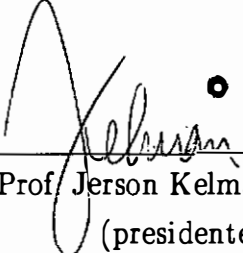


INCORPORAÇÃO DA INCERTEZA PARAMÉTRICA NA GERAÇÃO DE
SÉRIES SINTÉTICAS DE VAZÕES ATRAVÉS DE REAMOSTRAGEM

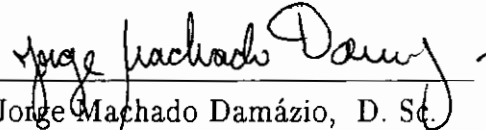
Rubem Luiz Darú

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO EM ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS
EM ENGENHARIA CIVIL.

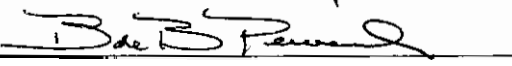
Aprovada por:



Prof. Jerson Kelman, Ph. D.
(presidente)



Jorge Machado Damázio, D. Sc.



Basílio Bragança Pereira, Ph. D.

DARÚ, RUBEM LUIZ

Incorporação da Incerteza Paramétrica na Geração de Séries Sintéticas de Vazões através de Reamostragem (Rio de Janeiro) 1992.

VI, 47 p. 29,7 cm (COPPE/UFRJ. M.Sc., Engenharia Civil, 1992)

Tese — Universidade Federal do Rio de Janeiro, COPPE.

1. Hidrologia Estocástica

2. Consideração da incerteza dos parâmetros

3. Estudos de regularização

I. COPPE/UFRJ

II. Título (série)

Agradecimentos

Jerson Kelman e Jorge Machado Damázio pelas explicações, dedicação, revisões de texto e inteligência que muito colaboraram para conversão deste em tese de mestrado

Fernanda da Serra Costa e Maria Elvira Piñeiro Maceira que participaram de algumas discussões sobre o trabalho.

Ao CNPq e ao CEPTEL por fornecer ajuda financeira e a infra-estrutura necessária para a elaboração desta tese.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para obtenção do grau de Mestre em Ciências (M.Sc.).

INCORPORAÇÃO DA INCERTEZA PARAMÉTRICA NA GERAÇÃO DE SÉRIES SINTÉTICAS DE VAZÕES ATRAVÉS DE REAMOSTRAGEM

Rubem Luiz Darú

Março, 1992

Orientador: Prof. Jerson Kelman

Programa: Engenharia Civil

Modelos de geração de vazões são muito utilizados em estudos de simulação de reservatórios de regularização de rios. No entanto, como os registros de vazões disponíveis são sempre curtos, existe sempre uma considerável incerteza na estimação dos parâmetros destes modelos. Os métodos Bayesiano e Clássico para consideração desta incerteza apresentam forte dependência da estrutura do modelo adotado.

Nesta tese desenvolve-se um procedimento geral para incorporar nos modelos de geração de vazões a incerteza dos parâmetros, procurando-se evitar os erros de modelagem. O procedimento se baseia num esquema de Reamostragem (semelhante ao Bootstrap) capaz de ser aplicado em séries que apresentam dependência temporal, como é o caso das séries de vazões naturais. Da série histórica se obtêm várias séries pseudo-históricas pela reamostragem de "blocos de observações" que por hipótese são independentes. Destas séries pseudo-históricas obtêm-se novos possíveis valores para os parâmetros.

O procedimento foi detalhado para aplicação num modelo de geração de vazões mensais do tipo AR(1) anual com desagregação pelo esquema de Meija e Rousselle (1976). Uma aplicação em estudos de regularização para o reservatório de Furnas no rio Grande, indicou que a consideração da incerteza dos parâmetros tem um efeito da mesma ordem de grandeza que o efeito obtido quando o intervalo de tempo da análise passa de anual para mensal.

Abstract of Thesis presented to COPPE/UFRJ as partial fulfillment of the requirements for the degree of Master of Science (M.Sc.).

INCORPORATION OF PARAMETER UNCERTAINTY IN SYNTHETIC STREAMFLOW SERIES GENERATION USING RESAMPLING

Rubem Luiz Darú

March, 1992

Thesis Supervisors: Jerson Kelman

Department: Civil Engineering

Streamflow generation models are very useful in simulation studies of river regulation. However there is always an uncertainty in their parameters due to fact that time period of the streamflow records are always short. The Bayesian and Clássic methods for consideration of parameter uncertainty show many dependence on the structure of the adopted model.

This thesis develops a general procedure for incorporation parameters uncertainty into streamflow generation models, avoiding modeling errors. The procedure is based on a Resampling scheme (similar to Bootstrap), that can be used in series that show time dependence, as the natural streamflow series. Based on historical series several "Bootstrap samples" are obtained by resampling "blocks of observations" which are independent by hypothesis. From these "Bootstrap samples" others possible parameters values are obtained.

The procedure is applied to a monthly streamflow generation model given by an annual (AR-1) process which is disaggregated into monthly values by the scheme of Meija e Rousselle (1976). An application in regulation studies for Furnas Reservoir in the Grande river, showed that the consideration of the uncertainty in the parameters model has an effect with the same magnitude that the effect obtained when the analisys' time interval is changed from annual to monthly.

SUMÁRIO

CAPÍTULO 1	INTRODUÇÃO	1
CAPÍTULO 2	DESCRIÇÃO DO MODELO DE SIMULAÇÃO HIDROLÓG.	6
2.1	Modelagem anual AR(1)	6
2.2	Esquema de desagregação mensal	9
CAPÍTULO 3	IMPLEMENTAÇÃO DA INCERTEZA NOS PARÂMETROS NO MODELO M.S.H UNIVARIADO	14
3.1	Bootstrap	14
3.2	Desenvolvimento de um esquema de Bootstrap em séries com dependência temporal	15
3.3	Utilização do esquema de reamostragem para implementação no M.S.H.	18
CAPÍTULO 4	ESTUDOS DE REGULARIZAÇÃO CONSIDERANDO A INCERTEZA DOS PARÂMETROS	21
4.1	Introdução	21
4.2	Descrição dos estudos	22
	Tabelas	24
4.3	Análise dos resultados	29
	Gráficos	30
CAPÍTULO 5	CONCLUSÃO	43
	BIBLIOGRAFIA	46

CAPÍTULO 1

INTRODUÇÃO

O uso da Hidrologia Sintética no planejamento de Sistemas de Recursos Hídricos têm origem na inadequação de análises baseadas simplesmente nos registros históricos de vazões. Isto porque, em geral, os registros não ultrapassam 60 anos de comprimento e portanto não exibem grande parte da gama de sequência de vazões possíveis de ocorrer.

Provavelmente a primeira aplicação de Hidrologia Sintética é o trabalho de Sudler (1927), que usando um baralho de 50 cartas criou novas sequências de vazões anuais a partir de um registro de 50 anos. O método tinha a desvantagem de ignorar a dependência temporal e de trabalhar sempre com apenas as vazões registradas, mudando-se somente a sua ordem, conseqüentemente os máximos e mínimos e os momentos das novas séries eram sempre iguais aos valores históricos. Desde então, com o desenvolvimento de vários modelos de geração de vazões incorporando diversas estruturas de correlação temporal e geração de números aleatórios, de técnicas de estimação de parâmetros e com o advento dos computadores, o uso de séries sintéticas foi grandemente expandido, tornando-se um instrumento importante em estudos de regularização de rios.

A metodologia geralmente utilizada é descrita basicamente por:

1 – A partir da análise da série histórica de vazões medidas na seção de interesse escolhe-se um modelo estocástico e estima-se seus respectivos parâmetros.

2 – A partir do modelo escolhido e ajustado na etapa anterior e de um gerador de ruídos aleatórios, obtem-se uma longa série sintética (≥ 1000 anos) de vazões. Admite-se que esta longa série sintética contenha toda a gama de sequência de vazões possíveis de ocorrer na seção estudada, assim como suas respectivas probabilidades de ocorrência. Isto porque considera-se que o gerador usado pela natureza comporta-se exatamente como o gerador definido pelo modelo.

3 – Pode-se então simular a operação do reservatório proposto por um longo período, submetendo-o às aflúências da série sintética, e calcular qualquer tipo de índice de performance, tal como o risco de não atendimento de uma certa demanda (risco de déficit).

Como se vê a geração de séries sintéticas envolve duas etapas preliminares, a escolha do modelo que descreve o fenômeno em estudo e a estimação dos seus parâmetros. A estimação de parâmetros têm sido objeto de preocupação em diversos estudos devido aos erros amostrais sempre presentes, pois as séries históricas têm sempre período limitado. Burges e Lettenmaier (1981) mostraram que o efeito da incerteza dos parâmetros é mais importante que a escolha do modelo. Stedinger e Taylor (1982) e Loucks et al(1981) mostraram a extensão dos erros amostrais nos valores dos parâmetros, e seus efeitos.

A consequência básica dos erros amostrais é que o gerador natural sempre difere do gerador definido pelo modelo. Logo a série sintética obtida poderá não conter toda a gama de sequência de vazões possíveis de ocorrer na seção estudada, e mais que isso, não conterá as verdadeiras probabilidades de ocorrência. Uma maneira de suplantar o problema consiste em:

1 – Construir a partir do histórico de vazões medidas na seção de interesse e de outras informações relevantes uma distribuição de probabilidades para os parâmetros do modelo.

2 – Através desta distribuição gerar por sorteio aleatório um possível vetor θ_i de parâmetros.

3 – Através do vetor de parâmetros θ_i e do modelo, gerar uma longa série sintética.

4 – Repete-se o item 2 e 3, um certo número de vezes obtendo-se a cada repetição uma série sintética. Com cada série sintética calcula-se por simulação um índice de performance (ex: risco de déficit). O conjunto de índices obtidos define a distribuição de probabilidades para este índice que considera a incerteza dos parâmetros.

As diversas metodologias disponíveis para a consideração da incerteza dos parâmetros diferem entre si pela maneira de obter a distribuição de probabilidades dos parâmetros. A formulação melhor fundamentada consiste em se adotar o ponto de vista Bayesiano considerando-se qualquer parâmetro para o qual não se tem certeza do seu valor como uma variável aleatória cuja distribuição de probabilidades é dada pelo produto entre a chamada distribuição a priori e a função Verossimilhança. A distribuição a priori fornece as probabilidades que o

hidrólogo, sem o conhecimento do registro histórico de vazões no local, atribuiria a cada possível valor para o parâmetro, enquanto que a função Verossimilhança fornece para os diferentes possíveis valores dos parâmetros as probabilidades de ocorrência, na modelagem adotada, da particular sequência de vazões correspondentes ao registro histórico. É sempre possível obter-se distribuições Bayesianas baseadas apenas nas informações do registro adotando-se distribuições a priori que fornecem a mesma probabilidade para qualquer possível valor para os parâmetros. O uso da função Verossimilhança faz com que as distribuições Bayesianas dependam fortemente da estrutura do modelo adotado.

Uma aproximação muito utilizada é o uso de distribuições estatísticas das estimativas de parâmetros (método Clássico). Esta abordagem tem o defeito de não incorporar a noção de distribuições a priori. Além disso, as distribuições estatísticas dependem não só da estrutura do modelo adotado como também dos valores de população dos parâmetros que são, por hipótese básica, desconhecidos. A abordagem usual para suplantar este problema é desprezar momentaneamente esta incerteza e adotar como valores de população as estimativas obtidas com a série histórica. Isto faz com que as distribuições estatísticas subestimem a incerteza nos parâmetros apresentando em consequência menor variância do que as correspondentes distribuições bayesianas baseadas apenas nas informações do registro.

Como vimos, tanto o método Bayesiano como o método Clássico têm como defeito o fato de que a distribuição de probabilidades dos parâmetros fornecida depende não só das vazões registradas mas também do modelo adotado. Outro problema comum a estes dois métodos é a dificuldade de se derivar resultados analíticos para modelos não Gaussianos e/ou com complexas estruturas de dependência temporal. Stedinger (1980) mostra que as distribuições de estimativas dos parâmetros podem ser sempre obtidas empiricamente, simplesmente gerando-se com o modelo ajustado à série histórica várias séries sintéticas com o mesmo comprimento da série histórica e aplicando-se o método de estimação a cada série gerada. A distribuição empírica destas estimativas aproxima-se da distribuição analítica sendo sempre possível diminuir o erro de aproximação pelo aumento do número de séries sintéticas. Stedinger (1980) apresenta ainda uma formulação geral para se proceder uma correção das estimativas obtidas em cada série sintética aumentando a sua variabilidade de forma a aproximar melhor a distribuição empírica resultante da distribuição Bayesiana dos parâmetros.

Uma terceira alternativa é o uso das técnicas de Reamostragem

inicialmente proposta por Quenouille (1949) e desenvolvida por Efron (1979, 1982). Em geral, as técnicas de Reamostragem consistem em esquemas de, a partir de uma amostra suposta formada por sorteios independentes de uma mesma população (desconhecida), obter-se por reamostragem destes sorteios, amostras derivadas chamadas de pseudo-amostras. Aplicando-se o método de estimação dos parâmetros a cada pseudo-amostra obtém-se um conjunto de estimativas para os parâmetros cuja distribuição empírica é considerada como distribuição dos parâmetros.

Analizando-se estes esquemas, pode-se concluir que os métodos de reamostragem são métodos similares ao método sugerido por Stedinger para obtenção de aproximações empíricas para as distribuições estatísticas e portanto sofrem do mesmo defeito de subestimação da incerteza dos parâmetros. Note que o mesmo método de correção das estimativas sugerido por Stedinger pode ser incorporado aos esquemas de Reamostragem, diminuindo o problema. Em relação ao método de Stedinger, os métodos de Reamostragem possuem uma desvantagem que é a limitação do número de estimativas ao número de diferentes pseudo-amostras possíveis de serem formadas a partir da amostra original, muito embora em geral este limite seja bastante grande. A vantagem dos métodos de Reamostragem é que a modelagem adotada, a exceção da hipótese de independência, não interfere na formação de pseudo-amostras; o que faz com que as distribuições obtidas com estes métodos não incorporem os erros de modelagem.

Pode-se dizer que o método de Reamostragem deve ser aplicado sempre que existam dúvidas quanto ao modelo que melhor representa o processo e se deseja que a avaliação da incerteza dos parâmetros independa do modelo adotado. Em geral, na modelagem de séries hidrológicas sempre existem dúvidas quanto ao modelo, o que indicaria o uso do método de Reamostragem. Por outro lado, em geral os esquemas de Reamostragem se baseiam na hipótese de independência dos dados, o que usualmente não é verdade para séries hidrológicas.

Costa (1988) aplicou um procedimento de Reamostragem para incorporar a incerteza dos parâmetros em modelos de geração de séries temporais do tipo ARMA(p,q). Para superar o problema da independência dos dados Costa trabalhou sobre os ruídos, obtidos em um ajuste prévio do modelo adotado à série histórica. Isto é, fez-se uma reamostragem nos ruídos amostrais obtendo-se pseudo-séries de ruídos. Usando-se a série histórica gera-se a partir das pseudo-séries de ruídos novas estimativas para os parâmetros cuja distribuição empírica é considerada como distribuição de probabilidades para os parâmetros. Evidentemente este

método tem basicamente os defeitos do método de Stedinger (subestimação da incerteza dos parâmetros e incorporação dos erros de modelagem) e dos métodos de Reamostragem (limitação no número de pseudo-séries). Cover e Unny (1985) apresentaram um procedimento para modelos ARMA(p,q) onde a obtenção das pseudo-séries é feita substituindo-se cada ruído a_t da pseudo-série de ruídos pelo conjunto de observações da série histórica ($y_t, y_{t-1}, \dots, y_{t-p}$). Desta forma as distribuições empíricas obtidas não dependem dos parâmetros históricos ao contrário dos métodos de Costa e Stedinger.

