anticiclone tropical marítimo do Atlântico Sul constitui uma massa de ar tropical marítima, que com sua subsidência superior e conseqüente inversão de temperatura, mantém a unidade limitada à camada superficial e a estabilidade do tempo.

As maiores enchentes da região são atribuídas a atividade das "Frentes Frias" que quando acontecem como sucessões, separadas por poucos dias, podem causar a ocorrência de situações críticas. A estabilidade do tempo na região de interesse do estudo pode também ser modificada pelas "correntes perturbadas do oeste" que se originam do forte aquecimento do interior do continente e do conseqüente decréscimo geral da pressão, fenômeno que dá origem às linhas de instabilidade ditas tropicais. Nestas, o ar em convergência provoca chuvas intensas geralmente de curta duração.

As precipitações médias anuais são superiores a 1400 mm em qualquer ponto da bacia, sendo a região do planalto no leste a menos chuvosa (média de 1400 mm por ano em Lages) e a região noroeste é a de maior precipitação (2200 mm por ano em Xanxerê).

#### 4.3. Dados disponíveis

Os dados utilizados neste estudo, foram os dados de vazão das bacias no rio Uruguai, conforme publicados no volume Boletim Fluviométrico Serie F-7.02: Bacia do rio Uruguai (1979). Esse volume apresenta dados de vazão até 1975. Uma análise mais completa teria incluído também os dados coletados depois desse ano, porém os dados publicados estavam mais prontamente disponíveis e atendiam aos propósitos do presente estudo.

Bacias com menos de 10 anos de registro de vazão (até 1975) não foram incluídas no estudo; as 29 bacias restantes foram aquelas apresentadas na tabela II.1, que também mostra as áreas das bacias. A figura II.1 ilustra a localização dos 29 postos fluviométricos correspondentes.

#### 4.4. Regionalização da vazão máxima

##### 4.4.1. Descrição do procedimento adotado

Preliminarmente, uma série anual de vazão máxima  $Y_1$  é construída para cada uma das 29 bacias do rio Uruguai. Os valores da variável  $Y_1$  são o maior valor de descarga registrada (de forma indireta) no posto fluviométrico em cada ano, isto é, a máxima vazão diária de cada ano.

As próximas etapas para regionalizar a vazão máxima  $Y_1$  com uso dos momentos L são mostrados na figura 4.1; onde se apresenta o método básico a ser investigado com a perspectiva de generalizá-lo. Paralelamente um método semelhante em que são usados momentos convencionais será processado a fim de fazer comparações.

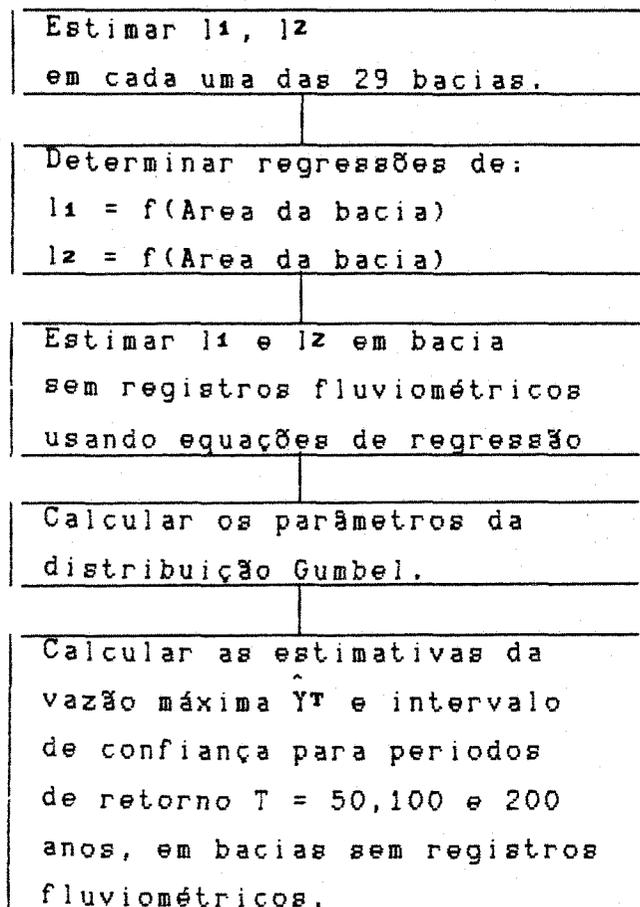


Figura 4.1- Fluxograma do método de regionalização básico

Assim, para a variável  $Y_1$  foram calculadas as estimativas  $l_1$  e  $l_2$  de momentos L em cada uma das 29 bacias do rio Uruguai. Adicionalmente para a vazão máxima  $Y_1$  também foram calculadas: as estimativas  $l_3/l_2$  e  $l_4/l_2^2$  das relações de momentos L, o desvio padrão e os coeficientes de assimetria e kurtosis. A tabela 11.2 mostra os valores das estimativas de momentos-L. Para extrapolar  $l_1$ ,  $l_2$  a outras bacias não medidas dentro do sistema do rio Uruguai é necessário determinar equações de regressão de  $l_1$  e  $l_2$  sobre características físicas e climáticas das bacias.

Nesta parte do trabalho e em vista de que estamos explorando o valor da utilização dos momentos L na regionalização da vazão máxima, duas simplificações foram adotadas: A primeira é a utilização de somente uma variável independente nas relações de regressão, esta é, a área da bacia.

Embora esta variável seja comumente testada em relações onde a variável dependente é uma estatística da cheia, existem outras características físicas e climáticas possíveis de serem incluídas como variáveis independentes. Assim por exemplo, com a finalidade de estabelecer relações de regressão entre a média e o coeficiente de variação da vazão máxima anual de rios na Inglaterra sobre características físicas e climáticas das bacias, NERC (1975) adotou um conjunto de índices dentre os quais escolhemos os mais representativos, estes são: área da bacia, precipitação média anual, densidade de drenagem, declividade do rio principal e percentagem da bacia drenada por lagos, reservatórios ou área urbana que poderão ser testados num estudo posterior como variáveis independentes.

A segunda simplificação é a utilização de uma regressão linear do tipo  $Y=a+bA+\epsilon$ , onde Y é o parâmetro estatístico que se deseja estimar, a e b são os parâmetros de interceptação e declividade da linha de regressão, A é a área da bacia e  $\epsilon$  são os erros da regressão. A adoção desta relação particularmente elementar facilitará a análise preliminar; posteriormente outras relações como as não lineares do tipo  $Y=aA^b\epsilon$ , que por transformação logarítmica são resolvidas como se se tratasse com relações de regressão lineares, poderão ser utilizadas.

A transferência de apenas dois dos primeiros momentos L como indicado na terceira etapa da figura 4.1, será realizada por serem estes suficientes para calcular a estimativa dos dois parâmetros da distribuição Gumbel no local sem medições.

Definidas as equações de regressão lineares de  $l_1$  e  $l_2$  sobre a área de drenagem, utilizaremos estas relações para estimar  $l_1$  e  $l_2$  em sub-bacias onde estes momentos L não podiam ser calculados por tratar-se de sub-bacias sem registros fluviométricos.

A distribuição Gumbel é inicialmente utilizada porque, por hipótese, seria a que descreve regionalmente a frequência de ocorrência da vazão máxima e por ser de fácil aplicação.

O próximo passo a fim de proporcionar um enfoque à investigação foi ainda pressupor que:

a) Descargas máximas anuais  $\hat{Y}_T$  com períodos de retorno  $T=50, 100$  e  $200$  anos foram requeridas para bacias individuais sem posto fluviométrico, cujas áreas de drenagem são iguais a  $0.5, 0.75, 1.0, 1.25,$  e  $1.5$  vezes a área média das 29 bacias usadas na análise. Na escolha dos valores do período de retorno foram considerados a quantidade de informação fluviométrica que seria utilizada, o objetivo do trabalho em sentido de que estes períodos de retorno seriam suficientes para analisar o valor do uso de momentos  $L$  no método de regionalização e finalmente, a utilidade já que são valores usados na prática.

b) Os intervalos de confiança aproximados de  $\pm 2 \sqrt{(\text{Var } \hat{Y}_T)}$ , eram necessários como estimativa da precisão de  $\hat{Y}_T$ . (O termo aproximados é usado para ressaltar que não sendo possível obter diretamente a distribuição de  $\hat{Y}_T$ , devemos considerar o uso do Teorema do Limite Central ao considerar a exatidão do valor calculado como intervalo de confiança de uma estatística  $\hat{U}$  como  $\hat{U} \pm 2 \sqrt{(\text{Var } \hat{Y}_T)}$ . Assim, em termos práticos consideramos que esta expressão proporciona os limites de confiança aproximados de 95%.)

A primeira modificação do procedimento básico de regionalização inicialmente proposto, seria provocado pela necessidade de utilização de uma função de distribuição diferente da de Gumbel. Efetivamente, se analisamos os valores das estimativas dos coeficientes de L-assimetria ( $t_3=I_3/I_2$ ) e L-kurtosis ( $t_4=I_4/I_2$ ) para as 29 bacias do estudo (tabela 1.2), verificamos que estas estatísticas com média  $\bar{t}_3=0.087$ , desvio padrão  $S_{t_3}=0.101$  e média  $\bar{t}_4=0.13$  com desvio padrão correspondente  $S_{t_4}=0.091$  não são bem representadas pela distribuição Gumbel que tem coeficientes de L-assimetria e L-kurtosis constantes, sendo  $\tau_3=0.170$  e  $\tau_4=0.150$ . Estas considerações nos levam a testar a utilização das distribuições Gamma (dois parâmetros) e Valores Extremos Generalizados (GEV) de uso corrente na tentativa de retratar a frequência de ocorrência da vazão máxima.

Embora para tratar estes dois últimos casos, seria desejável simplesmente adaptar o método básico de regionalização no que diz respeito a distribuição adotada, a obtenção de uma expressão para a variância da estimativa regional de vazão máxima, no caso em que a distribuição Gama e o

método de momentos L foram adotados, apresentou as seguintes dificuldades (Anexo V):

a) Os parâmetros da distribuição Gamma não podem ser escritos (de forma imediata), explicitamente em função dos momentos-L.

b) As expressões aproximadas dos parâmetros da distribuição em função dos momentos-L não favorecem a obtenção prática da expressão da  $\text{var}[\hat{Y}_T]$ .

c) A aproximação utilizada (equação V.3 do anexo V) a fim de conseguir uma expressão explícita para  $\hat{Y}_T$ , é válida apenas para os dados de algumas sub-bacias do Rio Uruguai.

Devido a isto, um método de regionalização alternativo será utilizado. Este consiste em:

1.- Adotar uma distribuição de probabilidades considerada adequada para descrever a distribuição de frequência da vazão máxima ( $Y_1$ ).

2.- Calcular; pelo método dos momentos L (e momentos), as estimativas  $\hat{Y}_0$ ,  $\hat{Y}_{100}$  e  $\hat{Y}_{200}$  da vazão máxima para cada uma das bacias; com dados fluviométricos consideradas no estudo.

3.- Estabelecer uma equação de regressão, para cada período de retorno escolhido ( $T=50, 100$  ou  $200$  anos), entre  $\hat{Y}_T$  e características hidrológicas da bacia (aqui só consideramos a área da bacia).

4.- Calcular a estimativa,  $\hat{Y}_T$ , da vazão máxima e seus intervalos de confiança para bacias sem registros fluviométricos cuja área de drenagem seja suposta igual a uma fração da área média, ou seja  $A=0.5\bar{A}$ ,  $A=1.0\bar{A}$  ou  $A=1.5\bar{A}$ .

A diferença básica está em que antes definíamos relações entre os momentos L e a área da bacia, agora estas relações são definidas entre a vazão máxima com T anos de período de retorno e a área das bacias.

#### 4.4.2. Distribuições de probabilidade utilizadas

Neste ítem, para cada distribuição de probabilidade adotada, isto é, para a distribuição GEV, para a distribuição Gamma e para a Gumbel, são discutidos alguns aspectos quanto ao seu uso na representação da frequência de ocorrência da vazão máxima; uma vez definidas através da sua função de densidade acumulada mostra-se como são definidos seus momentos L e suas

relações de momentos teóricos  $\tau_3$  e  $\tau_4$ ; ainda são apresentadas as equações resultantes da aplicação do método de momentos L para estimativa dos parâmetros.

A ordem de apresentação não coincide com a ordem em que as distribuições tinham sendo usadas, porque alguns aspectos da distribuição Gumbel são tratados de forma geral na distribuição GEV.