O objetivo desta tese é desenvolver e aplicar (em estudos de regularização) um método de Reamostragem que não incorpore os erros de modelagem, similar ao procedimento de Cover e Unny, porem mais geral, para a consideração da incerteza nos parâmetros em modelos de geração de séries de vazões mensais. Para tal utilizou-se uma versão univariada do Modelo de Simulação Hidrológica (M.S.H.), desenvolvido no CEPEL para estudos de planejamento do setor elétrico brasileiro. O M.S.H. (Pereira et al, 1984) é basicamente um modelo auto regressivo de ordem um AR(1) a nível anual com desagregação mensal pelo esquema proposto por Mejia e Rouselle (1976).

Uma descrição do M.S.H. é feita no capítulo 2. O desenvolvimento da metodologia para consideração da incerteza dos parâmetros é feita no capítulo 3. No capítulo 4 apresenta-se uma avaliação do efeitos da metodologia desenvolvida no capítulo anterior nos resultados de estudos de confiabilidade de regularização de vazões.

CAPÍTULO 2

DESCRIÇÃO DO MODELO DE SIMULAÇÃO HIDROLÓGICA

O Modelo de Simulação Hidrológica, M.S.H., é um modelo de geração de séries de vazões mensais, desenvolvido pelo CEPEL para estudos de planejamento do sistema hidroelétrico brasileiro. Basicamente ele é um modelo anual auto-regressivo de ordem 1, AR(1), (Matalas, 1967) com posterior desagregação em vazões mensais conforme esquema desenvolvido inicialmente por Mejia e Rouselle (1976). O M.S.H. é originalmente um modelo multivariado, podendo ser aplicado simultaneamente a vários postos. Neste trabalho foi usado o M.S.H. na sua versão simplificada que o transforma em um modelo univariado (aplicado a um posto de cada vez). Apesar desta simplificação o conceito teórico da implementação da incerteza dos parâmetros desenvolvido nesta tese pode ser facilmente estendido para o modelo geral multivariado do M.S.H.

Primeiramente apresentamos um breve resumo do que seja o modelo AR(1) e posteriormente uma descrição do esquema de desagregação (transformação de vazões anuais para mensais) utilizado.

2.1 MODELAGEM ANUAL AR(1)

O modelo estocástico AR(1) é um modelo que preserva a média, o desvio padrão e a auto correlação lag 1 da série utilizada para estimar seus parâmetros.

No modelo AR(1) as vazões anuais são geradas por uma equação do tipo:

$$Q_t = \mu + \rho \cdot (Q_{t-1} - \mu) + \sigma \cdot \sqrt{1 - \rho^2} \cdot Z_t \quad (2.1)$$

onde:

Q_t : Vazão anual do ano t

Z_t : Ruído com distribuição normal $N(0,1)$, Z_t independente de Z_τ , $t \neq \tau$

μ, σ, ρ : parâmetros

sendo:

μ : média da vazão anual

σ : desvio padrão da vazão anual

ρ : correlação entre Q_t e Q_{t+1}

Com Z_t sorteado de uma distribuição normal padrão as vazões geradas por este modelo terão uma distribuição normal tanto para a distribuição marginal de Q_t como para a distribuição condicionada de Q_t dado Q_{t-1} .

As distribuições das vazões anuais da maioria dos rios do mundo são melhor representadas por uma distribuição de probabilidades conhecida por Log-Normal de 3 parâmetros (LN3) fato este verificado por Klemes (1978) e explicado por Chow (1954) e Klemes (1978) que interpretaram os fenômenos hidrológicos como sendo o resultado de um conjunto de causas físicas que interagem tanto de forma multiplicativa (provável distribuição LN2) como aditiva (provável distribuição Normal). Portanto, é de se esperar que a provável verdadeira distribuição esteja situada entre os dois tipos (LN2 e N). A distribuição Log-normal de 3 parâmetros satisfaz essa propriedade, sendo dada por:

$$f_Q(q) = \frac{\exp\left[-\frac{1}{2} \left(\frac{\ln(q-\alpha) - \mu_n}{\sigma_n}\right)^2\right]}{(q-\alpha) \cdot \sigma_n \cdot \sqrt{2} \cdot \pi} \quad (2.2)$$

cujos momentos são:

$$\text{média} \quad : \quad \mu = \alpha + \exp\left(\frac{\sigma_n^2}{2} + \mu_n\right) \quad (2.3)$$

$$\text{variância} \quad : \quad \sigma^2 = \exp[2 \cdot (\sigma_n^2 + \mu_n)] - \exp(\sigma_n^2 + 2\mu_n) \quad (2.4)$$

$$\text{assimetria} \quad : \quad \gamma = \frac{\exp(3 \cdot \sigma_n^2) - 3 \cdot \exp(\sigma_n^2) + 2}{[\exp(\sigma_n^2) - 1]^{3/2}} \quad (2.5)$$

Para que o modelo AR(1) gere vazões anuais com distribuição marginal LN3 a equação (2.1) deve se transformar em:

$$\ln(Q_t - \alpha) = \mu_n + \rho_n \cdot (\ln(Q_{t-1} - \alpha) - \mu_n) + \sigma_n \cdot \sqrt{1 - \rho_n^2} \cdot Z_t \quad (2.6)$$

onde:

$$\mu_n = \frac{1}{2} \cdot \ln[\sigma^2 / (\delta^2 - \delta)] \quad (2.7)$$

$$\sigma_n = \sqrt{\ln \delta} \quad (2.8)$$

$$\rho_n = -\frac{\text{Ln}[1 + \rho \cdot (\delta - 1)]}{\text{Ln } \delta} \quad (2.9)$$

$$\alpha = \mu - [\sigma^2 / (\delta - 1)]^{1/2} \quad (2.10)$$

$$\delta = [(1 + \gamma^2/2) + (\gamma^2 + \gamma^4/4)^{1/2}]^{1/3} + [(1 + \gamma^2/2) - (\gamma^2 + \gamma^4/4)^{1/2}]^{1/3} - 1 \quad (2.11)$$

Pode-se considerar como parâmetros deste modelo as variáveis:

μ : Média da vazão anual

σ : Desvio padrão da vazão anual

ρ : Coeficiente de auto correlação lag 1 das vazões anuais.

γ : Coeficiente de assimetria das vazões anuais.

A obtenção das variáveis μ_n , σ_n , ρ_n e α é feita simplesmente substituindo-se em (2.7) a (2.11) μ , σ , ρ , e γ pelas suas estimativas obtidas da série histórica. A tabela 1 apresenta as fórmulas usadas para a obtenção destas estimativas.

Depois do cálculo dos parâmetros do modelo, se a probabilidade de ocorrer vazões negativas exceder o valor arbitrado de $0,25 \times 10^{-3}$, isto é, se

$$\text{Pr}(Q \leq 0) = \Phi \{[\text{Ln}(-\alpha) - \mu_n] / \sigma_n\} > 0,25 \times 10^{-3}$$

onde Φ é uma distribuição acumulada normal padrão, aumenta-se artificialmente a assimetria até que

$$\Phi \{[\text{Ln}(-\alpha) - \mu_n] / \sigma_n\} = 0,25 \times 10^{-3}$$

A expressão acima pode ser escrita em termos de δ , μ , σ utilizando as relações (2.7) à (2.10):

$$\Phi \left\{ \frac{\text{Ln}[\sqrt{\delta \cdot (\delta - 1)} \cdot (1/(\sqrt{\delta - 1}) + \mu/\sigma)]}{\sqrt{\text{Ln } \delta}} \right\} = 0,25 \times 10^{-3} \quad (2.12)$$

A equação acima é resolvida para δ através do método de Newton. O novo valor de δ é então aplicado à (2.6), (2.7), (2.8), (2.9) e (2.10) para recalcular os parâmetros.

2.2 ESQUEMA DE DESAGREGAÇÃO MENSAL

O modelo de desagregação usado no M.S.H. é o modelo de Mejia e Rouselle (1976) com $p=1$. Pode ser descrito por:

$$Y_t = X_t.A + Z_t.B + C.V_t \quad (2.13)$$

onde:

Y_t : vetor coluna de dimensão 12 cujos elementos fornecem a vazão de cada mês i do ano t menos a média das vazões do mês i .

X_t : vazão anual do ano t menos a média das vazões anuais.

Z_t : vazão mensal de dezembro do ano $t-1$ menos a média das vazões de dezembro.

V_t : vetor coluna de dimensão 11 contendo ruídos de distribuição normal de média zero e desvio padrão unitário, independentes entre si.

A : Matriz 12×1 de parâmetros.

B : Matriz 12×1 de parâmetros.

C : Matriz 12×11 de parâmetros.

Reescrevendo (2.13):

$$\begin{bmatrix} q_{t,1} - \mu_1 \\ q_{t,2} - \mu_2 \\ q_{t,3} - \mu_3 \\ \vdots \\ \vdots \\ q_{t,i} - \mu_i \\ \vdots \\ q_{t,12} - \mu_{12} \end{bmatrix} = (Q_t - \mu) \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ \vdots \\ a_i \\ \vdots \\ a_{12} \end{bmatrix} + (q_{t-1,12} - \mu_{12}) \cdot \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ \vdots \\ b_i \\ \vdots \\ b_{12} \end{bmatrix} + \begin{bmatrix} c_{1,1} & \dots & \dots \\ c_{2,1} & & \\ c_{3,1} & & \\ \vdots & & \\ \vdots & & \\ \vdots & & \\ c_{i,1} & \dots & c_{i,j} & \dots & \dots \\ \vdots & & & & \\ c_{12,1} & \dots & \dots & \dots & c_{12,11} \end{bmatrix} \cdot \begin{bmatrix} v_{t,1} \\ v_{t,2} \\ v_{t,3} \\ \vdots \\ \vdots \\ v_{t,j} \\ \vdots \\ v_{t,11} \end{bmatrix}$$

onde:

$q_{t,i}$: Vazão do mês i do ano t

μ_i : Média das vazões mensais do mês i

Q_t : Vazão anual do ano t

μ : Média das vazões anuais

$v_{t,i}$: Ruído do ano t sorteados de uma distribuição normal padrão, $v_{t,i}$ independentes de $v_{t,j}$; $i \neq j$

As matrizes A, B, C são calculadas de tal maneira que se preserve aproximadamente a matriz de covariância mensal dada por:

$$\text{cov} \begin{bmatrix} Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} S_{yy} & S_{yz} \\ S_{zy} & S_{zz} \end{bmatrix} \quad (2.14)$$

onde:

S_{yy} : matriz de covariância entre vazões de meses de um mesmo ano

S_{zz} : variância da vazão de dezembro do ano anterior

S_{yz} : matriz de covariância entre vazões do mês i e de dezembro do ano anterior

$$S_{zy} = S_{yz}^t$$

As matrizes S_{yy} , S_{yz} , S_{zy} , S_{zz} podem ser consideradas como os parâmetros do esquema de desagregação de Mejia. A tabela 1 fornece as equações usadas para a estimação das matrizes S_{yy} , S_{yz} , S_{zz} . Note que na estimação de S_{yy} abandonou-se os dados do primeiro ano do histórico enquanto na estimação de S_{zz} abandonou-se os dados do último ano do histórico, isto afim de garantir uma coerência entre estas estimativas e a estimativa de S_{yz} .

Seja D um vetor linha de dimensão 12 dado por:

$$D = [1/12, 1/12, 1/12, \dots, 1/12]$$

logo para preservação do processo anual previamente gerado têm-se:

$$(Q_t - \mu) = D \cdot Y_t \quad (2.15)$$

As equações para cálculo das matrizes A, B, C, a partir das estimativas de S_{yy} , S_{yz} , S_{zy} , S_{zz} são:

$$A = [S_{yy} - S_{yz} (S_{zz})^{-1} S_{zy}] D^t \{D [S_{yy} - S_{yz} (S_{zz})^{-1} S_{zy}] D^t\}^{-1} \quad (2.16)$$

$$B = (I - A D) S_{yz} (S_{zz})^{-1} \quad (2.17)$$

$$\Delta = S_{yy} - A D (S_{yy} D^t A^t + S_{yz} B^t) + B (S_{zy} D^t A^t + S_{zz} B^t) \quad (2.18)$$

$$C \cdot C^t = \Delta \quad (2.19)$$

onde I é a matriz de identidade.

As soluções de (2.16) e (2.17) fornecem as matrizes A e B que quando substituídas em (2.18) fornecem Δ . Dois casos podem então acontecer:

A) A matriz Δ tem 11 autovalores positivos e um nulo (devido à (2.15) sempre existirá um autovalor nulo).

Neste caso (2.19) é indeterminada, existindo uma família de soluções para C dada por:

$$C = C_0 \cdot \beta \quad (2.20)$$

onde

C_0 é uma solução qualquer de (2.19)
 β é uma matriz (12x12) tal que $\beta \cdot \beta^t = I$

Esta indeterminação não é problema pois qualquer solução pode ser igualmente usada no esquema de geração. Uma solução pode ser obtida por decomposição espectral através de

$$C = P \cdot \lambda^{1/2} \quad (2.21)$$

onde:

P : Matriz (12x11) de auto vetores da matriz Δ dispostos nas colunas na mesma ordem dos autovalores em λ
 λ : Matriz diagonal (11x11) de autovalores dispostos em ordem decrescente

Outra alternativa para obtenção de uma solução para C neste caso é o uso da decomposição de Choleski onde assume-se que a matriz C resultante é triangular inferior.

B) A matriz Δ têm mais de um autovalor nulo.

Neste caso (2.19) não tem solução. Este caso acontece quando existe dependência linear dentro de grupos de vazões mensais. Na estimação de S_{yy} , S_{yz} e S_{zz} estão envolvidos 13 meses (os 12 meses do ano mais o mês de dezembro do ano anterior). Como, no esquema de desagregação, a vazão total é conhecida podemos retirar um dos 12 meses do ano, restando para a análise 12 meses. Se o comprimento da série histórica for de 12 anos, existirão (na estimação de S_{yy} , S_{yz} , S_{zz}) para cada um dos 12 meses da análise apenas 11 valores (abandona-se para estimação os meses de janeiro a novembro do primeiro ano do registro). Neste caso será sempre possível exprimir qualquer destes meses como uma função linear dos outros 11. Em consequência Δ terá pelo menos mais um autovalor nulo e (2.19) não poderá ser resolvida. Se retirarmos deste histórico algum ano, Δ ganhará pelo menos mais um autovalor nulo e por mais razão (2.19) não poderá ser resolvida. Se ao contrário adicionarmos um ano ao histórico, provavelmente não existirá mais a dependência linear entre as vazões mensais e normalmente (2.19) poderá ser resolvida. Note que se o ano adicionado for na verdade uma repetição de algum ano já existente a dependência linear se manterá e (2.19) continuará sem solução.

Em conclusão para aplicação do esquema de desagregação do MSH é preciso se ter pelo menos 13 anos que formem 12 vetores de vazões mensais (incluindo dezembro do ano anterior) independentes linearmente. Finalmente vale lembrar que pode-se obter matrizes Δ com autovalores até mesmo negativos por problemas numéricos ou por problemas de inconsistências na estimação de S_{yy} , S_{yz} , S_{zz} . O uso das fórmulas da tabela 1 evita este último problema.

TABELA 1: Equações de estimação dos parâmetros de MSH a partir de uma série histórica com n anos.

Parâmetro	Equação
$\hat{\mu}$	$\frac{1}{n} \cdot \sum_{t=1}^n Q_t$
$\hat{\sigma}^2$	$\frac{1}{n} \cdot \sum_{t=1}^n (Q_t - \hat{\mu})^2$
$\hat{\rho}$	$\frac{1}{n \cdot \hat{\sigma}^2} \cdot \sum_{t=2}^n [(Q_t - \hat{\mu}) \cdot (Q_{t-1} - \hat{\mu})]$
$\hat{\gamma}$	$\frac{n}{\hat{\sigma}^3 \cdot (n-1) \cdot (n-2)} \cdot \sum_{t=1}^n (Q_t - \hat{\mu})^3$
$\hat{\mu}_i$	$\frac{1}{n} \cdot \sum_{t=1}^n q_{t,i} \quad \text{para } i=1, 12$
$\hat{s}_{yy}(i,j)$	$\frac{1}{n-1} \cdot \sum_{t=2}^n [(q_{t,i} - \mu_i^*) \cdot (q_{t,j} - \mu_j^*)] \quad \text{para } i=1, 12 \text{ e } j=1, 12$
	$\mu_i^* = \frac{1}{n-1} \sum_{t=2}^n q_{t,i}$
$\hat{s}_{yz}(i)$	$\frac{1}{n-1} \cdot \sum_{t=2}^n [(q_{t,i} - \mu_i^*) \cdot (q_{t-1,12} - \mu_{12}^{\#})] \quad \text{para } i=1, 12$
	$\mu_{12}^{\#} = \frac{1}{n-1} \sum_{t=1}^{n-1} q_{t,12}$
\hat{s}_{zz}	$\frac{1}{n-1} \cdot \sum_{t=1}^n (q_{t,12} - \mu_{12}^{\#})^2$

CAPÍTULO 3

IMPLEMENTAÇÃO DA INCERTEZA NOS PARÂMETROS NO MODELO MSH UNIVARIADO

Este capítulo descreve a metodologia usada para implementar a incerteza dos parâmetros na versão univariada do modelo M.S.H.. A metodologia escolhida para incorporar a incerteza dos parâmetros foi uma técnica de Reamostragem do tipo Bootstrap.

3.1 BOOTSTRAP

O Bootstrap (Efron, 1982) consiste num esquema de, a partir de uma amostra original $X = \{x_1, x_2, x_3, \dots, x_n\}$ (onde x_i independe de x_j ; $\forall i, j; i \neq j$), obter-se outras amostras, diferentes de X , de tamanho n por simples sorteios com reposição das observações da amostra original. Estas amostras derivadas são chamadas de pseudo-amostras. Note que o número de diferentes pseudo-amostras possíveis de se obter num esquema de reamostragem com reposição é bastante grande podendo ser calculado por $(2^n - 1)$ (Cover e Unny, 1985). Como exemplo, para $n=10$ temos 92378 pseudo-amostras possíveis.

No esquema do Bootstrap, a cada pseudo-amostra aplica-se os métodos de estimação do modelo adotado obtendo-se uma estimativa para o vetor de parâmetros. O conjunto de vetores assim obtido constitui a distribuição de probabilidades do vetor de parâmetros do modelo. Portanto o algoritmo para a determinação da distribuição de probabilidades dos parâmetros pode ser descrito por:

1 – Faz-se uma reamostragem com reposição das observações da amostra X obtendo-se uma pseudo-amostra $X_{boot}^1 = \{x_1^*, x_2^*, x_3^*, \dots, x_n^*\}$

2 – A partir da pseudo-amostra X_{boot}^i obtem-se uma estimativa do vetor de parâmetro.

$$\theta_{boot}^i = g(X_{boot}^i) = g(x_1^*, x_2^*, x_3^*, \dots, x_n^*)$$

3 – Repetições independentes dos passos 1 e 2, fornecem um conjunto de estimativas do vetor de parâmetros $\theta, \theta_{boot}^1, \theta_{boot}^2, \dots, \theta_{boot}^b$, cuja distribuição

empírica constitui a distribuição de probabilidades de θ .

3.2 DESENVOLVIMENTO DE UM ESQUEMA DE BOOTSTRAP EM SÉRIES COM DEPENDÊNCIA TEMPORAL

Uma condição para aplicabilidade do método Bootstrap é a independência dos dados a serem reamostrados. Entretanto as vazões fluviais não são independentes entre si, isto é, para instantes próximos teremos vazões próximas (por exemplo: se a vazão num certo instante for baixa, as vazões nos instantes seguintes tenderão também a serem baixas). A esse fenômeno dá-se o nome de persistência ou dependência temporal que pode ser expresso matematicamente pelo correlograma (função que relaciona a correlação entre valores de vazões em diferentes instantes com o intervalo de tempo entre elas). Em geral as séries de vazões fluviais podem ser representadas aproximadamente por um processo AR(k), com k grande, dado por

$$a_t = Q_t - \varphi_1 Q_{t-1} - \varphi_2 Q_{t-2} - \varphi_3 Q_{t-3} \dots - \varphi_k Q_{t-k} \quad (3.1)$$

onde a_t é um ruído aleatório independente no tempo e portanto possível de ser reamostrado. Cover e Unny (1985), sugeriram que a reamostragem do ruído a_t pode ser feita reamostrando-se as k-uplas

$$QM_t^k = (Q_t, Q_{t+1}, Q_{t+2}, \dots, Q_{t+k-1}) \quad (3.2)$$

. Um esquema de Bootstrap aproximado seria construir pseudo-séries anuais de tamanho n a partir de reamostragem com reposição de vetores QM_t^k usando-se para k um valor suficientemente grande para uma boa representação do processo estocástico.

Suponha que o registro de vazões no local de interesse abranja n anos, fornecendo uma série histórica de vazões anuais $QH = \{Q_1, Q_2, Q_3, \dots, Q_n\}$. Suponha ainda que haja interesse no estudo da incerteza dos parâmetros μ, σ, ρ .

Da série histórica QH pode-se obter uma amostra original X formada por n-k+1 observações independentes do processo multivariado QM_t^k , dada por

$$X = \{ QM_t^k ; t=1, n-k+1 \}. \quad (3.3)$$

A aplicação imediata do esquema tradicional de Bootstrap a partir desta amostra cria um problema na estimação dos parâmetros já que para $k > 1$ a probabilidade de uma vazão qualquer da série histórica pertencer à uma pseudo-amostra depende do seu ano. Por exemplo, a vazão do primeiro ano do histórico, Q_1 , só pertence a pseudo-amostra se QM_1^k for sorteado, enquanto que Q_2 pertencerá a pseudo-amostra se QM_1^k ou QM_2^k forem sorteados. Este problema pode ser resolvido adicionando-se à amostra X as observações mutiladas do processo QM_t^k , para $t=2-k$ à $t=0$ e para $t=n-k+2$ à $t=n$ criando-se uma amostra original estendida

$$X_{ex} = \{ QM_t^k ; t=2-k, n \} \quad (3.4)$$

O esquema de reamostragem será feito criando-se pseudo-amostras por sorteios com reposição dos elementos de X_{ex} . O número de sorteios necessários para formar uma pseudo-amostra não é definido a priori; considerando-se que uma pseudo-amostra estará formada quando os sorteios já feitos permitirem retirar $n-1$ pares de vazões anuais (Q_T, Q_{T+1}) realmente ocorridas consecutivamente. Com isso consegue-se reproduzir o número de pares disponíveis para a estimação de ρ existente no registro histórico. Obtida uma pseudo-amostra pelo esquema acima, teremos $n-1$ pares para estimar ρ , mas para a estimação de μ e σ estarão disponíveis em geral um número bem maior que n . Para simular o grau de informação histórico só utilizaremos as n primeiras vazões anuais obtidas na reamostragem.

O esquema de Reamostragem proposto acima foi testado num estudo de Monte Carlo (com 1000 series) para a avaliação das incertezas correspondentes aos parâmetros μ , σ e ρ ; considerou-se para tal a série histórica de 52 anos de afluências anuais ao reservatório de Furnas. As distribuições estatísticas dos parâmetros foram obtidos empiricamente conforme o método de Stedinger através do uso do modelo AR-1 normal e pelo esquema de Reamostragem proposto usando-se valores de $k = 2$. A tabela 3 apresenta as estimativas históricas de μ , σ , ρ além da média e do coeficiente de variação das distribuições empíricas destes parâmetros obtidas com cada método.

Pode-se notar que os momentos das distribuições obtidas com o esquema proposto apresenta resultados da mesma ordem de grandeza que os obtidos com o método de Stedinger. Em termos do parâmetro μ , os valores médios das distribuições dos 2 métodos foram muito próximos do valor histórico. Em relação ao parâmetro σ , todas duas distribuições obtidas apresentam valores médios inferiores ao valor histórico, sendo a diferença maior no método de Stedinger. Nota-se em relação ao parâmetro ρ que as estimativas obtidas pelo método de Stedinger se distribuem em torno de um valor inferior ao valor histórico o que não acontece com as estimativas obtidas pelo esquema de Reamostragem. Para todos os parâmetros as distribuições obtidas pelo esquema de Reamostragem apresentam menor variabilidade do que as obtidas com o método de Stedinger.

TABELA 3 Média e Coeficiente de Variação (entre parentesis) das distribuições empíricas dos parâmetros para diferentes métodos (obs: foram usadas 1000 pseudo amostras ou séries sintéticas).

Parâmetro	Estimativa Histórica	Método Stedinger	Reamostragem k=2
μ	917,7 (-)	916,2 (0,05)	916,2 (0,04)
σ	235,7 (-)	230,2 (0,11)	231,3 (0,09)
ρ	0,34 (-)	0,30 (0,45)	0,35 (0,39)

3.3 UTILIZAÇÃO DO ESQUEMA DE REAMOSTRAGEM PARA IMPLEMENTAÇÃO NO MSH

Suponha que o esquema de reamostragem proposto no item anterior tenha sido aplicado a um registro de vazões de n anos; então existirá disponível um número grande de pseudo-amostras formadas por sorteios com reposição dos elementos de X_{τ} , sendo possível retirar de cada pseudo-amostra $n-1$ pares de vazões anuais $(Q_{\tau}, Q_{\tau+1})$ realmente ocorridas consecutivamente. Suponha ainda que haja interesse no estudo da incerteza de todos os parâmetros do modelo MSH (ver tabela 1).

Assim como descrito no item anterior, a estimativa de ρ de cada pseudo-amostra será obtida usando-se os $n-1$ pares de vazões anuais realmente ocorridas consecutivamente disponíveis na pseudo amostra, enquanto que os parâmetros μ e σ serão estimados usando-se as n primeiras vazões anuais da pseudo-amostra. O parâmetro γ também será estimado usando-se as n primeiras vazões anuais da pseudo-amostra, assim como os parâmetros μ_i , $i=1,2$ que usarão as vazões mensais destes mesmos anos.

Os parâmetros restantes $s_{yy}(i,j)$, $i=1,2$, $j=1,2$; $s_{yz}(i)$, $i=1,2$ e s_{zz} , serão estimadas usando dados dos $n-1$ pares de anos $(\tau, \tau+1)$ usados para estimar ρ . Para $s_{yy}(i,j)$ se utilizará apenas as vazões mensais do anos $\tau+1$. Para $s_{yz}(i)$ se usará as vazões de dezembro dos anos τ e as vazões mensais dos anos $\tau+1$, enquanto que para s_{zz} se utilizará as vazões de dezembro dos anos τ . A Tabela 4 descreve com detalhes as equações usadas para a estimação dos parâmetros do M.S.H. a partir das pseudo-amostras.

Pode-se notar que o esquema proposto procura reproduzir para todos os parâmetros do MSH o grau de informação do histórico. Note ainda que a matriz de covariância mensal S_{yz} foi estimada de tal maneira que preserve uma coerência com a estimativa da correlação anual ρ , enquanto que as matrizes de covariância S_{yy} e S_{zz} mantêm uma coerência com a estimação de S_{yz} .

Conforme descrito no capítulo 2, para a estimação da matriz C é preciso que hajam pelo menos 12 pares de anos consecutivos cujos vetores de vazões mensais (incluindo dezembro do ano anterior) sejam independentes linearmente. Como nas pseudo-amostras os pares de anos consecutivos são obtidos num esquema de Reamostragem com reposição, diversas das possíveis pseudo-amostras

não conterão 12 diferentes pares de anos consecutivos. Note que quanto maior for o intervalo de agregação k , maior a percentagem de pseudo-amostras que possuirão menos de 12 diferentes pares de anos consecutivos. Como pseudo-amostras com menos de 12 diferentes pares de anos consecutivos não podem ser usadas para a estimação de parâmetros (em particular a matriz C) elas devem ser abandonadas. Este fenômeno limita aplicação do método proposto já que a eficiência de qualquer método de Reamostragem depende de um grande número de possíveis pseudo-amostras. Por exemplo, muito embora em geral o M.S.H. possa ser aplicado a séries com 13 anos, a aplicação do método de Reamostragem proposto à séries deste comprimento permitiria uma só pseudo-amostra dada pela própria série histórica, o que levaria a falsa conclusão de que os únicos parâmetros possíveis fossem exatamente os obtidos a partir da série histórica.

TABELA 4 Equações de estimação dos parâmetros a partir de uma pseudo amostra.

$t(p)$ = p-ésimo ano obtido na pseudo amostra.

$\tau(\ell)$ = ano inicial do ℓ -ésimo par de anos consecutivos obtidos na pseudo amostra

Parâmetro	Equação
$\hat{\mu}$	$\frac{1}{n} \cdot \sum_{p=1}^n Q_{t(p)}$
$\hat{\sigma}^2$	$\frac{1}{n} \cdot \sum_{p=1}^n (Q_{t(p)} - \hat{\mu})^2$
$\hat{\rho}$	$\frac{1}{n \cdot \sigma^2} \cdot \sum_{\ell=1}^{n-1} [(Q_{\tau(\ell)} - \hat{\mu}) \cdot (Q_{\tau(\ell)+1} - \hat{\mu})]$
$\hat{\gamma}$	$\frac{n}{\sigma^3 \cdot (n-1) \cdot (n-2)} \cdot \sum_{p=1}^n (Q_{t(p)} - \hat{\mu})^3$
$\hat{\mu}_i$	$\frac{1}{n} \cdot \sum_{p=1}^n q_{t(p),i} \quad \text{para } i=1,12$
$\hat{s}_{yy}(i,j)$	$\frac{1}{n-1} \cdot \sum_{\ell=1}^{n-1} [(q_{\tau(\ell)+1,i} - \mu_i^*) \cdot (q_{\tau(\ell)+1,j} - \mu_j^*)]$
	$\mu_i^* = \frac{1}{n-1} \cdot \sum_{\ell=1}^{n-1} q_{\tau(\ell)+1,i} \quad \text{para } i=1,12$
$\hat{s}_{yz}(i)$	$\frac{1}{n-1} \cdot \sum_{\ell=1}^{n-1} [(q_{\tau(\ell)+1,i} - \mu_i^*) \cdot (q_{\tau(\ell),12} - \mu_{12}^{\#2})]$
	$\mu_{12}^{\#2} = \frac{1}{n-1} \cdot \sum_{\ell=1}^{n-1} q_{\tau(\ell),12}$
\hat{s}_{zz}	$\frac{1}{n-1} \cdot \sum_{\ell=1}^{n-1} (q_{\tau(\ell),12} - \mu_{12}^{\#2})^2$

CAPÍTULO 4

ESTUDOS DE REGULARIZAÇÃO CONSIDERANDO A INCERTEZA DOS PARÂMETROS

4.1 INTRODUÇÃO

Reservatórios de regularização têm como objetivo garantir o suprimento de uma certa demanda hídrica durante a ocorrência de estiagens, quando a vazão natural cai abaixo das vazões demandadas. Isto é feito armazenando, na época fora da estiagem, as vazões naturais excedentes à demanda.