#### Distribuição de Valores Extremos Generalizada (GEV):

O uso da distribuição GEV em hidrologia tentou ser justificado pela teoria de valores extremos, que na sua parte mais destacada estabelece que a distribuição dos  $N$  maiores valores, cada um dos quais selecionado dentre os  $m$  valores encontrados em cada uma das  $N$  amostras, se aproximará de uma forma assintótica a medida que  $m$  cresce indefinidamente. Em hidrologia se adota  $N$  como o número de anos de registros de vazão diária e  $m$  corresponderia aos 365 dias do ano. O tipo de forma limite da distribuição dependerá do tipo da distribuição inicial dos  $Nm$  valores. Fréchet (1927) e Fisher e Tippett (1928), encontraram para três tipos diferentes de distribuição inicial as distribuições assintóticas de valores extremos denominadas tipo I, tipo II e tipo III que são casos particulares da distribuição GEV.

Assim, a cheia máxima anual constituiria uma série que poderia ser ajustada a uma das distribuições assintóticas; no entanto, não se pode esperar que a teoria dos valores extremos proporcione a distribuição exata para os picos da cheia pelos seguintes motivos:

- a) As descargas diárias não são eventos independentes.
- b) As  $m=365$  descargas diárias não constituem um número grande de observações.
- c) A decisão quanto ao tipo de distribuição a adotar se baseia na evidência empírica.

Apesar destes questionamentos à aplicação da teoria de valores extremos à vazão máxima, a distribuição GEV é bastante usada na modelação de fenômenos naturais extremos, já que ao ser uma distribuição de três parâmetros oferece bastante flexibilidade. A distribuição GEV tem sido recomendada em NERC

(1975) para modelar a distribuição do máximo anual das descargas diárias em rios da Inglaterra. A recomendação se baseia em ter sido apontada por testes estatísticos de qualidade de ajuste, de forma mais consistente que outras distribuições, para descrever a frequência de ocorrência da vazão máxima junto a constatação de que representa bem as curvas regionais adimensionais de vazão máxima. Além de que a estimativa dos seus parâmetros pelos métodos usuais, não é mais difícil que o cálculo das estimativas correspondentes para outras distribuições de três parâmetros.

A função de densidade acumulada para a distribuição de Valores Extremos Generalizada é:

$$F(Y) = \exp \left[ - \left( 1 - K(Y - \Theta) / \phi \right)^{1/K} \right], \quad K \neq 0 \quad (4.1)$$

$$= \exp \left[ - \exp \left( -(Y - \Theta) / \phi \right) \right], \quad K = 0 \quad (4.2)$$

onde,  $\Theta$  e  $\phi$  são os parâmetros de localização e escala, respectivamente. O parâmetro de forma  $K$  determina qual das distribuições de valores extremos é representada assim:

para tipo I	$K=0$	$-\infty < Y < \infty$
para tipo II	$K>0$	$-\infty < Y \leq \Theta + \phi / K$
para tipo III	$K<0$	$\Theta + \phi / K \leq Y < \infty$

notamos que quando  $K=0$ , estaríamos tratando com a distribuição Gumbel, também conhecida como a distribuição de valores extremos tipo I.

Para ajustar a distribuição GEV a um conjunto de observações é necessário estimar  $\Theta$ ,  $\phi$  e  $K$ . Estas estimativas quando usado o método de momentos  $L$  são funções de  $\lambda_1$ ,  $\lambda_2$  e  $\lambda_3$ .

Usando o método de momentos  $L$  descrito no item 2.3.1 as expressões para os momentos  $\lambda_1$  a  $\lambda_4$ , da distribuição GEV, são:

$$\lambda_1 = \Theta + \phi \left( 1 - \Gamma(1+K) \right) / K \quad (4.3)$$

$$\lambda_2 = \phi \left( 1 - 2^{-K} \right) \Gamma(1+K) / K \quad (4.4)$$

$$\lambda_3 = \phi \left( 1 - 3 \cdot 2^{-K} + 2 \cdot 3^{-K} \right) \Gamma(1+K) / K \quad (4.5)$$

$$\lambda_4 = \phi \left( 1 - 6 \cdot 2^{-K} + 10 \cdot 3^{-K} - 5 \cdot 4^{-K} \right) \Gamma(1+K) / K \quad (4.6)$$

onde  $\Gamma(\cdot)$  indica a função gamma do argumento entre parêntesis.

As relações de momentos L ou os coeficientes de L-assimetria e L-kurtosis são definidos por:

$$\tau_3 = \frac{\lambda_3}{\lambda_2} = 2 (1-3^{-K}) / (1-2^{-K}) - 3 \quad (4.7)$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2} = (5 (1-4^{-K}) - 10 (1-3^{-K}) + 6 (1-2^{-K})) / (1-2^{-K}) \quad (4.8)$$

as estimativas amostrais correspondentes são  $t_3 = l_3/l_2$  e  $t_4 = l_4/l_2$ .

Observamos que o coeficiente de L-assimetria apresentado na equação 4.7 é apenas função de K; portanto, se  $\tau_3$  é substituído pela estimativa amostral  $t_3$ , a solução desta equação pelo método iterativo de Newton nos proporcionará o valor de  $\hat{K}$ . Hosking e outros (1985) sugerem a seguinte solução aproximada:

$$\hat{K} \approx 7.8590 Z + 2.9554 Z^2$$

$$\therefore Z = 2/(3+t_3) - \log 2 / \log 3$$

Com  $\hat{K}$  conhecido a estimativa dos parâmetros  $\Theta$  e  $\phi$  são determinados pela solução das equações 4.4 e 4.3, sendo:

$$\phi = \frac{\lambda_2 K}{(1-2^{-K}) \Gamma(1+K)}, \quad \Theta = \lambda_1 - \phi (1-\Gamma(1+K)) / K \quad (4.9)$$

onde, a fim de determinar  $\hat{\phi}$  e  $\hat{\Theta}$ , é necessário substituir  $\lambda_1$  e  $\lambda_2$  pelas estimativas amostrais  $l_1$  e  $l_2$ .

Com os dados utilizados neste estudo tem-se obtido resultados praticamente iguais quando comparamos os quantiles obtidos quando na solução da equação 4.7 usamos o método de Newton ou a solução proposta por Hosking.

### Distribuição Gumbel:

A distribuição Gumbel é bastante usada na descrição probabilística da vazão máxima, talvez por ser de fácil aplicação e porque seus parâmetros permitem um bom ajuste aos dados em um intervalo baixo da curva de frequência; assim NERC (1985) recomenda seu uso quando o número de anos de registro é entre 10 e 25 anos e quando a extrapolação requerida é para  $T < 2N$ .

Como já temos indicado, esta distribuição é um caso particular da distribuição GEV e tem valores fixos de assimetria e kurtosis.

A função de densidade acumulada para a distribuição Gumbel é:

$$F(y) = \exp[-\exp(-(y-\theta)/\phi)] \quad \infty < y < \infty \quad (4.10)$$

onde  $\theta$  e  $\phi$  são os parâmetros de localização e escala respectivamente.

Usando o método de momentos L descrito no item 2.3.1 as expressões para os momentos-L de primeira e segunda ordem podem ser determinadas. Estas são:

$$\lambda_1 = \theta + \gamma \phi$$

$$\lambda_2 = \phi \ln 2$$

onde,  $\gamma=0.5772$  é o número de Euler.

Os coeficientes de L-assimetria e L-kurtosis são:

$$\tau_3 = 0.1699$$

$$\tau_4 = 0.1504$$

As expressões para calcular as estimativas dos parâmetros  $\theta$  e  $\phi$  foram obtidas no item 2.3.2 ao exemplificar a estimativa dos parâmetros de uma distribuição pelo método de momentos L.

### Distribuição Gamma:

A distribuição Gamma com dois parâmetros (Gamma incompleta); também conhecida como distribuição Pearson tipo III, é flexível e muito utilizada em hidrologia para descrever variáveis assimétricas com valores exclusivamente positivos, é o caso da vazão máxima anual. Seu uso neste caso não tem nenhum fundamento teórico, sendo adotada simplesmente por ser considerada de forma

apropriada e com expressão matemática conveniente.

A função de densidade acumulada para a distribuição Gamma é:

$$F(y) = \frac{1}{\Gamma(\alpha)} \int_0^{y/\beta} t^{\alpha-1} e^{-t} dt \quad 0 < y < \infty \quad (4.11)$$

onde;  $t$  é a variável de integração  $\Gamma(\cdot)$  é a função gamma,  $\alpha$  e  $\beta$  são respectivamente os parâmetros de forma e escala da distribuição. Usando o método de momentos L, os momentos  $\lambda_1$  e  $\lambda_2$  são definidos como:

$$\lambda_1 = \alpha \beta \quad (4.12)$$

$$\lambda_2 = \pi^{-1/2} \beta \Gamma(\alpha + 1/2) / \Gamma(\alpha) \quad (4.13)$$

ao combiná-las, obtemos:

$$\tau = \lambda_2 / \lambda_1 = \pi^{-1/2} \Gamma(\alpha + 1/2) / \Gamma(\alpha + 1) \quad (4.14)$$

cuja solução nos permitirá encontrar o valor de  $\alpha$ . Hosking (1986), sugere o seguinte processo iterativo.

$$\alpha_{(i+1)} = \frac{\pi^{-1/2} \Gamma(\alpha_{(i)} + 3/2)}{\tau \Gamma(\alpha_{(i)} + 1)} - \frac{1}{2}$$

com valor inicial

$$\alpha_{(0)} = \begin{cases} 1/(\pi \tau^2) & \text{se } \tau < 0.6 \\ -\log(\tau)/\log(4) & \text{se } \tau \geq 0.6 \end{cases}$$

depois poderá estimarse  $\beta$  com:

$$\beta = \lambda_1 / \alpha \quad (4.15)$$

#### 4.4.3. Resultados

##### a) Adoção da distribuição Gumbel

##### Transferência da informação.-

Como já temos indicado no caso em que a distribuição Gumbel é adotada para descrever a frequência de ocorrência da vazão máxima, a transferência da informação relativa aos registros de cheia na região de estudo é feita

através de equações de regressão dos momentos  $l_1$  (igual a média) e  $l_2$  sobre a área das bacias. Na tabela III.4 mostra-se detalhes destas regressões. Também é mostrado o resultado da regressão em que a variável dependente é o desvio padrão ( $S$ ); lembrando que  $l_2$  e  $S$  são medidas de dispersão, é de ressaltar que a variância residual da regressão de  $l_2$  é muito menor que a correspondente a regressão de  $S$ . Este resultado é importante na avaliação do intervalo de confiança da estimativa de  $Y_T$  por momentos  $L$  e momentos.

Estimativa da vazão máxima,  $\hat{Y}_T$ , e intervalo de confiança.-

Como a função de densidade acumulada para a distribuição Gumbel é dada por:

$$F(y) = \exp[-\exp -(y-\theta)/\phi]$$

na notação anteriormente usada, o evento com probabilidade  $P[Y>y]=1/T$  de ser excedido é o quantil  $Y_T$  da função de distribuição  $F(y)=P[Y\leq y]$ , onde  $F=(T-1)/T$ , e corresponde a estimativa da descarga máxima anual com período de retorno de  $T$  anos dada por:

$$\hat{Y}_T = \hat{\theta} - \hat{\phi} \ln \ln (T/[T-1]) \quad (4.16)$$

com variância dada por:

$$\text{var } \hat{Y}_T = \text{var}(\hat{\theta}) + [\ln \ln (T/(T-1))]^2 \text{var}(\hat{\phi}) - 2 [\ln \ln (T/(T-1))] \text{cov}(\hat{\theta}, \hat{\phi}) \quad (4.17)$$

Se os momentos- $L$  são utilizados, os parâmetros  $\hat{\theta}$  e  $\hat{\phi}$  na equação 4.16 são obtidos por suas relações com os momentos  $l_1$  e  $l_2$ , estes últimos estimados das equações de regressão de  $l_1$  e  $l_2$  sobre a área da bacia  $A$ . Neste caso as variâncias e covariâncias necessárias para o cálculo da  $\text{var}(\hat{Y}_T)$  com a expressão 4.17 são dadas por:

$$\text{var } \hat{\phi} = \text{var}(l_2)/(\ln 2)^2$$

$$\text{var } \hat{\theta} = \text{var}(l_1) + (\gamma/\ln 2)^2 \text{var } l_2 - 2(\gamma/\ln 2) \text{cov}(l_1, l_2)$$

$$\text{cov}(\hat{\theta}, \hat{\phi}) = (1/\ln 2) \text{cov}(l_1, l_2) - \gamma \text{var } l_2 / (\ln 2)$$

onde  $\gamma = 0.577\dots$  é a constante de Euler, a  $\text{var}(l_1)$ ,  $\text{var}(l_2)$  e  $\text{cov}(l_1, l_2)$  são obtidas também das equações de regressão de  $l_1$  e  $l_2$  sobre a área da bacia  $A$ , após substituição de  $A=0.5\bar{A}$ ,  $A=1.0\bar{A}$ , ou  $A=1.5\bar{A}$ .