Quanto maior for o volume do reservatório maior a vazão que pode ser regularizada, ou quanto maior a vazão que se deseja regularizar maior o tamanho do reservatório que se deve construir, configurando—se portanto uma relação entre o volume do reservatório e a vazão regularizada. Sendo as séries de vazões naturais afluentes em uma seção de um rio realizações de um processo estocástico, a intensidade das estiagens variam de um ano para outro não sendo conhecidas a priori. Por maior que seja o tamanho do reservatório, poderão sempre ocorrer estiagens tão severas e longas para as quais não se poderá garantir o atendimento da demanda, mesmo contando—se com o reservatório totalmente cheio ao seu início. Por outro lado, quanto maior o reservatório, mais severa e longa, portanto menos provável, é a estiagem capaz de provocar um déficit no atendimento. Logo, todo estudo de regularização envolve sempre três variáveis: o tamanho do reservatório, a vazão regularizada e o risco de déficit dado pela probabilidade da ocorrência de algum déficit no atendimento da demanda num intervalo de tempo pré—definido. Usualmente utiliza—se o intervalo de tempo anual (risco anual) ou o intervalo correspondente à vida útil (econômica ou física) do empreendimento, chamado de horizonte de planejamento.

Num estudo de regularização diversos são os tipos de problemas encontrados. Pode—se por exemplo desejar saber qual a vazão que é regularizada ao risco de 1% ao ano por um reservatório de volume conhecido ou então o problema pode ser projetar o volume do reservatório de uma forma que o risco de uma certa vazão fixada a priori seja igual à 5% ao ano. A solução destes problemas variam de um rio para outro sendo uma função das características do processo estocástico gerador das sequências de afluições naturais (que representa o fenômeno), tais como, valores médios, variabilidade, sazonalidade e persistência. Conforme visto nos capítulos anteriores, modelos de geração de séries sintéticas

procuram identificar nas séries históricas estas características em termos de parâmetros, de forma a reproduzi-las nas séries sintéticas. Neste capítulo as séries sintéticas geradas pelo modelo MSH, considerando e não considerando a incerteza dos parâmetros, serão aplicadas na solução de problemas de regularização avaliando-se desta forma o efeito da consideração da incerteza dos parâmetros.

4.2 DESCRIÇÃO DOS ESTUDOS

Como exemplo de aplicação do modelo MSH com incorporação da incerteza dos parâmetros, foram feitos estudos de regularização, avaliando a influência da incerteza dos parâmetros. Para tal usou-se a série de 52 anos (1931 a 1983) de afluições mensais ao reservatório de Furnas ($17,217 \times 10^9 \text{ m}^3$) localizado no rio Grande, cujos principais momentos anuais históricos são: média de $917,7 \text{ m}^3/\text{s}$, desvio padrão de $235,7 \text{ m}^3/\text{s}$ e auto correlação lag-1 igual a 0,34. Apesar de um estudo univariado, isto é, com apenas um posto, as conclusões sobre a influência da incerteza dos parâmetros fornece uma idéia do que seria esta influência num sistema de vários reservatórios.

Na metodologia de Reamostragem para a consideração da incerteza dos parâmetros no MSH usou-se $k=2$ (tamanho do bloco igual a 2 anos), isto é, considerou-se que as séries anuais no local possam ser bem representadas por um modelo AR(1). Note que, quanto maior o tamanho do bloco, maior o número de pares consecutivos iguais possuirá uma determinada pseudo amostra, aumentando o número de pseudo amostras abandonadas devido a impossibilidade de calcular a matriz C na desagregação do MSH. Em consequência diminui a eficiência da reamostragem com o aumento de k , pois diminuirão o número de pseudo amostras utilizadas.

Foram feitos estudos de regularização a nível de vazões anuais e mensais e para vários comprimentos da série histórica (10, 20, 30, 40 e 52 anos) obtidos dos trechos finais da série histórica de Furnas entre 1931 a 1983. Note que para 10 anos não foram feitos estudos a nível de vazões mensais devido a impossibilidade de desagregação. Foi avaliada a influência da incerteza dos parâmetros na vazão regularizada pelo reservatório existente em função do grau de informação da série histórica, obtendo-se as curvas comprimento da série histórica versus vazão regularizada para riscos anuais de déficits de 1%, 5% e 10%. Os resultados aparecem nas tabelas 4.1 e figuras 4.1

Foi também avaliada a influência da incerteza dos parâmetros na vazão

regularizada em função do tamanho do reservatório obtendo-se as curvas volume útil do reservatório versus vazão regularizada para comprimentos da série histórica de 20 e 52 anos e para riscos de déficits anuais de 1%, 5% e 10%. Estes resultados aparecem nas tabelas 4.2 e figuras 4.2

Os resultados foram obtidos por um esquema de Monte Carlo descrito por:

1— Do vetor de parâmetros históricos $\hat{\theta} = [\hat{\mu}, \hat{\sigma}, \hat{\rho}, \hat{\gamma}, (\hat{\mu}_i, i=1,12), \hat{S}_{yy}, \hat{S}_{yz}, \hat{S}_{zz}]$ gera-se 100 séries de 1000 anos cada.

2— Conhecido o tamanho do reservatório e o risco que se deseja correr, obtem-se a vazão regularizada correspondente através de um processo de tentativa e erro, onde para cada hipótese de vazão regularizada o risco é calculado pela média dos riscos de cada série que por sua vez são obtidos contando-se o número de anos com algum déficit no atendimento da vazão obtido na simulação da operação do reservatório (através da equação da continuidade) durante a ocorrência da série estudada, considerando-se sempre o reservatório inicialmente cheio e descrita por:

$$v_j = \max \{0, \min[V, v_{j-1} + i_j - v_{reg}]\}, \quad j=1,2.. \quad (4.1)$$

$$v_0 = V \quad (4.2)$$

onde

V é o volume do reservatório

v_j é o volume de água no reservatório ao final do j ésimo intervalo de tempo de simulação.

i_j é o volume afluente ao reservatório durante o j ésimo intervalo de tempo

v_{reg} é o volume correspondente à vazão regularizada mantida durante um período de tempo igual ao intervalo de tempo da simulação (mensal ou anual).

Sempre que $v_j=0$ assume-se a ocorrência de déficit durante o j ésimo intervalo de tempo. No caso da simulação a nível de vazões mensais, um ano será contando como ano com déficit se ocorrer déficit em algum dos seus meses. Note que os valores calculados de risco de déficit de cada série serão diferentes entre si já que cada valor é apenas uma estimativa (o tamanho da série sintética é igual 1000 e não infinito).

3— Da série histórica obtem-se pelo método de reamostragem proposto 100

possíveis vetores de parâmetros θ_i . Com cada um gera-se uma série de 1000 anos.

4- Repete-se o item 2 usando-se estas novas 100 séries de 1000 anos. Note que a variação dos riscos calculados para cada série neste caso reflete não só a variabilidade devido ao tamanho da série sintética usada como também a variabilidade dos parâmetros ou seja a incerteza no valor dos parâmetros.

Foi calculada, também por simulação, a distribuição de probabilidades do volume do reservatório necessário para a regularização, a nível de vazões anuais, de 80% da vazão média afluyente, sem nenhum déficit num horizonte de planejamento de 25 anos. Os resultados aparecem nas tabelas 4.3 e figuras 4.3. No esquema de Monte Carlo usado neste caso foram geradas 1000 séries com 25 anos cada. A cada série obteve-se o tamanho do reservatório necessário através de:

$$v_j = \min \{ 0, v_{j-1} + i_j - v_{reg} \}, \quad j=1,2 \dots \quad (4.3)$$

$$v_0 = 0 \quad (4.4)$$

$$V_{nec} = \left| \min_j \{ v_j \} \right| \quad (4.5)$$

TABELA 4.1—A Vazão regularizada (em % da vazão média histórica—917,7 m³/s) para o reservatório existente em Furnas e fazendo simulação com vazões anuais

Risco	Incerteza	Comprimento da série histórica				
		10 anos	20 anos	30 anos	40 anos	52 anos
1%	não	84,16	73,75	69,95	76,84	77,15
	sim	70,62	67,59	66,10	74,68	75,07
	Influência	13,54	6,16	3,85	2,16	2,08
5%	não	91,42	83,34	79,49	85,40	85,85
	sim	83,00	78,40	76,80	84,09	84,99
	Influência	8,42	4,94	2,69	1,31	0,86
10%	não	95,36	88,78	84,96	90,09	90,68
	sim	89,61	84,55	82,85	89,47	90,29
	Influência	5,75	4,23	2,11	0,62	0,39

TABELA 4.1–B Vazão regularizada (em % da vazão média histórica–917,7 m³/s) para o reservatório existente em Furnas e fazendo simulação com vazões mensais

Risco	Incerteza	Comprimento da série histórica			
		20 anos	30 anos	40 anos	52 anos
1%	não	70,28	67,03	73,40	73,59
	sim	66,06	63,46	71,58	71,73
	Influência	4,22	3,57	1,82	1,86
5%	não	79,98	76,54	82,19	82,53
	sim	77,55	74,15	81,21	81,81
	Influência	2,43	2,39	0,98	0,72
10%	não	85,59	82,10	87,14	87,65
	sim	83,53	80,24	86,62	87,26
	Influência	2,06	1,86	0,52	0,39

TABELA 4.1–A1 Desvio padrão do risco de déficit (em % da vazão média histórica 917,7 m³/s) para o reservatório existente em Furnas e fazendo simulação com vazões anuais

Risco	Incerteza	Comprimento da série histórica				
		10 anos	20 anos	30 anos	40 anos	52 anos
1%	não	0,23	0,57	0,59	0,57	0,57
	sim	2,80	1,73	1,18	1,52	1,24
5%	não	0,55	1,28	1,32	1,25	1,28
	sim	8,75	5,51	4,31	4,68	3,74
10%	não	0,80	1,77	1,86	1,74	1,76
	sim	12,76	8,77	7,09	7,28	5,87

TABELA 4.1—B1 Desvio padrão do risco de déficit (em % da vazão média histórica—917,7 m³/s) para o reservatório existente em Furnas e fazendo simulação com vazões mensais

Risco médio	Incerteza	Comprimento da série histórica			
		20 anos	30 anos	40 anos	52 anos
1%	não	0,55	0,53	0,49	0,49
	sim	1,32	1,12	1,43	1,15
5%	não	1,21	1,27	1,21	1,21
	sim	4,53	4,09	4,35	3,50
10%	não	1,59	1,69	1,53	1,59
	sim	7,54	6,61	6,79	5,40

TABELA 4.2—A Vazão regularizada (em % da vazão média histórica—917,7 m³/s) para diversos tamanhos de reservatório (em % do volume do reservatório existente em Furnas) e riscos de déficits, fazendo simulação com vazões anuais e usando série histórica de 20 anos.

Risco	Incerteza	Tamanho do reservatório				
		0	50	100	150	200
1%	não	46,63	66,08	73,75	79,05	82,75
	sim	45,53	61,43	67,59	71,39	73,99
	Influência	1,10	4,65	6,16	7,66	8,76
5%	não	59,06	76,29	83,34	87,51	90,33
	sim	58,60	73,03	78,40	81,67	83,89
	Influência	0,46	3,26	4,94	5,84	6,44
10%	não	66,72	82,49	88,78	92,47	94,88
	sim	65,88	79,43	84,55	87,60	89,58
	Influência	0,84	3,06	4,23	4,87	5,30

TABELA 4.2—B Vazão regularizada (em % da vazão média histórica—917,7 m³/s) para diversos tamanhos de reservatório (em % do volume do reservatório existente em Furnas) e riscos de déficits, fazendo simulação com vazões anuais e usando série histórica de 52 anos.

Risco	Incerteza	Tamanho do reservatório				
		0	50	100	150	200
1%	não	48,27	68,82	77,15	82,21	85,64
	sim	47,97	67,31	75,07	79,75	83,02
	Influência	0,30	1,51	2,08	2,46	2,62
5%	não	61,15	78,93	85,85	89,82	92,32
	sim	61,23	78,33	84,99	88,64	91,03
	Influência	-0,08	0,60	0,86	1,18	1,29
10%	não	68,47	84,57	90,68	93,95	95,97
	sim	68,88	84,44	90,29	93,39	95,30
	Influência	-0,41	0,13	0,39	0,56	0,67

TABELA 4.2—C Vazão regularizada (em % da vazão média histórica—917,7 m³/s) para diversos tamanhos de reservatório (em % do volume do reservatório existente em Furnas) e riscos de déficits, fazendo simulação com vazões mensais e usando série histórica de 20 anos.

Risco	Incerteza	Tamanho do reservatório				
		0	50	100	150	200
1%	não	< 0	59,60	70,28	76,61	80,84
	sim	< 0	56,79	66,06	71,10	74,71
	Influência		2,81	4,22	5,51	6,13
5%	não	4,23	69,75	79,98	85,28	88,73
	sim	2,51	68,53	77,55	81,74	84,53
	Influência	1,72	1,22	2,43	3,54	4,20
10%	não	12,38	75,91	85,59	90,38	93,31
	sim	12,36	74,91	83,53	87,45	89,96
	Influência	0,02	1,00	2,06	2,93	3,35

TABELA 4.2—D Vazão regularizada (em % da vazão média histórica—917,7 m³/s) para diversos tamanhos de reservatório (em % do volume do reservatório existente em Furnas) e riscos de déficits, fazendo simulação com vazões mensais e usando série histórica de 52 anos.

Risco	Incerteza	Tamanho do reservatório				
		0	50	100	150	200
1%	não	< 0	62,41	73,59	79,77	83,81
	sim	< 0	61,05	71,73	77,48	81,34
	Influência		1,36	1,86	2,29	2,47
5%	não	10,39	72,24	82,53	87,66	90,79
	sim	9,02	71,79	81,81	86,62	89,59
	Influência	1,37	0,45	0,72	1,04	1,20
10%	não	17,92	77,93	87,65	92,05	94,67
	sim	17,49	77,79	87,26	91,51	94,01
	Influência	0,43	0,14	0,39	0,54	0,66

TABELA 4.3 Distribuição do volume do reservatório (em 10⁶ m³) para demanda de 80% da vazão média histórica, e utilizando 10 anos da série histórica.

%	sem incerteza	com incerteza
95	28611,0	58679,0
90	23924,4	39968,8
75	18313,7	22616,4
50	12710,3	10899,1
25	9074,2	4482,4
10	6045,6	803,9
5	4583,5	0,0
μ	14380,0	17487,4
σ	7859,4	20976,5

4.3 ANÁLISE DOS RESULTADOS

Conforme verificamos por todas as tabelas e conforme o esperado, quando se leva em consideração a incerteza dos parâmetros a vazão regularizada obtida, para um certo volume útil do reservatório e para um certo risco de déficit, é em geral menor que a vazão obtida para as mesmas condições sem a consideração da incerteza dos parâmetros. Da mesma forma, o volume do reservatório necessário para garantir uma certa vazão a um certo risco obtido considerando-se a incerteza dos parâmetros é maior que o mesmo volume obtido sem a consideração da incerteza dos parâmetros. Isto porque ao considerarmos a incerteza dos parâmetros estamos adicionando mais uma variabilidade no processo de geração de sequências de vazões que é a correspondente variabilidade dos possíveis parâmetros.

Note pelas tabelas 4.1 e figuras 4.1 que conforme o esperado quanto menor o tamanho da série histórica maior é a influência da incerteza dos parâmetros na vazão regularizada. Note também que existe uma tendência para que quanto menor o risco de déficit almejado maior a influência da incerteza dos parâmetros. Isto pode ser explicado pelo fato de que em geral a inclusão da incerteza dos parâmetros em modelos probabilísticos tem maior influência nas probabilidades acumuladas para eventos mais extremos (estiagens severas ou períodos altamente úmidos), pouco alterando as probabilidades dos eventos mais corriqueiros.

As magnitudes da influência da incerteza dos parâmetros para o reservatório existente são da ordem de grandeza inferior a 8% da vazão média histórica a nível de vazões anuais e inferior a 5% quando se estuda o problema a nível de vazões mensais. Analizando as tabelas 4.2 e figuras 4.2 vemos que quanto maior o tamanho do reservatório maior a influência da incerteza dos parâmetros. Note ainda pela análise das tabelas 4.1 e 4.2 e suas figuras correspondentes que para o estudo a nível de vazões mensais a influência da incerteza dos parâmetros foi menor que a nível de vazões anuais.

A tabela 4.3 e figura 4.3 mostra que o efeito da incerteza dos parâmetros é maior na cauda superior da distribuição do reservatório necessário, confirmando a conclusão de que quanto menor o risco de déficit almejado, maior a influência da incerteza dos parâmetros.

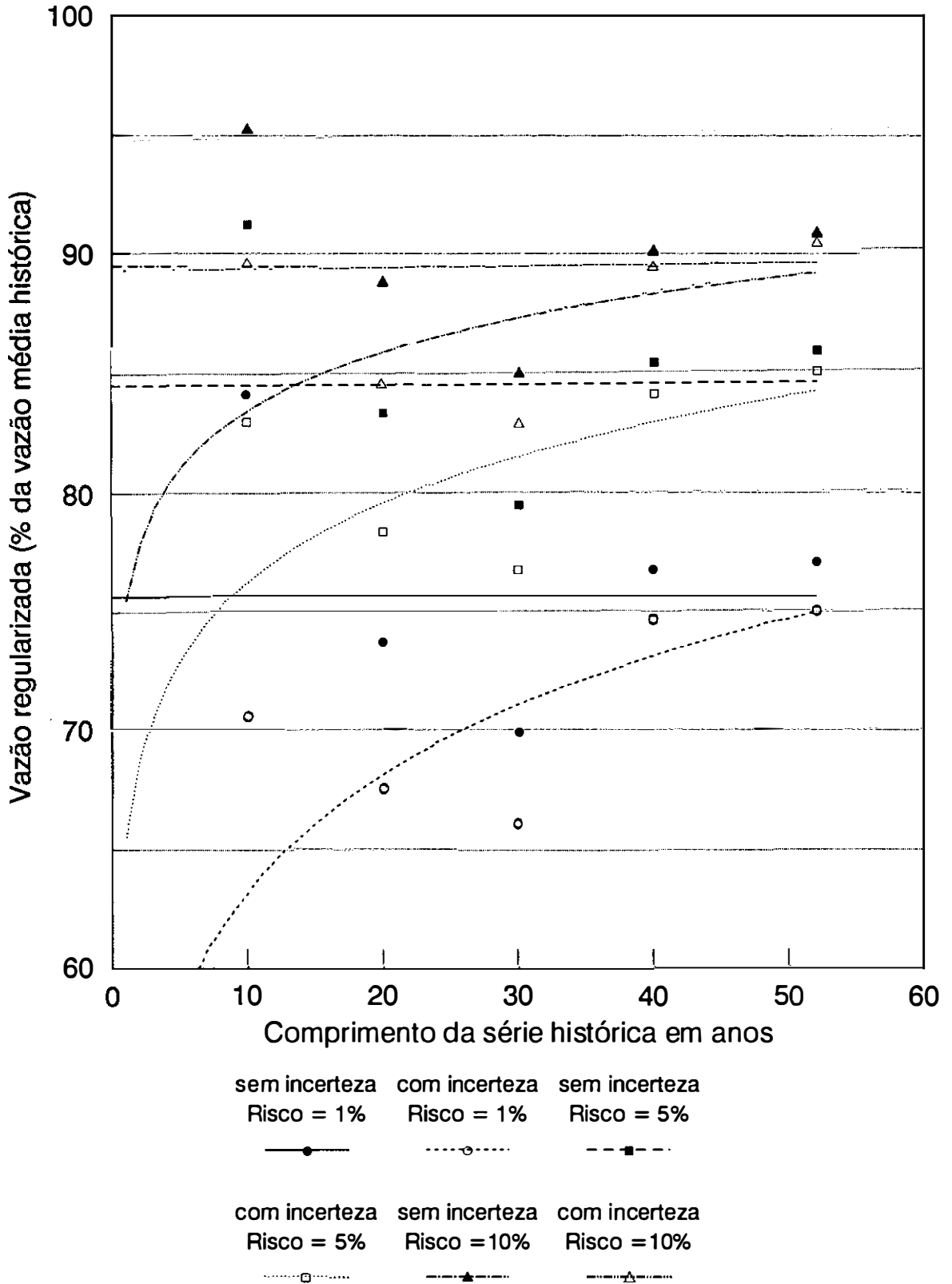


Figura 4.1-A Vazão regularizada versus tamanho da série histórica com simulação anual

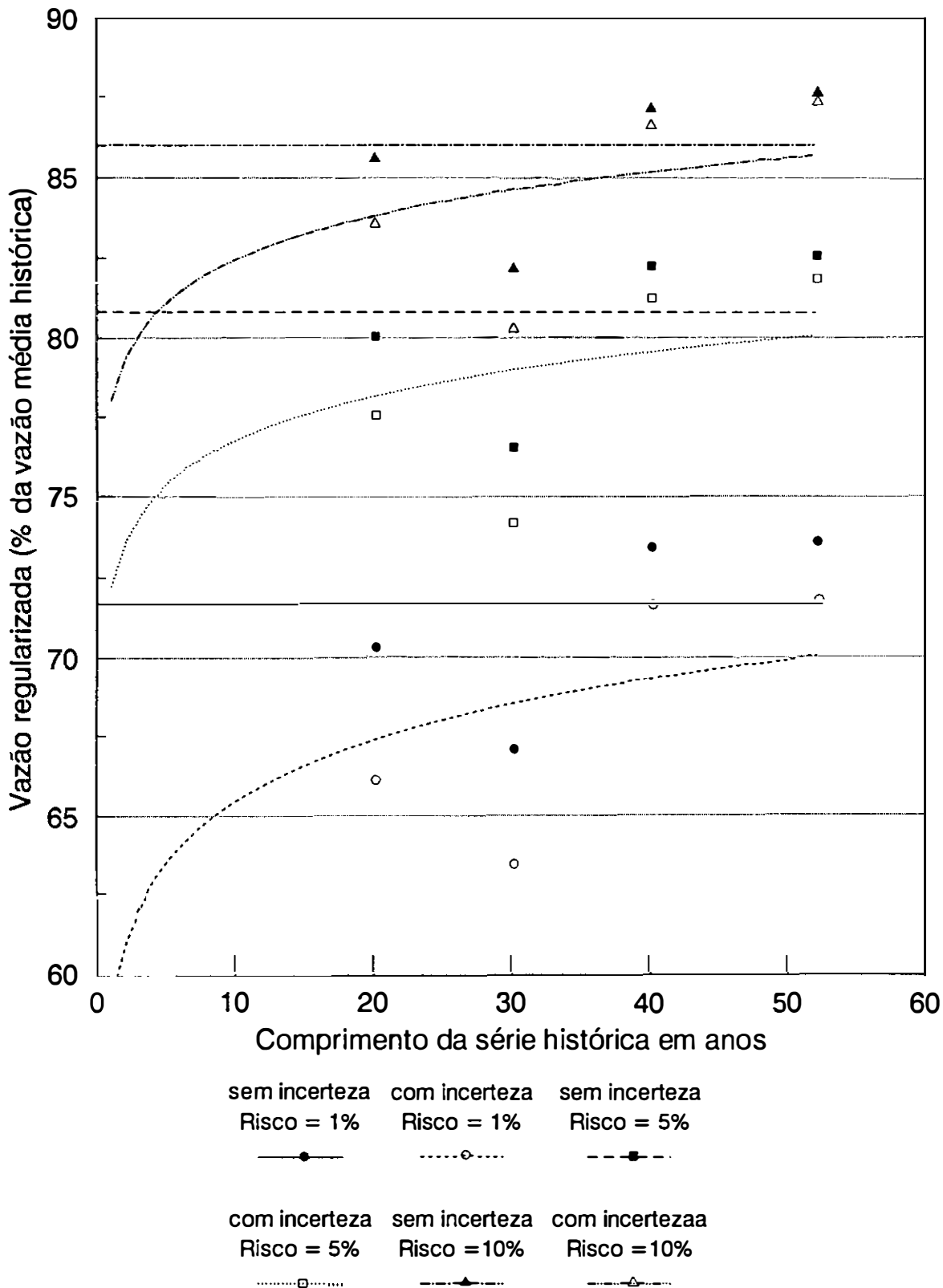


Figura 4.1-B Vazão regularizada versus tamanho da série histórica com simulação mensal

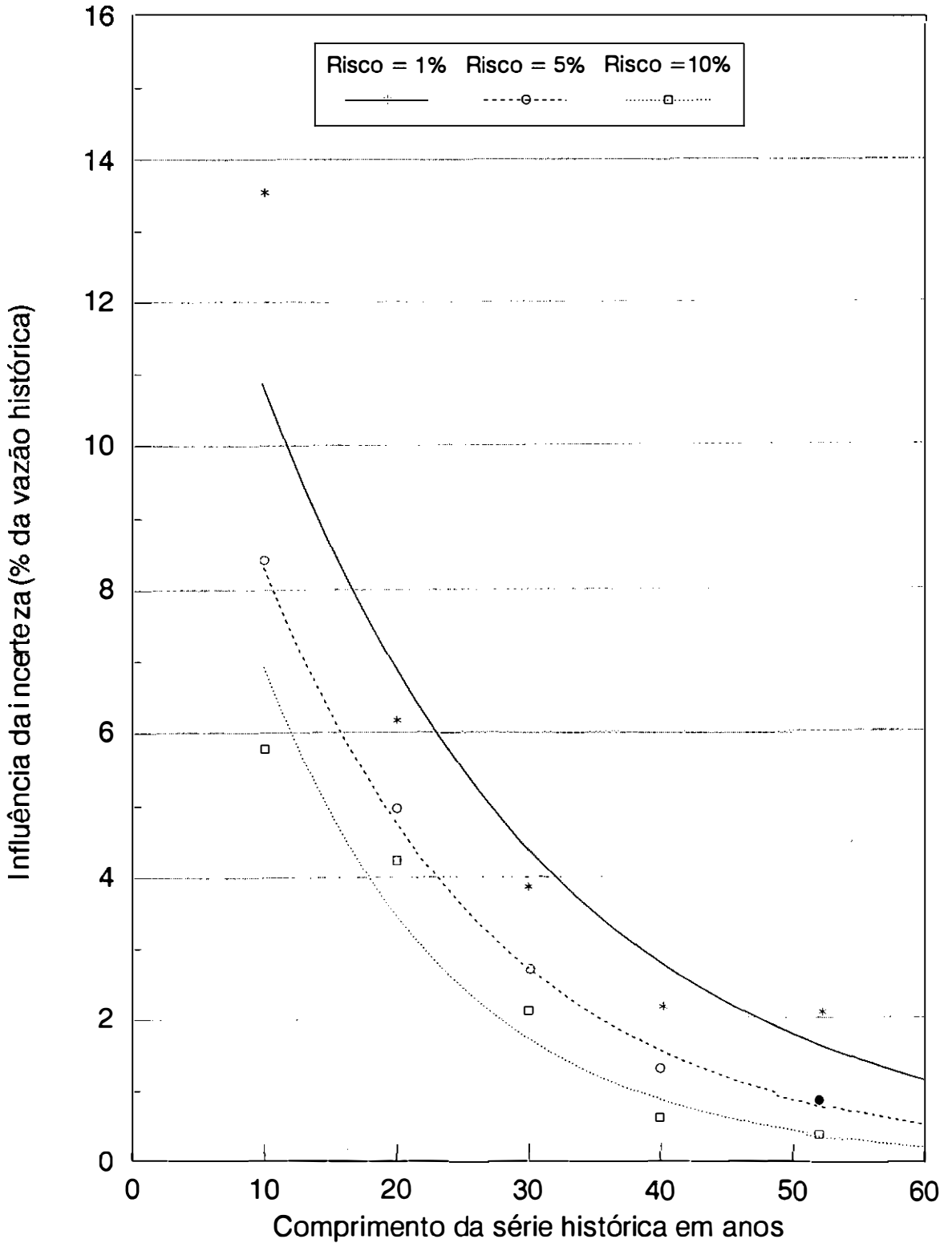


Figura 4.1-C1 Influência da incerteza dos parâmetros em função do comprimento da série histórica (com simulação anual)