Desta forma as equações 4.16 e 4.17 fornecerão estimativas e suas variâncias respectivamente, das descargas máximas anuais com períodos de

retorno T=50, 100 ou 200 anos para bacias sem registros de vazão, cuja área de drenagem hipotética seja igual a 0.5, 1 ou 1.5 vezes a área média das bacias usadas no estudo, quando se pressupõe uma distribuição Gumbel, e quando os parâmetros de Gumbel foram regionalizados através da regressão de momentos L sobre a área da bacia.

Um procedimento similar ao descrito acima foi seguido para o caso em que os momentos convencionais foram usados. Nesta situação a estimativa regionalizada da vazão máxima ( $\hat{Y}_T$ ) e sua variância ainda são dadas pelas expressões 4.16 e 4.17 sendo que os parâmetros de Gumbel da equação 4.16 são estimados pelo método de momentos convencionais, depois de obter a média (m) e o desvio padrão (s) das equações de regressão de m e S sobre a área da bacia A. A  $\text{var}(\hat{\phi})$ ,  $\text{var}(\hat{\Theta})$  e  $\text{cov}(\hat{\Theta}, \hat{\phi})$  a serem usadas na expressão 4.17 agora são:

$$\text{var } \hat{\phi} = 6 \text{ var}(s)/\pi^2$$

$$\text{var } \hat{\Theta} = \text{var}(m) + 6 \gamma^2 \text{ var}(s)/\pi^2 - 2 \gamma \sqrt{(6/\pi^2)} \text{ cov}(m, s)$$

$$\text{cov}(\hat{\Theta}, \hat{\phi}) = \sqrt{(6/\pi^2)} \text{ cov}(s, m) - 6 \gamma \text{ var}(s)/\pi^2$$

Pode-se ver que o uso adequado das expressões 4.16 e 4.17 nos proporcionará a estimativa regionalizada da vazão máxima  $\hat{Y}_T$  e sua variância ( $\text{var } \hat{Y}_T$ ) para os casos em que os momentos L ou os momentos convencionais sejam utilizados. Prosseguindo, se atribuirmos a distribuição Normal a estimativa regionalizada  $\hat{Y}_T$ , seus limites de confiança de 95% podem ser dados em termos práticos, aproximadamente por:

$$\hat{Y}_T \pm 2 \sqrt{\text{var}(\hat{Y}_T)}$$

e a amplitude desse intervalo de confiança poderá ser comparada quando obtida pelo uso de momentos L com a amplitude obtida quando os momentos convencionais são utilizados.

Assim, pressupondo que a vazão máxima anual tenha uma distribuição Gumbel, a tabela III.5 mostra as estimativas das máximas anuais com períodos de retorno de 50, 100 e 200 anos e seus intervalos de confiança, para sub-bacias que se encontram na área de estudo e que carecem de registros fluviométricos, com áreas de drenagem iguais aos múltiplos da área média das 29 bacias.

Pode ser visto que onde os momentos L regionalizados foram utilizados para estimar os parâmetros de Gumbel, as amplitudes aproximadas do intervalo de confiança para as estimativas de máximos anuais com períodos de retorno dados, foram consistentemente menores do que as amplitudes onde a média e desvio padrão regionalizados foram utilizados na estimativa dos parâmetros Gumbel. Este resultado é consistente com os resultados de Landwehr et al (1979) que baseados em experimentos de Monte Carlo obtiveram, para a estimativa de quantiles via momentos L, eficiências relativas a estimativa via momentos de 0.74 e 0.77 para períodos de retorno de T=50 e 100 anos, estes valores no entanto são considerados teóricos. Para efeitos de estudos de regionalização e sem considerar o desempenho pelo uso de momentos L ou momentos, relativo a outros aspectos que intervém na regionalização, os valores encontrados por Landwehr et al podem ser considerados ideais.

Contudo, com as suposições usadas até aqui, pode ser visto que o uso de momentos L não conduz a um grande aumento na eficiência das estimativas de  $Y_T$  para bacias sem registros fluviométricos na região de estudo.

A comparação das estimativas  $\hat{Y}_T$  da tabela III.5, denotadas agora como  $\hat{Y}_{P,l}$  para indicar as obtidas pela regionalização dos momentos-L e  $\hat{Y}_{P,c}$  as obtidas pela regionalização da média e do desvio padrão, mostra que as estimativas  $\hat{Y}_{P,l}$  são consistentemente maiores do que as estimativas  $\hat{Y}_{P,c}$ . O motivo desta diferença é analisado a seguir.

As expressões para o cálculo das estimativas da vazão máxima, suposta a distribuição Gumbel, podem também ser escritas como :

$$\hat{Y}_{P,l} = l_1 - 1.4427 l_2 ( \gamma + \ln \ln (T/(T-1)) ) \quad (4.18)$$

$$\hat{Y}_{P,c} = m - 0.7798 S ( \gamma + \ln \ln (T/(T-1)) ) \quad (4.19)$$

como as expressões entre (.) para os períodos de retorno usados proporcionam valores negativos, e pela proporção usualmente encontrada na magnitude dos valores de  $l_2$  e  $S$ , ficou evidente que em geral  $Y_{P,l} > Y_{P,c}$ . Porém esta diferença é atribuída em grande parte a tendenciosidade da estatística  $S = \sqrt{\sum (X_i - \bar{X})^2 / n}$  usada na elaboração da tabela III.5. Na tabela III.1 mostra-se resultados da estimativa da vazão máxima para T=100 anos nos casos em que a

estatística tendenciosa  $S$ , a estatística não tendenciosa  $S^* = \sqrt{\sum(X_i - \bar{X})^2 / (n-1)}$  e  $I_2$  são utilizadas. A menor diferença na magnitude entre  $\hat{Y}_{P,1}$  e  $\hat{Y}_{P,c}^*$  confirma a característica de não tendenciosidade de  $\hat{Y}_{P,1}$  decorrente do uso também de uma estatística não tendenciosa para  $I_2$ .

Estabelecer qual das estimativas é menos tendenciosa não tem interesse prático, já que  $\hat{Y}_{P,1}$  é aproximadamente igual a  $\hat{Y}_{P,c}^*$  para  $T=50$ ,  $100$  e  $200$  anos. Contudo se deve indicar que resultados baseados em experimentos de Monte Carlo, encontrados por Landwehr et al (1979), ressaltam que a diferença dos métodos de máxima verossimilhança e momentos, o método dos MPP; para o caso em que as amostras pertençam a um processo puramente aleatório, permitem obter estimativas de quantiles sem tendenciosidade. No caso em que no processo exista uma correlação na série, mas seja tratada como puramente aleatória, haverá nos resultados uma tendenciosidade que será menor quando se usa MPP que para os outros métodos.

Esclarecemos que, se no processo de estimativa regionalizada de  $Y_T$  e de seu intervalo de confiança por momentos  $L$  ou momentos convencionais, optássemos por estabelecer equações de regressão onde  $\hat{Y}_T$  (o quantil para cada sub-bacia) ou as estimativas dos parâmetros de Gumbel,  $\hat{\theta}$  e  $\hat{\phi}$  para cada sub-bacia, fossem a variável dependente usada no ajuste da relação de regressão, os resultados obtidos seriam iguais aos apresentados na tabela III.5, onde a variável dependente das equações de regressão é um dos momentos  $L$  ou dos momentos convencionais.

#### Análise da eficiência, por amostragem aleatória simples.-

Entendendo que as áreas das bacias supostas sem registros fluviométricos não corresponderiam a bacias verdadeiramente encontradas na região de estudo, e a fim de procurar resultados para situações reais, procedeu-se da seguinte maneira: Da amostra de 29 bacias anteriormente usadas, se retirou uma bacia. Para esta bacia, suposta agora sem registros fluviométricos, foram calculadas as estimativas da vazão máxima com período de retorno  $T=50$ ,  $100$  e  $200$  anos junto a seus intervalos de confiança.

Repetindo o procedimento acima para cada uma das 29 bacias se obteve os valores médios dos intervalos de confiança de  $\hat{Y}_T$  mostrados na tabela III.2. Nesta, observamos que o uso de momentos L proporciona um valor menor para o intervalo de confiança médio das estimativas que o uso de momentos. A diferença na amplitude do intervalo de confiança (Dif), no uso destes métodos aumenta para períodos de retorno maiores.

A fim de verificar o comportamento dos métodos quando na regionalização, se usa um menor número de postos. O procedimento acima descrito é executado com a diferença de que agora 14 postos são retirados também por amostragem aleatória. Repetindo este procedimento 29 vezes, obtivemos os intervalos de confiança médios das estimativas da vazão máxima mostrados na tabela III.2, onde observamos que o uso de momentos L proporciona um valor menor para o intervalo de confiança médio das estimativas que o uso de momentos. A diferença no uso destes métodos também aumenta para períodos de retorno maiores. Ainda se observa que a diferença nos intervalos de confiança médios é maior quanto menor for a quantidade de postos usados na regionalização.

Adicionalmente na tabela III.3 são apresentadas as estimativas regionalizadas ( $\hat{Y}_{100}$ ) da vazão máxima diária com 100 anos de período de retorno e seus intervalos de confiança em bacias da região do estudo supostas sem registros. Este tipo de resultados é também mostrado no caso em que registros de vazão destas bacias são utilizados. Observa-se que a amplitude do intervalo de confiança das estimativas é menor quando usamos momentos-L do que quando utilizamos momentos comuns nas tres bacias escolhidas.

#### Regressão ponderada. -

No ajuste das relações de regressão regionais para transferência da informação da cheia a variável dependente utilizada ( $l_1$  e  $l_2$ ,  $m$  e  $S$  ou  $\hat{Y}_T$ ) é uma estimativa da característica da vazão máxima obtida a partir de  $n$  anos de registros fluviométricos correspondentes a cada um dos postos utilizados. Como o número de anos de informação ( $n$ ) é diferente para cada posto e como a confiabilidade das estimativas de  $l_1$ ,  $l_2$ ,  $m$ ,  $S$  ou  $\hat{Y}_T$  depende do comprimento de registros utilizado, é interessante verificar os resultados da

transferência da informação da cheia depois de atribuir às estimativas utilizadas um peso  $w_i$  que é igual ao número de anos em que se baseia seu cálculo. Estas considerações são levadas em conta no ajuste da relação:

$$\hat{Y}_T = a + b A + v \quad 4.20$$

como uma regressão ponderada, onde a vazão máxima com T anos de período de retorno ( $\hat{Y}_T$ ), participa no ajuste da regressão com um peso  $w_i$  igual ao número de anos de registros usados na sua estimativa. a e b são os parâmetros a serem estimados pelo minimização dos erros ponderados, A é a área da bacia e v são os erros da regressão ponderada. Assim, a estimativa da vazão máxima com T anos de período de retorno para um local sem posto fluviométrico seria dada por :  $\hat{Y}_T = \hat{a} + \hat{b} A$ . As expressões para cálculo dos parâmetros  $\hat{a}$  e  $\hat{b}$  são apresentadas no anexo IV. A expressão que proporciona a variância da estimativa obtida pela regressão ponderada, dada a área da bacia  $A^*$  é:

$$\text{Var}(\hat{Y}_T) = \hat{S}^2 \left( \frac{1}{n^*} + \frac{1}{\sum_{i=1}^N n_i} + \frac{(A^* - \bar{A})^2}{\sum_{i=1}^N n_i (A_i - \bar{A})^2} \right)$$

onde  $\hat{S}^2$  é a estimativa da variância residual da regressão ponderada cuja expressão é encontrada no anexo IV,  $\bar{A}$  é a média ponderada das áreas de drenagem das N bacias usadas na regionalização,  $A^*$  é a área da bacia sem registros de vazão,  $n_i$  é o número de anos de registro da vazão máxima em cada posto; finalmente observamos que seria preciso conhecer o número de anos,  $n^*$ , de registros fluviométricos justamente na bacia onde não se dispõe dos mesmos. Assumimos neste ponto que a estimativa  $\hat{Y}_T$  a partir das características físicas da bacia resulta em uma variância de estimativa semelhante a que seria obtida para um registro  $n^*=1$  ano. No anexo IV justificamos esta adoção.