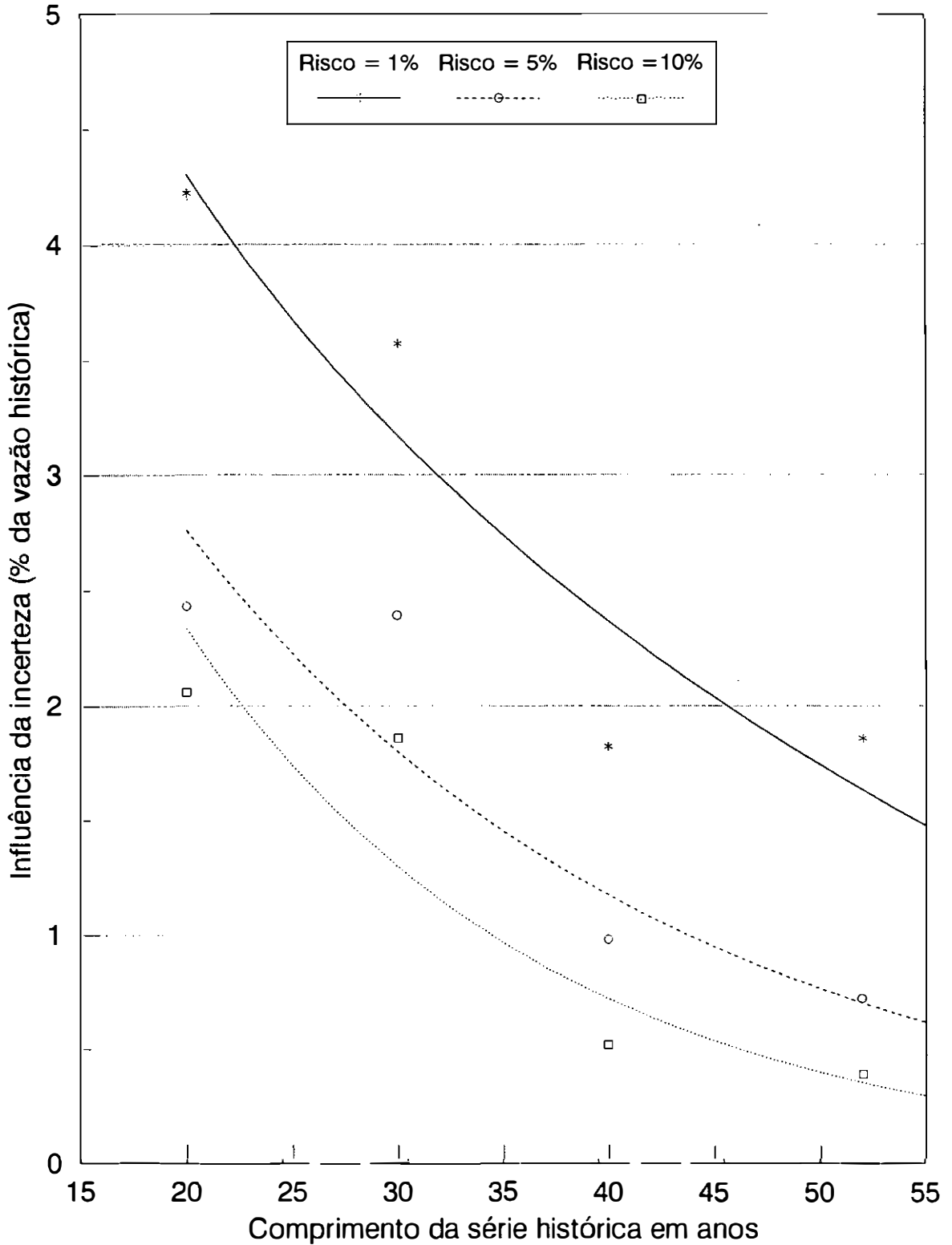


Figura 4.1-C2 Influência da incerteza dos parâmetros em função do comprimento da série histórica (simulação com vazões mensais)

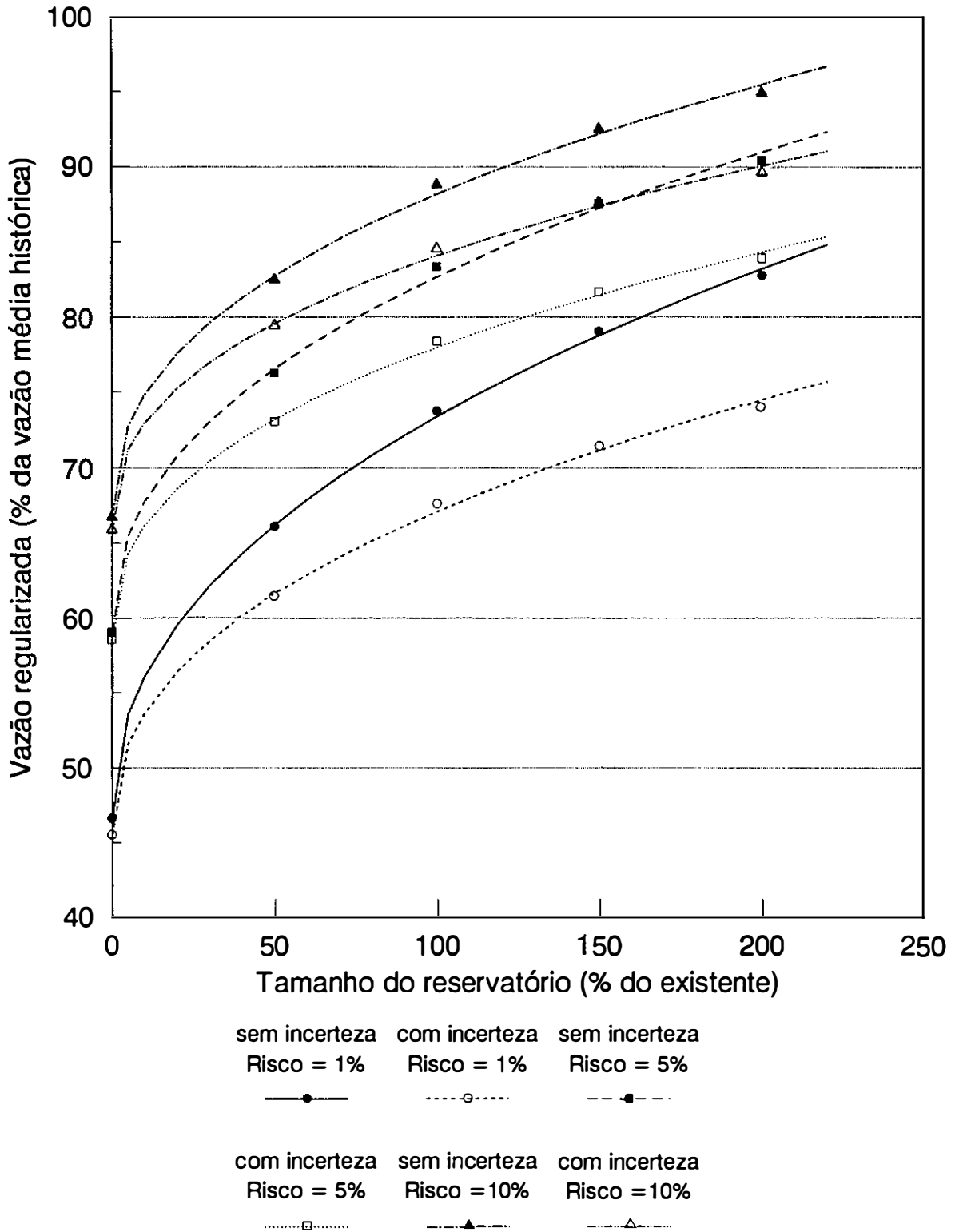


Figura 4.2-A Vazão regularizada versus tamanho do reservatório com simulação anual e série histórica de 20 anos

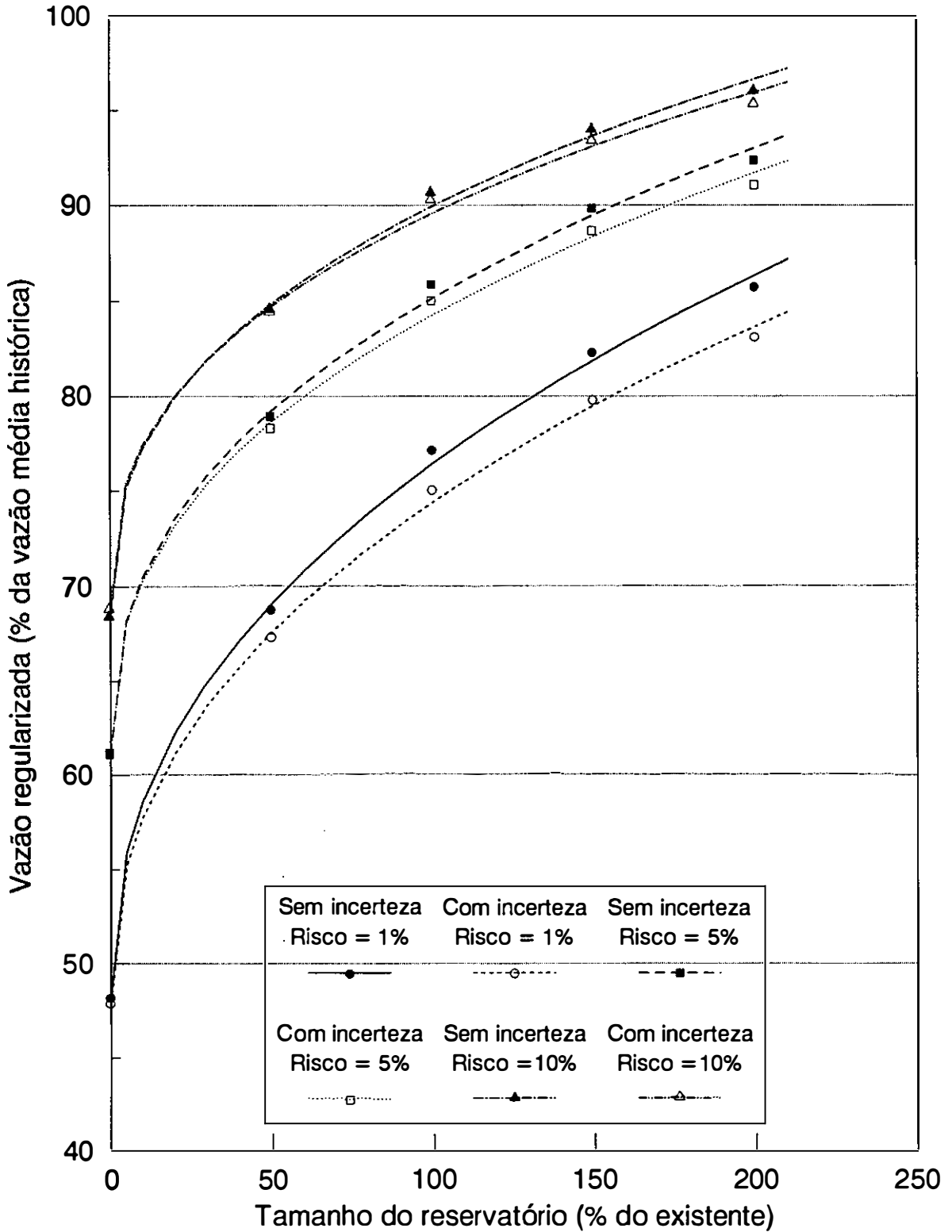


Figura 4.2-B Vazão regularizada versus tamanho do reservatório com simulação anual e série histórica de 52 anos

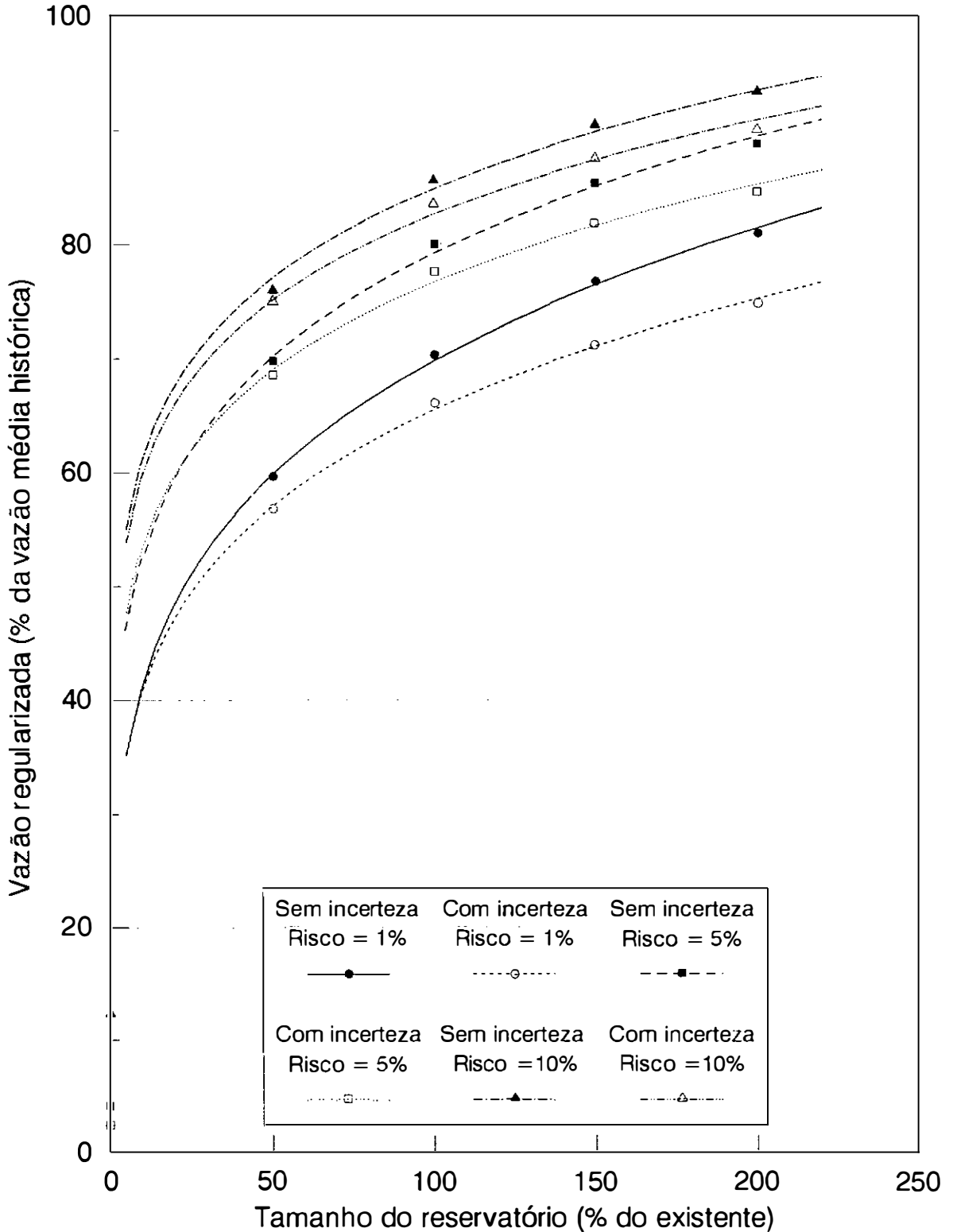


Figura 4.2-C Vazão regularizada versus tamanho do reservatório com simulação mensal e série histórica de 20 anos

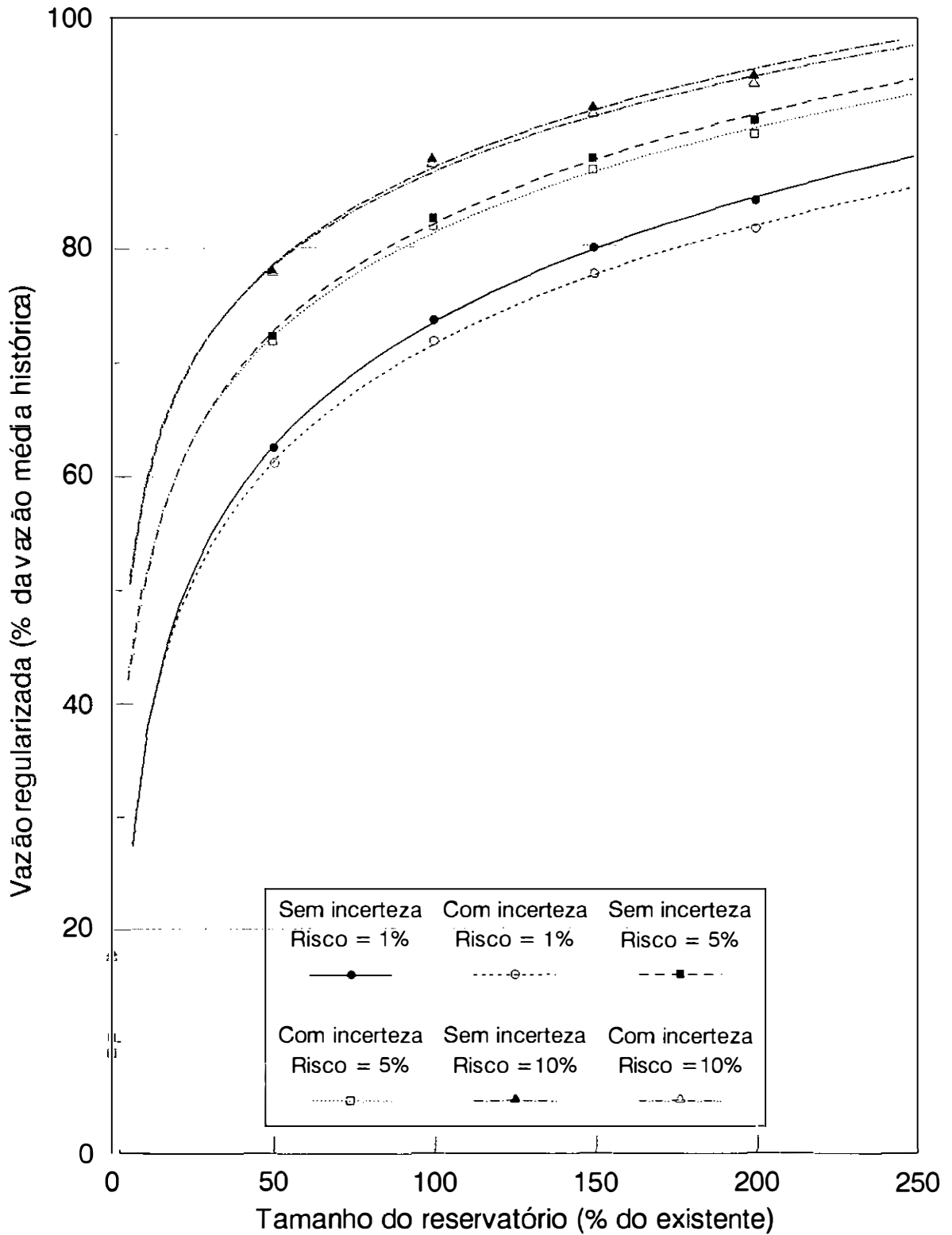


Figura 4.2-D Vazão regularizada versus tamanho do reservatório com simulação mensal e série histórica de 52 anos

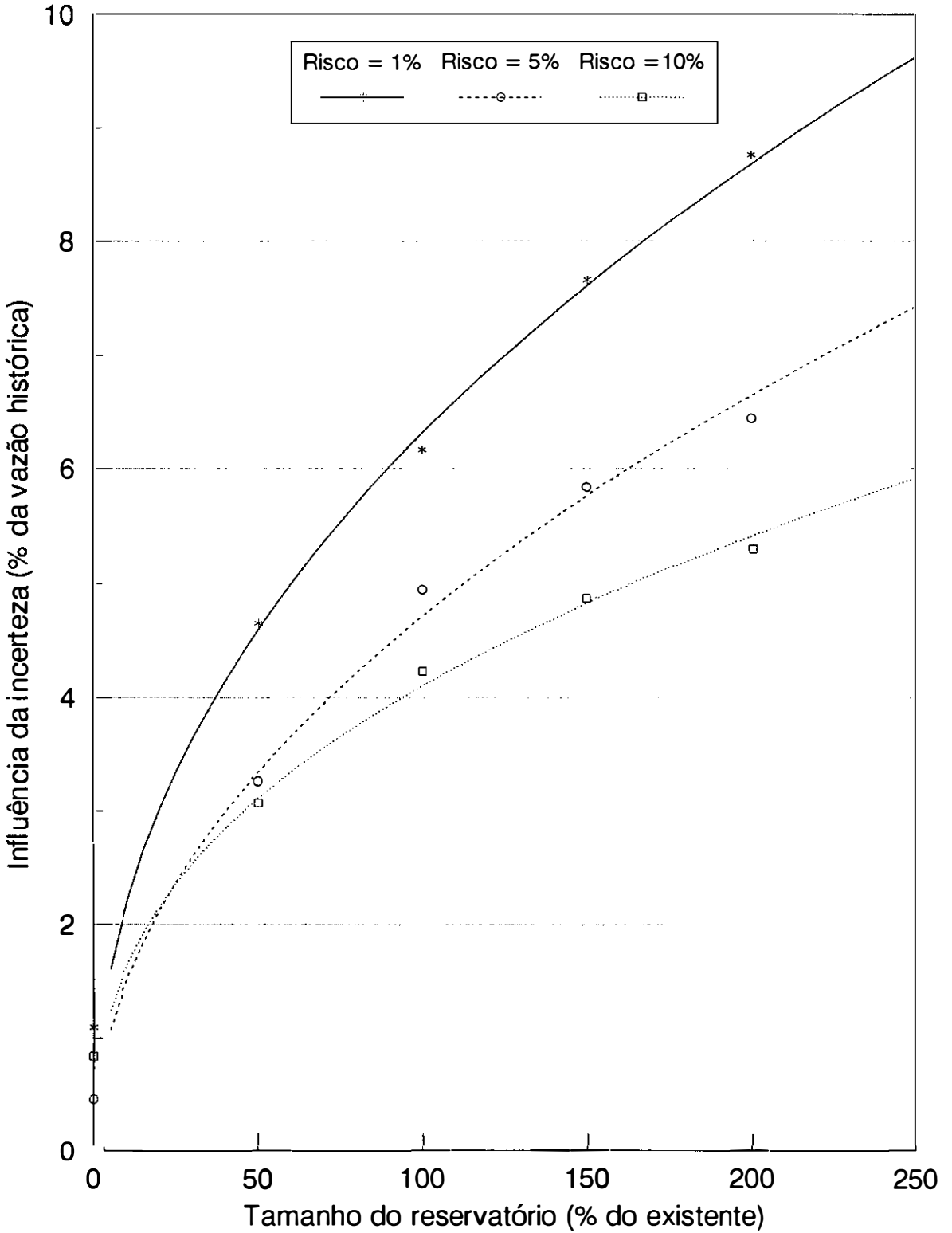


Figura 4.2-A1 Influência da incerteza dos parâmetros em função do tamanho do reservatório simulação com vazões anuais e 20 anos históricos

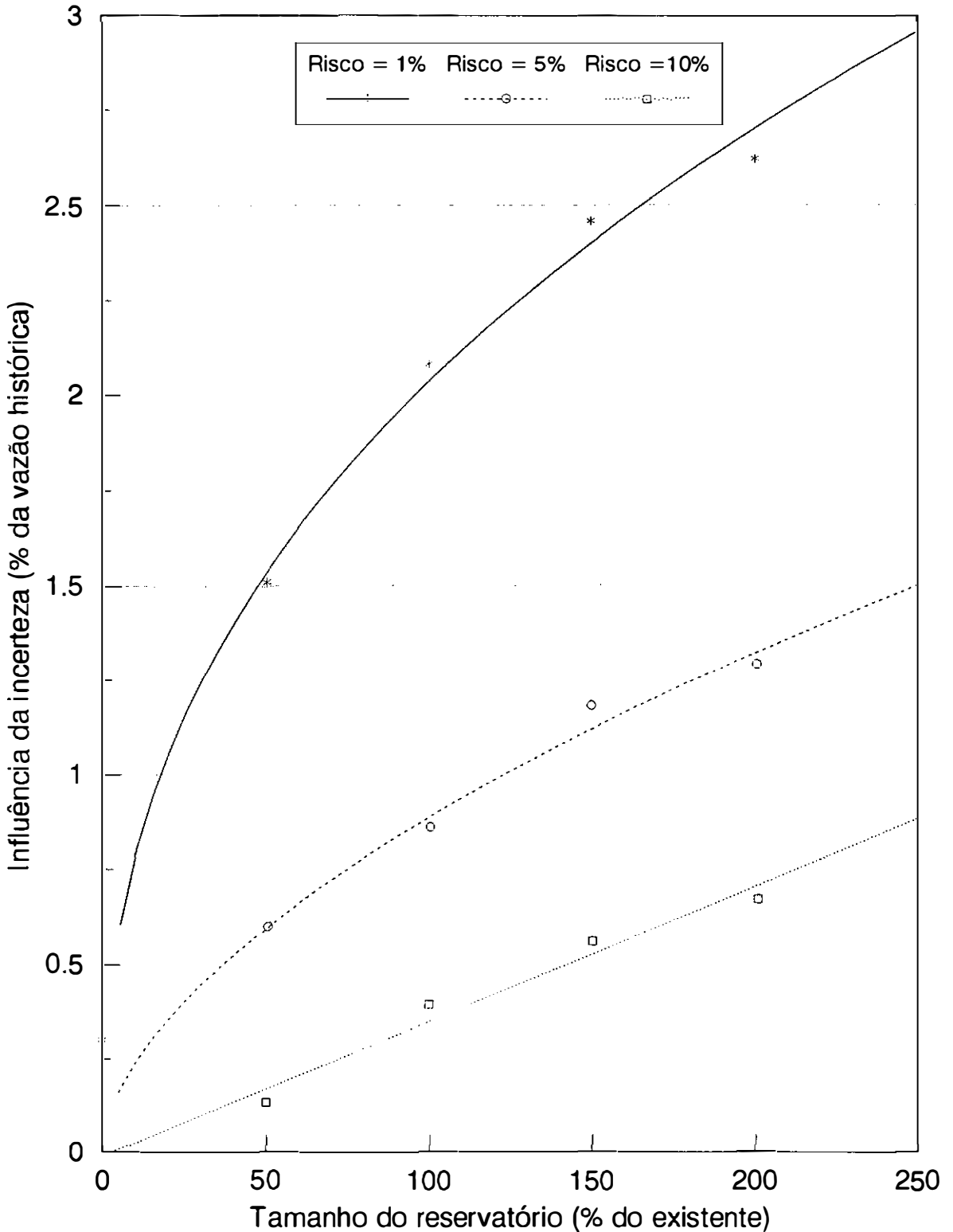


Figura 4.2-B1 Influência da incerteza dos parâmetros em função do tamanho do reservatório simulação com vazões anuais e 52 anos históricos

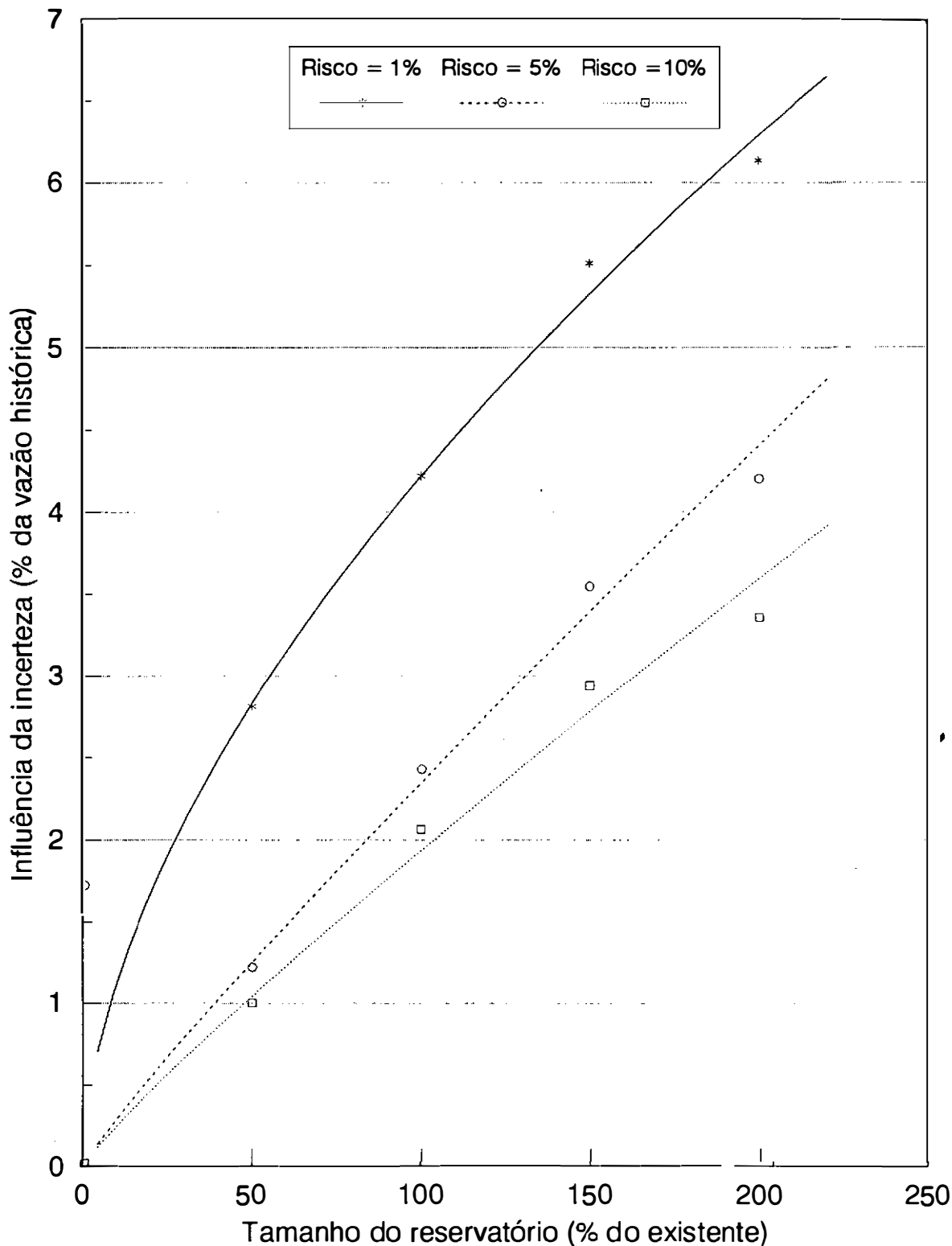


Figura 4.2-C1 Influência da incerteza dos parâmetros em função do tamanho do reservatório simulação mensal e série histórica de 20 anos