Ainda adotando a distribuição Gumbel para representar a frequência de ocorrência da vazão máxima, as estimativas  $\hat{Y}_T$ , calculadas pelo uso dos momentos L e pelo uso dos momentos convencionais em cada um dos postos da região do estudo, são transferidas através da equação de regressão ponderada a locais sem postos fluviométricos com área de drenagem  $A^*=0.5\bar{A}$ ,  $A^*=1.0\bar{A}$ ,  $A^*=1.5\bar{A}$ . Os resultados correspondentes da estimativa da vazão máxima para períodos de retorno T=50,100 e 200 anos junto com seu intervalo de confiança

são apresentados na tabela III.6. Observa-se que o uso de momentos L produz estimativas um pouco mais eficientes que quando se usam momentos convencionais e que a magnitude da cheia estimada pelo uso de momentos L é consistentemente maior que a magnitude estimada pelo uso de momentos convencionais. Entretanto, a amplitude dos intervalos de confiança quando se usa regressão ponderada são excessivamente grandes em relação às amplitudes dos intervalos quando a regressão não é ponderada, como pode verificar-se pela comparação dos resultados das tabelas III.5 e III.6.

Embora o uso da regressão ponderada seja justificada pelas características da variável dependente já discutidas, observamos que a excessiva amplitude nos intervalos de confiança encontrados para o caso do estudo, é atribuído ao efeito de  $w_i = n_i$  nos erros da regressão ponderada, analisado a seguir.

Se consideramos que o erro residual  $v$  em 4.20 pode decompor-se em um erro  $\epsilon$  atribuído ao modelo e um erro de amostragem  $\eta$ , onde  $\eta$  é atribuído ao fato da variável independente ser estimada a partir de uma amostra finita com diferentes comprimentos de registros, notamos que  $w_i = n_i$  deveria afetar apenas uma parcela do erro  $v$ . Porém, a equação IV.1 que nos leva a ajustar a regressão ponderada por mínimos quadrados é  $\sum w_i v^2$ . Uma forma de resolver o problema seria procurar uma função para  $w_i$  sensível a magnitude das parcelas de erro que estariam sendo representadas por  $v$ . Tasker (1980), sugere uma função deste tipo para o caso em que a distribuição log Pearson tipo III seja adotada. Neste sentido, Stedinger e Tasker (1985, 1986 a b), ampliaram a consideração dos fatores que intervêm no erro de amostragem incluindo também a correlação existente entre as séries de vazão extrema anual. A técnica dos mínimos quadrados foi ainda usada, porém de uma forma mais geral. Exemplos da aplicação desta técnica de mínimos quadrados generalizados (GLS) são encontrados no trabalho de Tasker et al (1986) que estabeleceu uma relação regional entre a vazão máxima com T anos de período de retorno e características fisiográficas de bacias numa região dos Estados Unidos. Vogel e Kroll (1990) usaram a técnica para estabelecer relações de regressão entre a vazão mínima com diferentes durações e período de retorno T com características fisiográficas de bacias em Massachusetts. Uma conclusão importante, comum destes autores, é que a consideração do erro de amostragem

teria pouco impacto na estimativa dos parâmetros e no erro de predição da regressão quando se usam postos com mais de 10 anos de observações e quando o erro de amostragem não é muito incrementado pela existência de grande correlação (superior a 0.8) entre as séries de vazão extrema.

A consideração dos erros de amostragem na regressão da forma como é descrita neste item não é mais explorada neste trabalho. Continuamos o estudo com o uso de regressões do tipo mais usual, deixando a abordagem de mínimos quadrados generalizados para um estudo posterior específico.

#### b) Adoção da distribuição Gamma

##### Estimativa das curvas de frequência individuais.-

Como já temos mencionado, no caso em que adotamos a distribuição Gamma, serão estabelecidas equações de regressão das cheias com período de retorno  $T=50, 100$  e  $200$  anos sobre características hidrológicas das bacias. Para isto, as cheias correspondentes serão obtidas das curvas de frequência ajustadas para cada posto.

Assumindo que a série anual  $Y_t$  em determinado posto tem como elementos valores aleatórios e independentes,  $y_i$ , onde  $i=1, \dots, n$ , e  $n$  é o número de anos de registro, o evento com  $T$  anos de período de retorno é o evento com probabilidade  $P[Y > y] = 1/T$  de ser excedido e corresponde ao quantil  $Y_T$  da função de distribuição  $F(y) = P[Y \leq y]$ , onde  $F = (T-1)/T$ . Adotando que a função de distribuição Gamma, definida pela expressão 4.11, é a que descreve a frequência de ocorrência da descarga máxima anual, a vazão máxima com  $T$  anos de período de retorno ( $Y_T$ ) é estimada através de:

$$\frac{T-1}{T} = \frac{1}{\Gamma(\alpha)} \int_0^{Y_T/\beta} t^{\alpha-1} e^{-t} dt \quad 0 < Y < \infty$$

onde  $\alpha$  e  $\beta$  são os parâmetros de forma e escala da distribuição Gamma e podem ser estimados pelo método de momentos L através das equações 4.14 e 4.15. Como  $Y_T$  na equação acima não tem expressão analítica explícita, o uso do método de Newton resultará no seguinte processo iterativo:

$$x_{(i+1)} = x_{(i)} + \frac{T - \int_0^{Y_T/\beta} t^{\alpha-1} e^{-t} dt}{(T-1)(x_{(i)}^{\alpha-1} e^{-x_{(i)}})}$$

com valor inicial  $x_0 = \lambda_1/\beta$ , cuja solução nos proporciona o valor de  $Y_T/\beta$ . Desta forma é possível calcular as estimativas de vazão máxima  $\hat{Y}_T$  para o período de retorno escolhido ( $T=50, 100, 200$  anos).

Usando o procedimento acima descrito em cada uma das 29 bacias do rio Uruguai, foram obtidos os valores da variável independente a serem usados no ajuste das equações de regressão.

A fim de realizar comparações, o procedimento acima será repetido com os parâmetros da distribuição Gamma estimados pelo método de momentos.

#### Transferência da informação.-

A tabela III.7 mostra detalhes das regressões de  $Y_{50}$ ,  $Y_{100}$  e  $Y_{200}$  sobre a área da bacia quando os momentos L foram usados na estimativa dos parâmetros da distribuição Gamma. Com as mesmas considerações a tabela III.8 mostra o caso em que o método dos momentos fora utilizado. Observa-se que a variância residual da regressão é menor para o caso em que os momentos L são usados.

#### Estimativa regional da vazão máxima $\hat{Y}_T$ e intervalo de confiança.-

A tabela III.9 mostra as estimativas das descargas máximas anuais com períodos de retorno 50, 100 e 200 anos, obtidos com as equações de regressão da tabela III.7 (correspondentes ao uso de momentos L) após a substituição da variável independente por  $A=0.5\bar{A}$ ,  $A=1.0\bar{A}$ , e  $A=1.5\bar{A}$  que são as áreas de drenagem supostas para bacias sem registros fluviométricos. Os intervalos de confiança aproximados das estimativas de  $Y_T$  são também mostradas e foram calculadas como  $\pm 2 \sqrt{\text{Var}(\hat{Y}_T)}$ , onde  $\text{Var}(\hat{Y}_T)$  é obtida a partir da equação de

regressão como:

$$\text{Var}(\hat{Y}_T) = \hat{\sigma}^2 \left( 1 + \sum_{i=1}^N \left[ \frac{1}{N} + \frac{(A^* - \bar{A})(A_i - \bar{A})}{\sum_{i=1}^N (A_i - \bar{A})^2} \right]^2 \right)$$

onde; N é o número de postos fluviométricos usados na regionalização,  $\bar{A}$  a média das áreas de drenagem das N bacias,  $A^*$  é uma fracção de  $\bar{A}$ , por exemplo a área da bacia sem registros fluviométricos onde se quer a estimativa da vazão máxima é  $A^* = 0.50\bar{A}$ ,  $A_i$  é a área de cada uma das bacias e  $\hat{\sigma}^2$  é a estimativa da variancia residual da regressão.

Seguindo um procedimento similar ao acima descrito e a fim de realizar comparações, foram calculadas estimativas regionalizadas da vazão máxima e seus intervalos de confiança, para o caso em que o método de momentos fora utilizado. Os resultados são também mostrados na tabela III.9, onde se observa, que a amplitude do intervalo de confiança das estimativas obtidas via momentos L é um pouco menor.

### c) Adocção da distribuição de Valores Extremos Generalizada

#### Estimativa das curvas de frequência individuais.-

No caso em que adotamos a distribuição de Valores Extremos Generalizada (GEV), também serão estabelecidas equações de regressão, das cheias com período de retorno T=50, 100 e 200 anos sobre características hidrológicas das bacias. Para isto os parâmetros  $\Theta$ ,  $\phi$  e K da distribuição GEV são estimados para cada posto a partir da sua série anual de vazão máxima através do método de momentos L e momentos convencionais. Detalhes com relação ao primeiro método usado com a distribuição GEV foram apresentados no item 4.4.2. Qualquer que seja o método de estimativa dos parâmetros da distribuição GEV, a descarga máxima anual com período de retorno de T anos é:

$$Y_T = \Theta + \phi \left( 1 - \left( \ln \frac{T}{T-1} \right)^K \right) / K \quad K \neq 0$$

$$Y_T = \Theta - \phi \ln \left( \ln \frac{T}{T-1} \right) \quad K = 0$$

que nos proporciona os valores da variável dependente a serem usados no ajuste das equações de regressão mencionadas.

Quanto a utilização dos métodos de momentos-L e momentos na estimativa dos parâmetros da distribuição GEV, uma maior simplicidade de aplicação do primeiro método fora observada.

#### Transferência da informação.-

A tabela III.10 mostra detalhes das regressões de  $Y_{50}$ ,  $Y_{100}$ , e  $Y_{200}$  sobre a área da bacia quando os momentos L foram usados na estimativa dos parâmetros da distribuição GEV. Com as mesmas considerações a tabela III.11 mostra o caso em que o método dos momentos fora utilizado. Observa-se que no uso de momentos L somente as relações regionais para  $Y_{50}$  e  $Y_{100}$  tem menor variancia residual da regressão. No caso de  $Y_{200}$  o uso de momentos se mostra favorável.

## Estimativa regional da vazão máxima $\hat{Y}_T$ e intervalo de confiança.-

Com um procedimento análogo ao descrito para o caso em que a distribuição Gamma foi adotada para os mesmos períodos de retorno e bacias sem registros fluviométricos ali considerados, mas adotando agora a distribuição GEV, são obtidos os resultados da tabela III.12 que mostra as estimativas da vazão máxima e intervalos de confiança para o caso em que os parâmetros da distribuição são estimados pelo método dos momentos L. Resultados para o caso em que a distribuição GEV e o método de momentos é usado são mostrados na mesma tabela. Nota-se, que os intervalos de confiança das estimativas de 50 e 100 anos de período de retorno são menores quando se usa momentos L. No caso das estimativas  $\hat{Y}_{200}$  o uso de momentos se mostra vantajoso.

### 4.5.- Regionalização da vazão mínima

#### 4.5.1 Distribuições de probabilidade adotadas

Uma das formas de descrever as características da descarga mínima em um rio é através da curva de frequência da vazão mínima. Assim, a série de descargas mínimas anuais de certa duração é suposta ser uma amostra retirada de certa distribuição de probabilidades  $F(Y)$  cujos parâmetros são estimados a partir da amostra.

Desta forma a magnitude da vazão mínima de certa duração é associada empiricamente a uma frequência de ocorrência. A única modificação com relação ao que fora executado para a vazão máxima é que agora estamos interessados no evento cuja probabilidade de não ser superado é pequena. Isto é  $P\{Y \leq y\} = 1/T$ , onde T em anos é o período de retorno desejado.

As distribuições de probabilidade aqui adotadas são a distribuição Gamma e a distribuição Weibull, esta última é a distribuição de valores extremos tipo III. Ambas distribuições já foram discutidas no item 4.4.2.

Embora as distribuições Weibull e Gamma permitam o ajuste de três

parâmetros, usamos a parametrização que permite usar só dois parâmetros, tornando o limite inferior da função de densidade igual a zero. Isto evita a obtenção de valores negativos para as estimativas de descarga mínima com períodos de retorno elevado, em algumas bacias.

#### 4.5.2. Caraterização da descarga mínima

A vazão mínima ( $Y_z$ ), neste caso, é o menor valor obtido do cálculo da média das descargas em dado período de tempo em um ano hidrológico. Assim, só um evento correspondente a vazão mínima é selecionado para cada ano hidrológico de uma série de dados.

O período de tempo dado é aqui adotado como sendo 1 dia, embora outros períodos de tempo poderão ser usados dependendo da aplicação e o problema específico que esteja sendo analisado. Assim, por exemplo, Matias (1963) sugere a adoção de um período de 7 dias se for necessário remover os efeitos de flutuações de certa regularização no rio.

#### 4.5.3. Método de regionalização adotado

A fim de analisar o desempenho do uso de momentos L na regionalização da vazão mínima na região do estudo na bacia do rio Uruguai, se adotou em termos gerais, o procedimento alternativo descrito para o caso em que a variável estudada era a vazão máxima, e que pode ser assim resumido:

- a) Adotar uma distribuição de probabilidades considerada adequada para descrever a distribuição de frequência da vazão mínima ( $Y_z$ ).
- b) Calcular pelo método dos momentos- L (e momentos), as estimativas  $\hat{Y}_0$ ,  $\hat{Y}_{100}$  e  $\hat{Y}_{200}$  da vazão mínima para cada uma das bacias com dados fluviométricos consideradas no estudo.
- c) Estabelecer uma equação de regressão para cada período de retorno escolhido ( $T=50, 100$  ou  $200$  anos), entre  $\hat{Y}_T$  e características hidrológicas da bacia (aqui só consideramos a área da bacia).
- d) Calcular a estimativa,  $\hat{Y}_T$ , da vazão mínima e seus intervalos de confiança para bacias sem registros fluviométricos cuja área de drenagem

seja suposta e igual a uma fração da área média, ou seja  $A=0.5\bar{A}$ ,  $A=1.0\bar{A}$  ou  $A=1.5\bar{A}$ .

#### 4.5.4. Resultados

Seguindo o procedimento descrito no item anterior foram calculadas as estimativas da vazão mínima de duração 1 dia ( $\hat{Y}_T$ ) para períodos de retorno pre-estabelecidos em bacias sem registros de vazão na região do estudo. Na tabela III.13 são mostrados os valores destas estimativas junto aos seus intervalos de confiança para os casos em que os momentos-L e os momentos foram utilizados junto a adoção da distribuição Weibull. Resultados para os mesmos casos mas quando a distribuição Gamma é adotada são mostrados na tabela III.14.

Observa-se que o uso de momentos-L independente da distribuição adotada proporciona estimativas cujas amplitudes do intervalo de confiança são um pouco menores que as amplitudes correspondentes quando os momentos são utilizados.

A adoção da distribuição Weibull junto ao uso de momentos-L mostra o melhor desempenho em termos da amplitude do intervalo de confiança. Neste caso a magnitude das estimativas  $\hat{Y}_T$  são consistentemente menores.

## 5.- CONCLUSões

### 5.1. Estimativa regional da vazão máxima com T anos de período de retorno:

Com a finalidade de facilitar a comparação dos resultados, as estimativas regionalizadas da vazão máxima anual na bacia do rio Uruguai apresentadas neste trabalho, serão referidas como GU/1, GA/1 ou GEV/1 quando se trate de estimativas onde os momentos L foram usados e a distribuição Gumbel, Gamma ou Valores Extremos Generalizada é utilizada. Para as mesmas distribuições, mas quando os momentos são usados as estimativas são denominadas GU/m, GA/m ou GEV/m.

As seguintes observações podem ser feitas, com base no uso dos momentos L em regionalização e com base nos resultados obtidos.

1.- O simples uso de momentos L não garante estimativas regionais de vazão máxima com maior eficiência que as conseguidas pelo uso de momentos, isto é verificado pelo intervalo de confiança da estimativa GEV/m com período de retorno de 200 anos que é menor que o intervalo de confiança da estimativa GEV/1 correspondente. Além deste resultado particular também se deve indicar, que embora na análise de eficiência por amostragem aleatória simples (Tab.III.2), os valores médios da amplitude do intervalo de confiança são menores para as estimativas GU/1, não se garante que em um caso particular isto não venha a acontecer, já que os valores médios ocultam a ocorrência destes casos.

2.- Apesar das observações acima, se verifica que as melhores estimativas em termos de intervalo de confiança (para todos os casos estudados) são: a estimativa GEV/1 para 50 anos de período de retorno, a GU/1 para 100 anos de período de retorno e a GA/1 para T=200 anos.

3.- A perda de eficiência das estimativas GU/1 dentro da gama de períodos de retorno considerados e distribuições adotadas é pequena. A comparação no mesmo sentido de GA/1 e GEV/1 favorece as estimativas GU/1. Conseqüentemente, em termos práticos e dentro dos limites das situações consideradas no estudo

as estimativas GU/1 seriam escolhidas por serem estimativas robustas; nota-se que os valores correspondentes para  $\hat{Y}_T$  são os mais conservadores.

4.- Não existe nenhuma vantagem teórica (relacionada ao fenômeno em si) em descrever a distribuição máxima usando momentos L. Nossos resultados indicam simplesmente que em alguns casos o uso de momentos L permite obter quantidades  $Y_T$  (variável dependente obtida em cada um dos postos) cuja variação entre bacias é melhor explicada pela (variável independente) área da bacia, com uma relação do tipo  $Y_T = a + b A + \epsilon$ , onde  $\epsilon$  é o erro da regressão. Deve-se ressaltar que este é o único aspecto medido pelo intervalo de confiança. Não se pode assegurar que esta vantagem (pelo uso de momentos L em vez de momentos) em termos de intervalo de confiança, permanecerá se a relação for do tipo  $Y_T = a + b A + c B + \dots + \epsilon$  ou  $Y_T = a + b A + c B + \dots + \epsilon$ , onde a, b, c são parâmetros a serem ajustados e A, B ... são características hidrológicas da bacia.

5.- Para os dados utilizados tem-se observado uma marcada associação linear entre  $I_z$  e S.

6.- As vantagens básicas do uso de momentos-L são referidas a uma melhoria na representação de uma amostra aleatória de dados enquanto aos erros de amostragem (erros de medição e erros devido a escassez de registros da vazão em um rio). A repercussão desta vantagem na aplicação do estudo é pequena comparada com o desempenho do uso de momentos comuns. Seria interessante verificar se o incremento de erros de amostragem favoreceria o uso de momentos-L com maior intensidade. Isto poderia ser observado se aos postos aqui utilizados fossem agregados àqueles que possuem mais de 5 anos de registros de vazão.

7.- A questão no sentido de indicar se os valores da variável dependente, usados para ajuste da relação regional, obtidos pela adoção de certa distribuição de probabilidades e pelo uso de momentos L, são menos tendenciosos que os obtidos pelo uso de momentos é mais difícil de ser respondida, já que não se conhece o valor verdadeiro de  $Y_T$  nos postos fluviométricos.

8.- As características de relevo e a influência do mesmo no regime fluvial do rio Canoa, que o diferenciam bastante do resto das bacias na região do estudo, sugerem uma possível divisão da região de interesse em que a bacia do rio Canoa seja tratada como uma sub-região. Em vista do método de regionalização usado no estudo, em que não é necessário supor uma única distribuição para toda a região, a divisão proposta poderia ser equivalente a considerar outras variáveis independentes na regressão. Por exemplo, a declividade do curso principal.

9.- O critério fundamental do uso da regionalização é o de não considerar que tratamos com conjuntos de bacias quase homogêneas, mas com bacias com características físicas diferentes que estão sendo relacionadas para estimar o risco da cheia. Poderão existir bacias dentro da região considerada que provavelmente não correspondam às condições médias que se tenta representar. O caráter preliminar das estimativas obtidas por regionalização em locais sem registros fluviométricos deve estar sempre presente.

10.- Uma dificuldade quanto a definição da relação regional em função da área de drenagem é que dentro da gama de magnitude da área das bacias a disponibilidade de dados intermediários é escassa. Quer dizer, temos bacias com áreas muito grandes ou muito pequenas.

11.- A vantagem do uso de momentos-L, no sentido de que a existência de elementos estranhos na amostra não influencia demasiado nas relações de frequência nos postos individuais, é uma qualidade que favorece seu uso ao contrário de técnicas gráficas.

#### 5.2. Estimativa regional da vazão mínima com T anos de Período de retorno:

As estimativas de vazão mínima regionalizada são denominadas WE/L ou GA/L quando a distribuição Weibull ou Gamma foram adotadas e quando o método de momentos-L fora utilizado. Para as mesmas distribuições, mas para o caso em que o método de momentos fora usado, as estimativas são denotadas WE/m e GA/m.

Embora a largura dos intervalos de confiança das estimativas WE/L e GA/L para as bacias sem registros na região estudada da bacia do rio Uruguai, sejam um pouco menores que as correspondentes as estimativas WE/m e GA/m, sua magnitude é muito grande, e a tentativa de diminuir estas larguras se mostra indispensável.

Uma forma de melhorar o intervalo de confiança das estimativas da vazão mínima é através do uso de outras características hidrológicas (apropriadas) da bacia. Não se deve no entanto, esquecer o sucesso limitado desta prática, pelas dificuldades em se obter um índice que descreva adequadamente a influência da geologia na vazão mínima, o mesmo pode ser dito da evapotranspiração.

O uso de uma relação regional diferente da utilizada neste trabalho também poderia ser tentada. As observações quanto ao ganho na precisão das estimativas da vazão mínima em cada um dos postos pelo uso de momentos L, e as implicações na regionalização são em geral válidas às mencionadas para o caso da vazão máxima.

## 6.- BIBLIOGRAFICA

- BENSON, M.A. (1962), "Evolution of methods for evaluating the occurrence of floods.", U.S. Geol. Surv. Water Supply Pap. 1580-A, 30 pp.
- CHONG, S. e MOORE, S.M. (1983), "Flood frequency analysis for small watersheds in southern Illinois. Wat. Resour. Bull. 19(2), 277-282
- CHOQUETTE, A.F. (1988), "Regionalization of peak discharges for streams in Kentucky.", Water Resources Investigations Report 87-4209, U.S. Geol. Surv., Louisville District, Louisville, Ky.
- CHOW, V.T. (1964), Handbook of Applied Hydrology, Mc Graw-Hill.
- CRUFF, R.W. e RANTZ, S.E. (1965), "A comparison of methods used in flood-frequency studies for coastal basins in California.", U.S. Geol. Surv. Water Supply Pap. 1580-E, 56pp.
- DALRYMPLE, T. (1960), "Flood-frequency analyses.", U.S. Geol. Surv. Water Supply Pap. 1543-A, 80 pp.
- DAMAZIO, J.M. (1984), Estimaco robusta de vazes milenares., Relatrio Tcnico CEPEL 002/84, Rio de Janeiro, 27 pginas.
- DING, J., SONG, D., YANG, R. e HOU, Y. (1989 a), "Expressions relating probability weighted moments to parameters of several distributions inexpressible in inverse form." , J. Hydrol., 110, 259-279.
- DING, J., SONG, D. e YANG, R. (1989 b), "Further research on application of probability weighted moments in estimating parameters of Pearson Type Three distribution.", J. Hydrol., 110, 239-257.
- ELETRORAS, IPH, (1983), Metodologia para regionalizao de vazes, ELETRORAS, DPE, Rio de Janeiro.

- ELETROSUL, CENEC. (1979), Estudo do inventário hidroenergético, bacia hidrográfica do rio Uruguai, ELETROSUL, Florianópolis SC.
- GREENWOOD, J.A, LANDWEHR, J.M., MATALAS, N.C. e WALLIS, J.R. (1979), "Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form.", Water Resour. Res., 15, 1049-1054.
- GREIS, N.P. e WOOD, E.F. (1981), "Regional flood frequency estimation and network design .", Water Resour. Res., 17, 1167-1177. (Correction. Water Resour. Res., 19, 589-590, 1983.)
- HOSKING, J.R.M. (1986), The theory of probability weighted moments, Research Report RC12210, IBM Research, Yorktown Heights, N.Y.
- HOSKING, J.R.M. (1989), Some theoretical results concerning L-moments, Research Report RC 14492, IBM Research, Yorktown Heights, N.Y.
- HOSKING, J.R.M. (1990), "L-Moments: Analysis and estimation of distributions using linear combinations of order statistics.", J. R. Statist. Soc., 52, No. 1.
- HOSKING, J.R.M., e WALLIS, J.R. (1988), "The effect of intersite dependence on regional flood frequency analysis.", Water Resour. Res., 24, 588-600
- HOSKING, J.R.M., WALLIS, J.R. e WOOD E.F. (1985 a), "An appraisal of the regional flood frequency procedure in the UK Flood Studies Report.", Hydrol. Sci. J., 30, 85-109.
- HOSKING, J.R.M., WALLIS, J.R. e WOOD,E.F. (1985 b), "Estimation of the generalized extreme-value distribution by the method of probability-weighted moments.", Technometrics, 27, 251-261.
- HOUGHTON, J.C.(1978). " Birth of a parent: the Wakeby distribution for modeling flood flows.", Water Resour. Res., 14, 1105-1109.