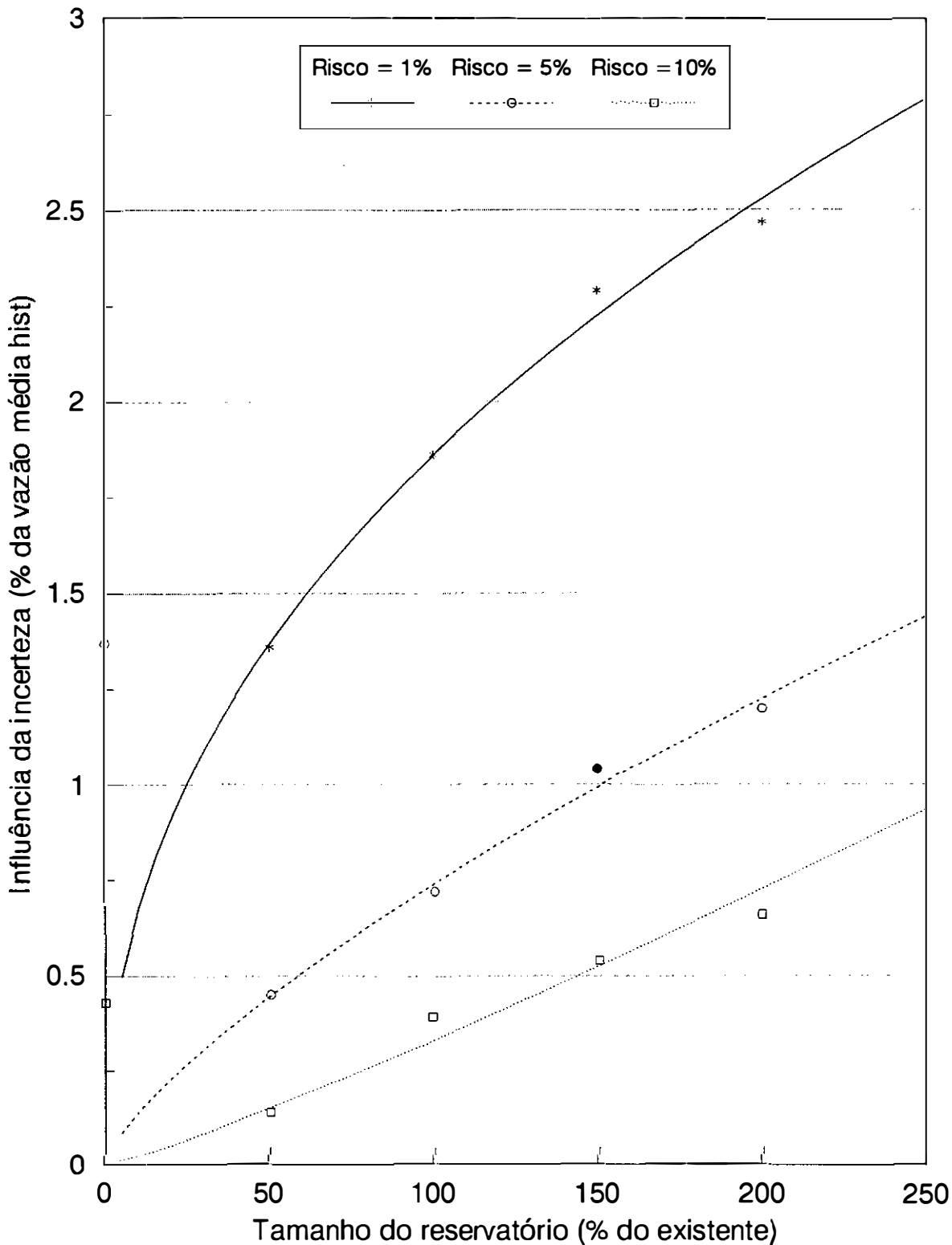


Figura 4.2-D1 Influência da incerteza dos parâmetros em função do tamanho do reservatório simulação mensal e série histórica de 52 anos

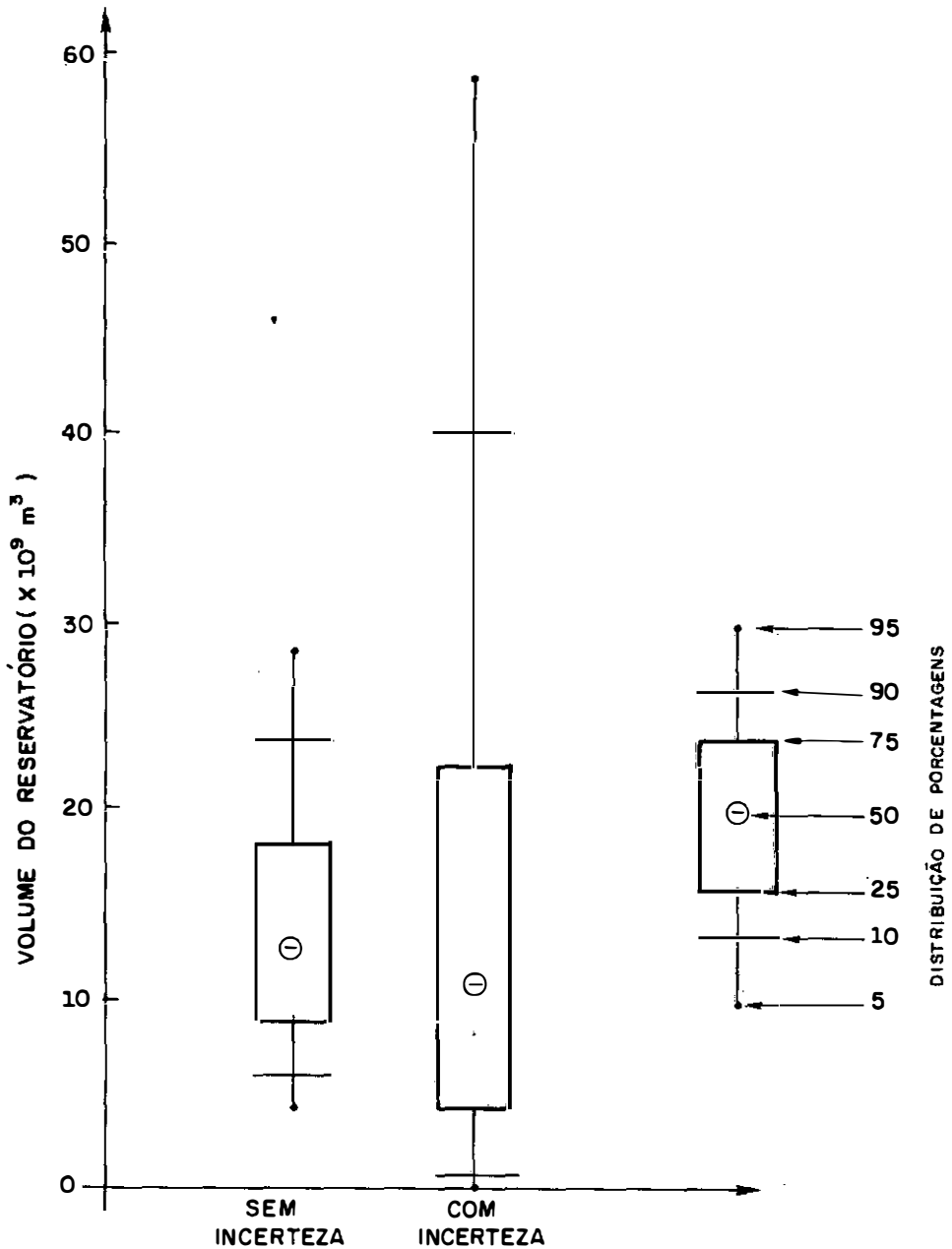


FIGURA 4.3 - DISTRIBUIÇÃO DO VOLUME DO RESERVATÓRIO NECESSÁRIO PARA REGULARIZAR 80% DA VAZÃO MÉDIA HISTÓRICA UTILIZANDO 10 ANOS DA SÉRIE HISTÓRICA E SIMULAÇÃO ANUAL

CAPÍTULO 5 CONCLUSÃO

Modelos de geração de vazões são bastante usados em estudos de simulação de reservatórios de regularização de rios. No entanto como os registros de vazões disponíveis são sempre curtos existe sempre uma considerável incerteza na estimação dos parâmetros dos modelos. A consideração da incerteza dos parâmetros na geração de séries sintéticas é feita definindo-se uma distribuição de probabilidades para os parâmetros de onde possíveis valores são sorteados. Os diversos métodos existentes para consideração da incerteza dos parâmetros podem ser classificados de acordo com o método de definição da distribuição de probabilidades dos parâmetros. O método Bayesiano, que pode ser considerado como o melhor fundamentado, tem como desvantagem a sua forte dependência da modelagem adotada. A mesma dificuldade é apresentada pelo método Clássico. O método de Reamostragem, pelo contrário, não depende da modelagem adotada, mas apresenta duas dificuldades: a limitação no número de reamostragens possíveis e a necessidade de se considerar observações independentes, o que não é o caso das séries de vazões naturais. Em geral, os esquemas de Reamostragem já desenvolvidos (Costa, 1988 e Cover, 1985) procuram suplantar este último problema incorporando no procedimento de reamostragem a estrutura do modelo adotado.

Nesta tese desenvolveu-se um procedimento para incorporar em modelos de geração de série de vazões naturais a incerteza dos parâmetros, procurando-se evitar os erros de modelagem. O procedimento se baseia num esquema de Reamostragem (semelhante ao Bootstrap) capaz de ser aplicado em séries que apresentam dependência temporal, como é o caso das séries de vazões naturais. Da série histórica se obtém várias pseudo-amostras pela reamostragem de "blocos de observações" que por hipótese são independentes. Destas pseudo-amostras obtém-se novos possíveis valores para os parâmetros.

A aplicação do procedimento proposto para um processo AR(1) foi comparado num estudo de Monte Carlo com um método Clássico proposto por Stedinger (1980). Verificou-se que os momentos das distribuições obtidas com o esquema proposto apresenta para qualquer dos valores para k (tamanho do bloco) resultados da mesma ordem de grandeza que os obtidos com o método de Stedinger. Em termos do parâmetro μ (média do processo), os valores médios das distribuições de todos os métodos foram muito próximos do valor histórico. Em

relação ao parâmetro σ (desvio padrão do processo), todas as distribuições obtidas apresentam valores médios inferiores ao valor histórico, sendo a diferença maior no método de Stedinger. Notou-se em relação ao parâmetro ρ (auto-correlação lag-1 do processo) que as estimativas obtidas pelo método de Stedinger se distribuem em torno de um valor inferior ao valor histórico o que não acontece com as estimativas obtidas pelo esquema de Reamostragem proposto.

O procedimento com $k=2$ foi detalhado para aplicações num modelo de geração de vazões mensais do tipo AR(1) anual com desagregação pelo esquema de Meija e Rousselle(1976). Notou-se que o esquema de desagregação utilizado só é aplicável a séries de observações de no mínimo 13 anos, isto para que haja pelo menos 12 pares de anos consecutivos cujos vetores de vazões mensais (incluindo dezembro do ano anterior) sejam linearmente independentes. Portanto a aplicação neste modelo do procedimento desenvolvido nesta tese sofre a mesma restrição. Como nas pseudo-amostras os pares de anos consecutivos são obtidos num esquema de Reamostragem com reposição, diversas das possíveis pseudo-amostras não conterão 12 diferentes pares de anos consecutivos, principalmente em séries pequenas e próximas a 13 anos (14, 15, 16 ..., 19 anos); note ainda que quanto maior for o intervalo de agregação k (tamanho do bloco), maior a percentagem de pseudo-amostras que possuirão menos de 12 diferentes pares de anos consecutivos. Como pseudo-amostras com menos de 12 diferentes pares de anos consecutivos não podem ser usadas para a estimação de parâmetros (em particular a matriz C) elas devem ser abandonadas. Este fenômeno limita a aplicação do método proposto já que a eficiência de qualquer método de Reamostragem depende de um grande número de possíveis pseudo-amostras.

Uma aplicação em estudos de regularização para o reservatório de Furnas no rio Grande, indicou que a consideração da incerteza dos parâmetros tem um efeito da mesma ordem de grandeza que o da consideração do intervalo de tempo mensal ao invés do anual, o que sugere a importância de incorporarmos as incertezas dos parâmetros nos modelos de geração de vazões naturais. Temos como exemplo desta importância o fato que em certos casos o risco de déficit que pensávamos ser de 5% é, considerando a incerteza dos parâmetros, de 10%. Além disto algumas tendências foram observadas: quanto menor o tamanho da série histórica maior é a influência da incerteza dos parâmetros; quanto maior o tamanho do reservatório maior a influência da incerteza dos parâmetros; para o estudo a nível de vazões mensais a influência da incerteza dos parâmetros foi menor que a nível de vazões anuais; e, por último, quanto menor o risco de déficit almejado maior a influência da incerteza dos parâmetros, que pode ser explicado

pelo fato de que em geral a inclusão da incerteza dos parâmetros em modelos probabilísticos tem maior influência nas probabilidades acumuladas para eventos mais extremos (ex: estiagens severas).

BIBLIOGRAFIA

BURGES, S. J.; LETTENMAIER, D. P.

"Reliability Measures of Water Supply Reservoirs and Significance of Long-term Persistence",

paper presented at International Symposium on Real-Time Operation of Hydrosystems, Waterloo, Ont, June/1981.

CHOW, V. T.

"The Log-probability law and its Engineering Applications",

Proc. Am. Soc. Civ. Eng. 80, Paper nº 536, 1954.

COSTA, FERNANDA S.

"Aplicação de Técnicas Estatísticas de Reamostragem em Hidrologia",

Tese de mestrado, COPPE, UFRJ, 1988.

COVER, K. A.; UNNY, T. E.

"Application of Computer Intensive Statistics to Parameters Uncertainty in Streamflow Synthesis",

Simposia on statistics in honours of professor V. W. Joshi's 70th Birthday, University of western, Ontario, Canadá, maio/1985.

EFRON, BRADLEY

"The Jackknife, the Bootstrap and others Resampling Plans",

Department of Statistics Stanford University,

Society for industrial and applied mathematics, Philadelphia, Pennsylvania,

Monograph nº 38, 1982.

EFRON, BRADLEY

"Computers and the theory of Statistics: Thinking the Unthinkable",

SIAM, Rev. 21, pp. 460-480, 1979.

KELMAN, J.; PEREIRA, M. V. F.; OLIVEIRA, G. C.; COSTA, C. C. G.

"Stochastic Streamflow Models for Hydroelectric Systems",

Water Resources Research, vol 20, no 3, pag 379-390, 1984.

KLEMES, V.

"Physically Based Stochastic Hidrologic Analysis",
Advances of Hydrosience, nº 11, Academic Press, 1978.

LOUCKS, D. P.; STEDINGER, J.R., HAITH, D.A.

"Water Resources Systems Planning and Analysis",
Prentice-Hall, Englewood Cliffs, N. J., 1981.

MATALAS, N. C.

"Mathematical Assessment of Synthetic Hidrology",
Water Resour. Res., 3(4), 937-945, 1967.

MEJIA, JOSÉ M.; ROUSSELE, JEAN

"Disagregation Models in Hydrology Revisited",
Water Resourcer Research, vol 12, no 2, pag 1985, 1976.

QUENOUILLE, M.

"Approximate Tests of Correlations in Time Series",
J. Roy. Statist. Soc., Ser. B, 11, pp 18-84, 1949.

SANGIT, C.

"Bootstrapping ARMA Model: Some Simulations",
IEEE Transactions on systems, man and cybernetics, vol SMC-16, no 2,
março/abril 1986.

STEDINGER, R. J.; TAYLOR, M. R.

"Synthetic Stream Flow Generation 2. Effect of Parameter Uncertainty",
Water Resources Research, vol 18, no 4, pag 919-924, ago/1982.

STEDINGER, J. R.; SCHAAKE, J. C. Jr.

"Adding Parameter Uncertainty to Streamflow Generating Algorithms",
Department of Environmental Engineering, Cornell University, Ithaca, N. Y.,
14853, 1980.

SUDLER, C. E.

"Storage Required for the Regulation of Stremflow",
Trans. Amer. Soc. Civil Eng., 91, 622-660, 1927.