KENDALL, M.G., e STUART, A. (1969), The Advanced Theory of Statistics, vol. 1, Griffin, London.

KITE, G.W (1977), Frequency and Risk Analysis in Hydrology, Water Resources Publication, Fort Collins Co.

LANDWEHR, J.M, MATALAS, N.C., e WALLIS, J.R. (1979), "Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles.", Water Resour. Res., 15, 1055-1064

LANNA, A.E. (1983), "Províncias hidrológicas e bacias representativas.", Anais do V Simposio Brasileira de Recursos Hídricos, Vol.1, 3-23

LANNA, A.E., da SILVEIRA e TUCCI, C.E. (1983). "Regionalização das vazões mínimas na bacia do rio Jacuí, RS.": Anais do V Simpósio Brasileiro de Recursos Hídricos, Vol. 3, 109-131

— LETTENMAIER, D.P., WALLIS J.R., e WOOD, E.F. (1987), "Effect of regional heterogeneity on flood frequency estimation.", Water Resour. Res., 2, 313-323

MATALAS, N.C. (1963), "Probability distribution of low flows.", U.S. Geol. Surv. Professional Pap. 434-A, 27 pp.

MATALAS, N.C., e SLACK, J.R. (1975), "Regional skew in search of a parent.", Water Resour. Res., 6, 815-826.

MATALAS, N.C., e WALLIS, J.R. (1973), "Eureka it fits a Pearson type III distribution.", Water Resour. Res., 9, 281-289

MOSLEY, M.P. (1981). "Delimitation of New Zealand Hydrologic Regions.", J. of Hydrology, 49, 173-192

NAGESHWAR, R.B. (1989). "Comparison of method of residuals and cluster analysis for flood regionalization.", Water Resour. Planning and Management, 115 (6), 793-808

- NERC (1975), Flood Studies Report. Institute of Hydrology Wallingford.
- RAD, C.R. (1973), Linear Statistical Inference and Its Applications, John Wiley & Sons, London. pagina 575.
- RIGGS, H.C. (1973), Regional analyses of streamflow characteristics, U.S. Geol. Surv. Techniques of Water Resources Inv., book 4, chap. b3, 15p.
- SONG, D. e DING, J. (1988), "The application of probability weighted moments in estimating parameters of Pearson Type Three distribution.", J. Hydrol., 101, 47-61.
- STEDINGER, J.R. (1983), "Estimating a regional flood frequency distribution.", Water Resour. Res., 19, 503-510.
- STEDINGER, J.R., e TASKER, G.D. (1985), "Regional hydrologic analysis, 1. Ordinary, Weighted and Generalized Least Squares compared.", Water Resour. Res., 21(9), 1421-1432
- STEDINGER, J.R., e TASKER, G.D. (1986.a), "Correction to regional hydrologic analysis, 1. Ordinary, Weighted and Generalized Least Squares compared.", Water Resour. Res., 22(5), 844
- STEDINGER, J.R., e TASKER, G.D. (1986.b), "Regional hydrologic analysis, 2. Model error estimates, estimation of Sigma and Log Pearson Type 3 distributions.", Water Resour. Res., 22(10), 1487-1489
- TASKER, G.D. (1980), "Hydrologic regression with weighted least squares.", Water Resour. Res., 16(6), 1107-1113
- TASKER, G.D., EYCHANER, J.H., e STEDINGER, J.R. (1986), "Application of Generalized Least Squares in regional hydrologic regression analysis.", U.S. Geological Survey Water-Supply Paper 2310, pp. 107-115

TIAGO de OLIVEIRA, J. (1963), "Decision results for the parameters of the extreme value (Gumbel) distribution based on the mean and the standard deviation.", *Trabajos de Estadística*, 14, 61-81

— VOGEL, R.M. e KROLL, C.N. (1990), "Generalized low-flow frequency relationships for ungaged sites in Massachusetts.", *Wat. Resour. Bull.* 26(2), 241-253

WALLIS, J.R. (1980), "Risk and uncertainties in the evaluation of flood events for the design of hydrologic structures.", paper presented at the Seminar on Extreme Hydrologic Events: Floods and Droughts, Centro di Cult. Sci. "E. Majorana", Erice, Italy

WALLIS, J.R. (1988), "Catastrophes, computing, and Containment: Living With Our Restless Habitat, Research Report RC13406, IBM Research, Yorktown Heights, N.Y.

— WILTSHIRE, S.E. (1986 a), "Regional flood frequency analysis, 1, Homogeneity Statistics." , *Hydrol. Sci. J.*, 31, 321-333.

— WILTSHIRE, S.E. (1986 b), "Regional flood frequency analysis, 11, Multivariate classification of drainage basins in Britain." , *Hydrol. Sci. J.*, 31, 335-346.

ANEXO I  
TABELA DE INTERPRETAÇÃO DOS MOMENTOS-L  
E COMPARAÇÃO DE OUTROS AUTORES COM O MÉTODO DE MÁXIMA VEROSSIMILHANÇA

Tabela 1.1.- Correspondência em quanto a interpretação dos momentos L e momentos convencionais, obtidos de uma amostra.

Momentos L amostrais	Interpretação	Correspondência	Momentos amostrais
$l_1$	localização	= →	Media (m)
$l_2$	dispersão	→	Desvio padrão (S)
$l_3/l_2$	assimetria	→	Coef assimetria (g)
$l_4/l_2$	kurtosis	→	Coef kurtosis (k)

Tabela 1.2.- Desvio padrão das estimativas de quantiles para a distribuição Generalizada de Valores Extremos com parâmetro de forma  $K=0.2$ . (Hosking et al, 1985 b)

n	método	F=P[Y≤y]		
		0.9	0.99	0.999
15	L	0.23	0.32	0.56
	MV	0.23	0.79	>1
25	L	0.18	0.24	0.39
	MV	0.17	0.29	0.86
50	L	0.12	0.16	0.25
	MV	0.12	0.16	0.28

Nota: L denota o método de momentos L. MV significa máxima verossimilhança, o desvio padrão foi obtido de  $\hat{Y}(F)/Y(F)$ .

ANEXO II  
INFORMAÇÃO SOBRE OS POSTOS USADOS NA REGIONALIZAÇÃO



Tabela II.1- As 29 bacias do rio Uruguai utilizadas no estudo

No.	Posto fluviométrico	Rio	Número de anos de registro	Area (Km <sup>2</sup> )
1	Invernada Velha	Pelotas	10	2847
2	Fazenda Mineira	Lava Tudo	22	1108
3	Coxilha Rica	Pelotas	24	497
4	Passo Socorro	Pelotas	31	9009
5	Vila Canoas	Canoas	17	955
6	Ponte do Rio João Paulo	João Paulo	16	564
7	Rio Bonito	Canoas	33	1990
8	Encruzilhada	Canoas	24	2980
9	Ponte Alta do Sul	Canoas	19	4600
10	Ponte do Rio Antinhas	Antinhas	11	41
11	Ponte Marombas	Marombas	16	355
12	Ponte do Rio Correntes	Correntes	11	643
13	Passo Marombas	Marombas	35	3800
14	Passo Caru	Canoas	25	10180
15	Colônia Santana	Canoas	12	13460
16	Passo do Rio Tapejara	Tapejara	18	1110
17	Passo do Ligeiro	Ligeiro	18	537
18	Passo do Granzotto	Forquilha	18	1756
19	Passo São Geraldo	Forquilha	16	1499
20	Passo Santa Tereza	Ligeiro	18	2841
21	Bonito	Irani	24	630
22	Passo Colombelli	Ligeiro	36	3718
23	Videira	do Peixe	22	1103
24	Rio Uruguai	do Peixe	35	5239
25	Marcellino Ramos	Uruguai	35	41267
26	Engenho Velho	Jacutinga	24	938
27	Itá	Uruguai	24	43901
28	Passo Alto Irani	Irani	17	900
29	Passo Caxambú	Uruguai	33	52832

Tabela II.2 - Os momentos L  $l_1$ ,  $l_2$  e as razões de momento  $\hat{\tau}_2 = l_1/l_2$ ,  $\hat{\tau}_4 = l_4/l_2$  (a identificação da sub-bacia é feita através da sua área, mostrada entre parênteses).

(Área):	$l_1$	$l_2$	$\hat{\tau}_2$	$\hat{\tau}_4$
(2847)	806.6	137.42	0.1099	-0.1694
(1108)	444.3	100.06	-0.0907	-0.0017
( 497)	309.1	62.21	-0.0629	-0.0115
(9009)	1960.9	511.61	0.1751	0.1280
( 955)	169.6	17.73	-0.0263	0.2470
( 564)	97.0	10.15	0.0147	0.1808
(1990)	210.5	39.21	0.1125	0.1363
(2980)	322.8	68.51	0.3003	0.1037
(4600)	466.7	97.32	0.1225	0.0219
(1110)	443.6	104.72	0.0783	0.0840
( 41)	8.4	2.35	0.2280	0.2208
( 537)	153.7	31.95	0.0483	0.1286
( 355)	100.2	24.30	0.0154	0.0626
(1756)	332.4	49.94	-0.0088	0.2296
( 643)	76.7	20.60	0.1975	0.1531
(1499)	292.6	44.13	-0.0531	0.1623
(3800)	582.5	135.86	0.1523	0.2293
(10180)	1330.8	333.44	0.1380	0.1554
(13460)	1849.0	521.26	0.2088	0.1564
(2841)	871.3	170.37	-0.0230	0.1538
( 630)	226.1	38.94	0.0541	0.1249
(3718)	987.4	199.56	-0.0366	0.1212
(1103)	480.3	149.72	0.1010	0.0726
(900)	434.2	109.65	0.0629	0.1484
(5239)	1570.3	461.02	0.0790	0.1196
(41267)	7489.0	2096.28	0.1829	0.2555
(938)	414.7	99.40	0.0734	0.1157
(43901)	8398.6	2006.82	0.2591	0.2335
(52832)	9619.8	2673.61	0.1350	0.2098

ANEXO III  
TABELAS DE ESTIMATIVAS DA VAZÃO MÁXIMA COM T ANOS  
DE PERÍODO DE RETORNO, EM BACIAS SEM DADOS DE VAZÃO.  
USO DE MOMENTOS-L E MOMENTOS

Tabela III.1 - Estimativas regionalizadas de descargas diárias médias máximas anuais ( $\hat{Y}_T, m^3 s^{-1}$ ) com períodos de retorno de 50, 100 e 200 anos, para bacias com áreas  $0.5\bar{A}$ ,  $1.0\bar{A}$ , e  $1.5\bar{A}$ , com limites de confiança; regionalização (regressão sobre a área da bacia) da média e desvio padrão, da média e desvio padrão não tendencioso e dos momentos L. Suposta a distribuição Gumbel para  $Y_1$ .

Regionalização de	Estimativa $\hat{Y}_{100}$ e intervalo de confiança ( $m^3/s$ )				
	A=0.5 $\bar{A}$	A=0.75 $\bar{A}$	A=1.0 $\bar{A}$	A=1.25 $\bar{A}$	A=1.5 $\bar{A}$
m, S	1719 ±1472	2568 ±1471	3417 ±1470	4265 ±1471	5114 ±1472
m, S*	1742 ±1469	2600 ±1468	3458 ±1467	4316 ±1468	5174 ±1469
l <sub>1</sub> , l <sub>2</sub>	1767 ±1432	2613 ±1431	3460 ±1431	4306 ±1431	5152 ±1432

Tabela III.2.- Intervalo de confiança médio em  $m^3/s$  das estimativas de descargas máximas anuais ( $\hat{Y}_T$ ). A bacia sem dados é a do posto retirado. L e C indicam o uso de momentos L e momentos respectivamente.

Número de postos retirados	Período de Retorno (anos)								
	50			100			200		
	L	C	Dif.	L	C	Dif.	L	C	Dif.
1	±1303	±1335	32	±1469	±1508	40	±1633	±1683	50
14	±1469	±1509	40	±1651	±1701	50	±1834	±1894	60

Tabela III.3.- Estimativas regionalizadas e estimativas individuais de descargas diárias médias máximas anuais ( $\hat{Y}_T$ , m<sup>3</sup>/s) com período de retorno 100 anos, para bacias escolhidas, com limites de confiança. Suposta a distribuição Gumbel para  $Y_1$ , com parâmetros estimados pelo método de momentos-L (MO-L) e momentos comuns (M-C). Regionalização (regressão de  $Y_{100}$  sobre a área da bacia).

No. para identificação posto	Area (Km <sup>2</sup> )	Estimativas regionalizadas:		Número de anos de registro no posto
		Uso MO-L	Uso M-C	
		$\hat{Y}_{100}$ $\pm 2\sqrt{\text{Var } \hat{Y}_T}$	$\hat{Y}_{100}$ $\pm 2\sqrt{\text{Var } \hat{Y}_T}$	
3	497	289 $\pm 1457$	236 $\pm 1498$	
7	1990	1023 $\pm 1444$	971 $\pm 1488$	
29	52832	24244 $\pm 1851$	24466 $\pm 1928$	
Estimativas individuais:				
3	497	670 $\pm 170$	639 $\pm 196$	24
7	1990	438 $\pm 91$	423 $\pm 104$	33
29	52832	25137 $\pm 6651$	24906 $\pm 8091$	29

Tabela III.4- Regressões de: momentos L, razão de momentos  $l_2/l_2$  e  $l_4/l_2$  sobre a área da sub-bacia. Também são mostradas regressões do desvio padrão, coeficientes de assimetria e kurtosis sobre a área das sub-bacias. Ajuste da regressão com 29 sub-bacias.

Variável dependente:	Declividade:	Intercepto:	Variância residual:	$r^2$ :
$l_1$	0.1814	72.8470	76535	0.988
$l_2$	0.0488	0.3442	6739	0.986
$l_3$	0.0091	-10.0319	1252	0.930
$l_4$	0.0112	-13.7795	740	0.971
$\hat{\tau}_3$	0.(5x0)30	0.0656	0.008809	0.173
$\hat{\tau}_4$	0.(5x0)25	0.1119	0.007309	0.141
SD	0.0907	-16.7000	25629	0.985
Assime.	0.(4x0)25	0.3093	0.129576	0.492
Curtoz.	0.(4x0)78	2.6005	0.751265	0.617

Tabela III.5- Estimativas regionalizadas de descargas diárias médias máximas anuais ( $\hat{Y}_T, m s^{-1}$ ) com períodos de retorno de 50, 100 e 200 anos, para bacias com áreas 0.5 $\bar{A}$ , 1.0 $\bar{A}$ , e 1.5 $\bar{A}$ , com limites de confiança; regionalização usando momentos L e momentos comuns. (Suposta a distribuição Gumbel para  $Y_1$ )

		Regionalização momentos L:		
		Período de retorno (anos)		
		50	100	200
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	1588	1767	1946
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1271$	$\pm 1432$	$\pm 1594$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	3101	3460	3817
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1269$	$\pm 1431$	$\pm 1592$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	4614	5152	5687
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1271$	$\pm 1432$	$\pm 1594$
		Regionalização momentos:		
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	1548	1719	1890
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1302$	$\pm 1472$	$\pm 1643$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	3066	3417	3766
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1301$	$\pm 1470$	$\pm 1641$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	4583	5114	5642
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1302$	$\pm 1472$	$\pm 1642$

Tabela III.6- Estimativas regionalizadas de descargas diárias médias máximas anuais ( $\hat{Y}_T, m^3 s^{-1}$ ) com períodos de retorno de 50, 100 e 200 anos, para bacias com áreas  $0.5\bar{A}$ ,  $1.0\bar{A}$ , e  $1.5\bar{A}$ , com limites de confiança; regionalização (regressão ponderada sobre a área da bacia) de  $Y_{50}$ ,  $Y_{100}$  e  $Y_{200}$  (supondo a distribuição Gumbel para  $Y_1$ , com parâmetros estimados pelo método de momentos L e momentos comuns).

Uso de momentos L:

		Período de retorno (anos)		
		50	100	200
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	1635	1821	2005
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 6141$	$\pm 6943$	$\pm 7749$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	3156	3522	3887
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 6141$	$\pm 6942$	$\pm 7748$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	4678	5224	5769
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 6141$	$\pm 6942$	$\pm 7748$

Uso de momentos:

A=0.5 $\bar{A}$ :	$\hat{Y}_T$	1597	1774	1950
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 6192$	$\pm 7019$	$\pm 7850$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	3124	3484	3842
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 6192$	$\pm 7018$	$\pm 7850$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	4652	5194	5733
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 6192$	$\pm 7018$	$\pm 7850$

Tabela III.7.- Regressões lineares de  $\hat{Y}_T$  sobre a área da bacia, os parâmetros da distribuição Gamma em cada posto foram estimados a partir dos momentos L. Ajuste da regressão com 29 sub-bacias

	Período de retorno (Anos)		
	50	100	200
Intecepto	60.82	55.76	50.75
Declividade	0.4074	0.4499	0.4911
Variância residual	391453	494854	610313
$r^2 \times 100$	98.8	98.7	98.7

Tabela III.8.- Regressões lineares de  $\hat{Y}_T$  sobre a área da bacia, os parâmetros da distribuição Gamma em cada posto foram estimados a partir dos momentos. Ajuste da regressão com 29 sub-bacias.

	Período de retorno (Anos)		
	50	100	200
Intercepto	15.86	0.62	-14.78
Declividade	0.4122	0.4558	0.4982
Variância residual	417163	531743	660303
$r^2 \times 100$	98.8	98.7	98.6

Tabela III.9 - Estimativas regionalizadas de descargas diárias médias máximas anuais ( $\hat{Y}_T$ ,  $m^3 s^{-1}$ ) com períodos de retorno de 50, 100 e 200 anos, para bacias com áreas  $0.5\bar{A}$ ,  $1.0\bar{A}$ , e  $1.5\bar{A}$ , com limites de confiança; regionalização (regressão sobre a área da bacia) de  $Y_{50}$ ,  $Y_{100}$  e  $Y_{200}$  supondo a distribuição Gamma para  $Y_t$ , com parâmetros estimados pelo método de momentos L e momentos.

Uso de momentos L:

		Período de retorno (anos)		
		50	100	200
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	1545	1695	1839
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 1274$	$\pm 1433$	$\pm 1591$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	3029	3334	3629
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 1273$	$\pm 1431$	$\pm 1589$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	4513	4973	5418
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 1274$	$\pm 1433$	$\pm 1591$
Uso de momentos:				
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	1517	1661	1800
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 1315$	$\pm 1485$	$\pm 1655$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	3019	3322	3615
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 1314$	$\pm 1483$	$\pm 1653$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	4521	4982	5430
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 1315$	$\pm 1485$	$\pm 1655$

Tabela III.10.- Regressões lineares de  $\hat{Y}_T$  sobre a área da bacia, os parâmetros da distribuição GEV em cada posto foram estimados a partir dos momentos L. Ajuste da regressão com 29 sub-bacias.

	Período de retorno (Anos)		
	50	100	200
Intercepto	-0.0587	-40.77	-87.41
Declividade	0.4221	0.4766	0.5326
Variância residual	368761	546113	866153
$r^2 \times 100$	98.9	98.8	98.4

Tabela III.11.- Regressões lineares de  $\hat{Y}_T$  sobre a área da bacia, os parâmetros da distribuição GEV em cada posto foram estimados a partir dos momentos. Ajuste da regressão com 29 sub-bacias.

	Período de retorno (Anos)		
	50	100	200
Intercepto	-42.26	-103.15	-172.92
Declividade	0.4223	0.4780	0.5340
Variância residual	423438	593585	838001
$r^2 \times 100$	98.8	98.7	98.5

Tabela III.12 - Estimativas regionalizadas de descargas diárias médias máximas anuais ( $\hat{Y}_T, m^3 s^{-1}$ ) com períodos de retorno de 50, 100 e 200 anos, para bacias com áreas  $0.5\bar{A}$ ,  $1.0\bar{A}$ , e  $1.5\bar{A}$ , com limites de confiança; regionalização (regressão sobre a área da bacia) de  $Y_{50}$ ,  $Y_{100}$  e  $Y_{200}$  supondo a distribuição GEV para  $Y_t$ , com parâmetros estimados pelo método de momentos L e momentos.

		Uso de momentos L:		
		Período de retorno (anos)		
		50	100	200
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	1538	1695	1853
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1237$	$\pm 1505$	$\pm 1895$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	3075	3432	3793
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1235$	$\pm 1503$	$\pm 1893$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	4613	5168	5733
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1237$	$\pm 1505$	$\pm 1893$
		uso de momentos:		
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	1496	1638	1772
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1325$	$\pm 1569$	$\pm 1864$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	3035	3380	3718
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1324$	$\pm 1567$	$\pm 1862$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	4573	5121	5663
	$\pm 2\sqrt{(\text{var } \hat{Y}_T)}$	$\pm 1325$	$\pm 1569$	$\pm 1864$

Tabela III.13.- Estimativas regionalizadas de descargas diárias médias mínimas anuais ( $\hat{Y}_T$ ,  $m^3 s^{-1}$ ) com períodos de retorno de 50, 100 e 200 anos, para bacias com áreas  $0.5\bar{A}$ ,  $1.0\bar{A}$ , e  $1.5\bar{A}$ , com limites de confiança: regionalização (regressão sobre a área da bacia) de  $Y_{50}$ ,  $Y_{100}$  e  $Y_{200}$  supondo a distribuição Weibull para  $Y_z$ , com parâmetros estimados pelo método de momentos L e momentos.

Uso de momentos L:

		Período de retorno (anos)		
		50	100	200
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	3.9	3.2	2.7
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 9.2$	$\pm 8.4$	$\pm 7.7$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	7.5	6.1	5.0
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 9.2$	$\pm 8.4$	$\pm 7.7$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	11.1	9.0	7.3
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 9.2$	$\pm 8.4$	$\pm 7.7$

Uso de momentos:

A=0.5 $\bar{A}$ :	$\hat{Y}_T$	4.1	3.4	2.8
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.0$	$\pm 9.3$	$\pm 8.5$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	7.7	6.3	5.2
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.0$	$\pm 9.3$	$\pm 8.5$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	11.4	9.2	7.5
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.0$	$\pm 9.3$	$\pm 8.5$

Tabela III.14.- Estimativas regionalizadas de descargas diárias médias mínimas anuais ( $\hat{Y}_T, m^3 s^{-1}$ ) com períodos de retorno de 50, 100 e 200 anos, para bacias com áreas  $0.5\bar{A}$ ,  $1.0\bar{A}$ , e  $1.5\bar{A}$ , com limites de confiança; regionalização (regressão sobre a área da bacia) de  $Y_{50}$ ,  $Y_{100}$  e  $Y_{200}$  supondo a distribuição Gamma para  $Y_2$ , com parâmetros estimados pelo método de momentos L e momentos.

Uso de momentos L:

		Período de retorno (anos)		
		50	100	200
A=0.5 $\bar{A}$ :	$\hat{Y}_T$	4.9	4.3	3.9
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.2$	$\pm 9.9$	$\pm 9.7$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	9.4	8.3	7.4
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.2$	$\pm 9.9$	$\pm 9.7$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	14.0	12.3	10.9
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.2$	$\pm 9.9$	$\pm 9.7$

Uso de momentos:

A=0.5 $\bar{A}$ :	$\hat{Y}_T$	5.1	4.5	4.1
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.8$	$\pm 10.6$	$\pm 10.4$
A=1.0 $\bar{A}$ :	$\hat{Y}_T$	9.9	8.7	7.8
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.8$	$\pm 10.6$	$\pm 10.4$
A=1.5 $\bar{A}$ :	$\hat{Y}_T$	14.6	12.9	11.5
	$\pm 2\sqrt{\text{var } \hat{Y}_T}$	$\pm 10.8$	$\pm 10.6$	$\pm 10.4$

**ANEXO IV**  
**REGRESSÃO PONDERADA**

#### Anexo IV.- Regressão ponderada

Se entre as variáveis independentes usadas na análise de regressão umas forem menos confiáveis que outras a equação de regressão simples é:

$$Y_i = a + b A_i + v_i \quad \text{para } i=1, 2, \dots, N$$

onde  $a$  e  $b$  são parâmetros a determinar,  $A_i$  é a característica da bacia na estação  $i$ ,  $Y_i$  é a estimativa da característica da descarga na estação  $i$ ,  $N$  é o número de estações,  $v_i$  é o erro e é considerada uma variável aleatória com média zero e variancia  $\sigma^2(v_i)$  proporcional a variancia de  $Y_i$ . Assumindo que  $Y_i$  é independente, a estimativa dos parâmetros  $a$  e  $b$  obtidos pela minimização do somatório ao quadrado dos erros ponderados

$$S = \sum_{i=1}^N w_i (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^N w_i v_i^2 \quad \text{IV.1}$$

é:

$$\hat{a} = \frac{\sum_{i=1}^N w_i (Y_i - \bar{Y})(A_i - \bar{A})}{\sum_{i=1}^N w_i (A_i - \bar{A})^2}$$
$$\hat{b} = \bar{Y} - \hat{a} \bar{A}$$

onde

$$\bar{Y} = \frac{\sum_{i=1}^N w_i Y_i}{\sum_{i=1}^N w_i}$$

$$\bar{A} = \frac{\sum_{i=1}^N w_i A_i}{\sum_{i=1}^N w_i}$$

$w_i$  é o fator de ponderação.

A estimativa da variancia residual da regressão é:

$$s^2 = \frac{\sum w_i (Y_i - \bar{Y})(A_i - \bar{A})^2 - [\sum w_i (A_i - \bar{A})(Y_i - \bar{Y})]^2}{\sum w_i (A_i - \bar{A})^2}$$

N-K

onde  $K$  é o número de parâmetros a serem estimados na regressão. Os índices dos somatórios foram suprimidos, entende-se que estes são de  $i=1$  até  $N$ .

#### Intervalo de confiança de $\hat{Y}$

Caso o peso  $w_i$  seja igual ao número de anos ( $n_i$ ) com que  $Y_i$  é estimado, a variancia do valor de  $\hat{Y}$  estimado pela regressão a partir de certo valor  $A^*$  é:

$$\text{Var}(\hat{Y}) = \hat{S}^2 \left( \frac{1}{n^*} + \frac{1}{\sum_{i=1}^N n_i} + \frac{(A^* - \bar{A})^2}{\sum_{i=1}^N n_i (A_i - \bar{A})^2} \right) \quad \text{IV.2}$$

No capítulo 4 assumimos que a estimativa  $\hat{Y}_T$  a partir de características físicas da bacia resulta em uma  $\text{var}[\hat{Y}_T]$  semelhante a que seria obtida para  $n^*=1$  ano. Uma forma de verificar esta adoção consiste em obter o número de anos ( $n$ ) da amostra aleatória que seria necessário para obter uma estimativa cuja precisão em termos de variancia é igual a precisão representada pelo valor da variancia obtida com a expressão IV.2. Como, a variancia da estimativa do evento com  $T$  anos de período de retorno obtida a partir de uma amostra proveniente da distribuição Gumbel onde os momentos convencionais são utilizados é dada por:

$$\text{Var}(\hat{Y}_T) = \frac{\sigma^2}{n} (1 + 1.14 k + 1.10 k^2)$$

onde  $\sigma^2$  será substituído pela estimativa amostral da variancia ( $S^2$ ),  $k$  é o fator de frequência proposto por Chow (1964). Então:

$$n = \frac{\sigma^2}{\text{var}(\hat{Y}_T)} (1 + 1.14 k + 1.10 k^2) \quad \text{IV.3}$$

Exemplo.- Para a bacia do rio Uruguai o valor medio regional do erro padrão da vazão máxima é  $\bar{S}=836 \text{ m}^3/\text{s}$ . Adotada a distribuição Gumbel junto ao uso dos momentos convencionais a vazão máxima com período de retorno  $T=100$  anos (consequentemente  $k=3.17$ ) estimada a partir da área da bacia  $A^* = 0.5 \bar{A}$  através da regressão ponderada tem precisão dada por  $[\text{var}(\hat{Y}_T)]^{1/2}=736 \text{ m}^3/\text{s}$ . Nestas circunstâncias o número de anos da amostra será proporcionado pelo uso da expressão IV.3. Onde obtemos  $n = 0.9$  anos que não invalida a adoção de  $n^*=1$  ano.

Uma expressão análoga a IV.3 pode ser conseguida para o caso em que os momentos  $L$  sejam utilizados, no entanto a sua validade é restrita para casos em que  $n > 20$  anos já que a  $\text{var}(l_2)$  e  $\text{cov}(l_1, l_2)$  necessários no desenvolvimento podem ser obtidos segundo Hosking (1986) apenas com base na teoria assintótica para combinações lineares cujos resultados (no caso da distribuição Gumbel) apresentam-se adequados para  $n > 20$  anos.

ANEXO V  
VARIANÇA DA ESTIMATIVA REGIONAL  $\hat{Y}_T$ ,  
CASO EM QUE A DISTRIBUIÇÃO GAMMA É ADOTADA E  
O MÉTODO DE MOMENTOS L É UTILIZADO

Anexo V. - Variância da estimativa regional, caso em que a distribuição Gamma é adotada e o método de momentos L é utilizado.

A estimativa da  $\text{var}(\hat{Y}_T)$  de acordo com Kendall (1969) e para o caso em que  $\hat{Y}_T = f(l_1, l_2)$  pode ser calculada aproximadamente por:

$$\text{Var}(\hat{Y}_T) = \left(\frac{\partial Y_T}{\partial l_1}\right)^2 \text{var}(l_1) + \left(\frac{\partial Y_T}{\partial l_2}\right)^2 \text{var}(l_2) + 2\left(\frac{\partial Y_T}{\partial l_1}\right)\left(\frac{\partial Y_T}{\partial l_2}\right)\text{cov}(l_1, l_2) \quad \text{V.1}$$

como o método de regionalização que se deseja adotar prevê o estabelecimento de equações de regressão de  $l_1$  e  $l_2$  sobre características físicas e climáticas das bacias, os termos  $\text{var}(l_1)$ ,  $\text{var}(l_2)$  e  $\text{cov}(l_1, l_2)$  serão obtidos dessas regressões. Tendo em vista de que a função de distribuição Gamma não permite uma representação imediata do quantil  $Y_T$ , na forma  $\hat{Y}_T = f(l_1, l_2)$  para o cálculo das derivadas parciais acima requeridas, recorreremos a um procedimento pelo qual obtemos  $\hat{Y}_T \cong f(l_1, l_2)$ . Assim, para a função de distribuição Gamma

$$F(Y) = \frac{1}{\Gamma(\alpha)} \int_0^{Y_0} t^{\alpha-1} e^{-t} dt \quad Y_0 = \frac{Y_T}{\beta}$$

Abramowitz e Stegun (1965) mostram que

$$F(Y) = F(\kappa^2/v)$$

onde  $F(\kappa^2/v)$  é a distribuição chi-quadrado com  $v=2\alpha$  graus de liberdade e  $\kappa^2=2Y$ . Portanto, para uma probabilidade  $(1-1/T)$  e  $2\alpha$  graus de liberdade o evento com magnitude  $Y_0$  é obtido como:

$$Y_0 = \kappa^2/2$$

e a magnitude da cheia esperada com probabilidade de excedência  $1/T$  é dada por:

$$Y_T = \frac{\beta \kappa^2}{2} \quad \text{V.2}$$

a variável  $\kappa^2$  pode ser obtida da expressão:

$$u = \left( \left(\frac{\kappa^2}{v}\right)^{1/3} + \frac{2}{9v} - 1 \right) \left( \frac{9v}{2} \right)^{1/2}$$

apresentada por Kendall (1969), que mostra que para  $v > 30$ ,  $u$  tem aproximadamente distribuição Normal, com média zero e variância unitária. Assim o valor de  $\kappa^2$  para certo nível de probabilidade e número de graus de liberdade  $v$ , é aproximado por:

$$\kappa^2 \cong v \left( 1 - \frac{2}{9v} + u \sqrt{\frac{2}{9v}} \right)^3 \quad \text{V.3}$$

levando V.3 em V.2, depois de substituir o número de graus de liberdade  $v=2\alpha$ , temos:

$$Y_T \cong \beta \alpha \left( 1 - \frac{1}{9\alpha} + u \sqrt{\frac{1}{9\alpha}} \right)^3$$

que em termos de momentos-L e como  $\beta = \lambda_1/\alpha$  é:

$$Y_T \cong \lambda_1 \left( 1 - \frac{1}{9\alpha} + u \sqrt{\frac{1}{9\alpha}} \right)^3 \quad V.4$$

Neste ponto deve ressaltar-se que a relação entre o parâmetro  $\alpha$  e os momentos  $\lambda_1$  e  $\lambda_2$  não é explícita tal como mostrado na equação 4.14, no entanto, a aproximação:

$$\hat{\alpha} \cong (1 - 0.3080 Z) / (Z - 0.05812 Z^2 + 0.01765 Z^3),$$

onde,  $Z = \pi t^2$  e  $t = \lambda_2/\lambda_1$ , quando  $0 < t < 1/2$ .

Se,  $1/2 \leq t < 1$  então  $Z = 1 - t$  e

$$\hat{\alpha} \cong (0.7213 Z - 0.5947 Z^2) / (1 - 2.1817 Z + 1.2113 Z^2),$$

$$\hat{\beta} = \lambda_1 / \hat{\alpha}$$

proposta por Hosking (1989), já em termos das estimativas dos parâmetros da distribuição Gamma, poderia ser usada. Contudo é fácil perceber a extensão algébrica que resultaria do cálculo das derivadas parciais  $\frac{\partial Y_T}{\partial \lambda_1}$ ,  $\frac{\partial Y_T}{\partial \lambda_2}$  requerida na equação V.1 para estimativa da  $\text{var}(\hat{Y}_T)$ . Estes inconvenientes não são encontrados quando a distribuição Gamma e os momentos convencionais são usados, em vista da representação simples de  $Y_T$  função da média ( $m$ ) e desvio padrão ( $S$ ).

ANEXO VI

CARATERÍSTICAS DESEJAVEIS DE UM "BOM" ESTIMADOR

## Anexo VI.- Características desejáveis de um "bom" estimador.

Quando conhecemos ou assumimos como hipóteses de trabalho, que certa variável aleatória  $Y$  se comporta de acordo com certa distribuição de probabilidades cujos parâmetros  $\theta$  não são conhecidos mais podem ser estimados a partir de uma amostra  $y_1, y_2, \dots, y_n$ , utilizamos na estimativa funções dos valores da amostra (estatísticas) que são consideradas "bons" estimadores. Deve-se ressaltar que a estimativa é uma variável aleatória, conseqüentemente, a formulação de um bom estimador é no sentido de que este proporcione bons resultados em média. Assim, as características desejáveis mais utilizadas para se ter um bom estimador são:

### Não tendenciosidade.-

Se para qualquer  $n$  e  $\theta$  desejamos que o valor esperado de  $\hat{\theta}$  seja igual a  $\theta$ , definimos o estimador não tendencioso como:

$$E[\hat{\theta}] = \theta$$

### Eficiência.-

Seja  $\hat{\theta}^*$  um estimador de máxima verosimilhança de  $\theta$ . Seja  $\hat{\theta}$  um outro estimador de  $\theta$ . Sendo ambos estimadores consistentes com distribuição assintoticamente Normal. Uma medida da eficiência relativa de  $\hat{\theta}$  é dada por:

$$E = \frac{\text{Var}[\hat{\theta}^*]}{\text{Var}[\hat{\theta}]}$$

### Consistência.-

O estimador  $\hat{\theta}_n$  calculado a partir de uma amostra com  $n$  valores é denominado estimador consistente de  $\theta$  se para dado  $\epsilon$  positivo e pequeno verifica-se que:

$$\lim_{n \rightarrow \infty} P[|\hat{\theta}_n - \theta| < \epsilon] = 1$$

Isto é,  $\hat{\theta}_n$  é um estimador consistente de  $\theta$  se uma sucessão de probabilidades converge a 1 a medida que o tamanho da amostra aumenta.

Uma forma prática de ver a consistência e que pode ser obtida da definição anterior é:

$$\text{Se, } \lim_{n \rightarrow \infty} E[\hat{\theta}_n] = \theta \quad \text{e} \quad \lim_{n \rightarrow \infty} \text{Var}[\hat{\theta}_n] = 0$$

então  $\hat{\theta}_n$  é um estimador consistente de  $\theta$ .