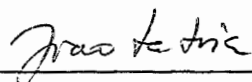


UM MODELO DETERMINÍSTICO - ESTOCÁSTICO PARA
PREVISÃO DE SÉRIES HIDROLÓGICAS TEMPORAIS

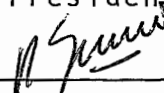
PAULO ROBERTO DE HOLANDA SALES

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE
PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JA
NEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO
GRAU DE MESTRE EM CIÊNCIAS (M.Sc.).

Aprovada por:



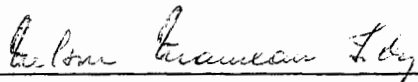
JOÃO LUIZ MAURITY SABOIA
Presidente



PEDRO GUERRERO-SALAZAR



FERNANDO YASSUO CHIYOSHI



NELSON MACULAN FILHO

RIO DE JANEIRO, RJ - BRASIL
MARÇO DE 1977

A G R A D E C I M E N T O S

À COPPE, na pessoa de seu Diretor, Prof. Sêrgio Neves Monteiro.

Aos Profs. João Luiz Maurity Saboia e Pedro Guerrero, seguros e eficientes orientadores da Tese.

Aos Profs. Fernando Yassuo Chiyoshi e Nelson Maculan Filho e ao Engenheiro Luiz Augusto Lattari Barretto, pelo exame crítico a que submeteram o manuscrito.

À ELETROBRÁS, na pessoa dos Engenheiros Fausto de Barros Pinto e Celso Ferreira, respectivamente, chefe do Departamento de Operação Energética e chefe da Divisão de Estudos de Operação pelos apoios e estímulos prestados.

Aos Engenheiros Sêrgio Mario Amado e Luiz Fernando Couto Amaro da Silva pelas valiosas contribuições prestadas no desenvolvimento computacional da Tese.

Aos Engenheiros Roberto Pereira D'Araujo e Reinaldo Castro de Souza pela gentil cessão do "package" Box e Jenkins, bem como pelas diversas sugestões importantes ao bom desenvolvimento da Tese.

Aos Engenheiros Leslie Afranio Terry e Jerson Kelman pelas contribuições aos novos testes de significância de harmônicos em parâmetros periódicos e constância de parâmetros.

À secretária Maria Helena Correia e ao desenhista Ricardo dos Santos Mattos, ambos da ELETROBRÁS, pelos excelentes trabalhos de datilografia e execução gráfica, respectivamente.

Finalmente, um agradecimento especial a meus pais, Laerte e Zélia, pelo carinho com que sempre acompanharam meus estudos.

R E S U M O

Foi desenvolvido um modelo determinístico-estocástico univariado para séries hidrológicas temporais, pela integração das metodologias apresentadas por YEVJEVICH (1972) e BOX e JENKINS (1970). A série foi transformada de acordo com BOX e COX (1964), identificando-se uma transformação logarítmica.

Foram desenvolvidos, na análise da estrutura de séries hidrológicas, novos testes para a determinação do número de harmônicos em parâmetros periódicos e constância de parâmetros, baseados na teoria dos intervalos de confiança.

Como exemplo de aplicação das técnicas desenvolvidas, utilizou-se a série temporal referente às vazões médias mensais medidas no posto fluviométrico de Salto Osório, no rio Iguaçu, no período de janeiro de 1941 a dezembro de 1971.

Previsões para 1971 (não pertencente à amostra), 1968 (ano seco), 1970 (ano médio) e 1957 (ano úmido) foram obtidas pelos diferentes modelos baseados nas metodologias de Yevjevich, Box e Jenkins, assim como na metodologia mista proposta neste trabalho. Procedeu-se, então, a uma avaliação destas previsões por três diferentes critérios, a qual mostrou a adequabilidade da metodologia mista proposta.

A B S T R A C T

An univariate deterministic-stochastic model for hydrologic time series was developed by the combination of the methodologies presented by YEVJEVICH (1972) and BOX and JENKINS (1970). The time series has been transformed according to BOX and COX (1964), and a logarithmic transformation has been identified.

New tests have been developed for the determination of the number of harmonics in periodic parameters and constancy of parameters, based on the theory of confidence intervals.

As an exemple of the application of the techniques developed, a time series has been used referring to the monthly mean streamflows taken at Salto Osório fluviometric post, at Iguaçu River, from January 1941 until December 1971.

Forecasts for 1971 (not belonging to the sample), 1968 (dry year), 1970 (average year) and 1957 (wet year) have been obtained by different models based on Yevjevich's, Box and Jenkins', as well as on the proposed methodology. Then a comparison among these forecasts was made using three different schemes. This comparison was very favorable to the mixed methodology proposed in our work.

Í N D I C E

CAPÍTULO I	- INTRODUÇÃO	1
I.1	- Importância do Estudo	1
I.2	- Objetivos do Estudo	5
CAPÍTULO II	- REVISÃO DA LITERATURA	8
II.1	- Modelos Pioneiros	8
II.2	- Modelos Baseados na Análise Es trutural Proposta Por YEVJEVICH	15
II.3	- Modelos Baseados na Metodologia de BOX e JENKINS	23
CAPÍTULO III	- INTEGRAÇÃO DO MODELO PROPOSTO POR YEVJEVICH COM OS MODELOS ARIMA DE BOX E JENKINS ...	39
III.1	- Introdução	39
III.2	- Transformação de Parâmetros	40
III.3	- Hipóteses Básicas	43
III.4	- Separação das Componentes Periódica e Estocástica	44
III.5	- Análise Harmônica em Parâmetros Periódicos	45
III.6	- Significância de Harmônicos em Parâmetros Periódicos	46

III.7	- Teste Para Constância de Parâmetros	58
III.8	- Modelos de Dependência Para a Componente Estocástica	60
III.9	- Identificação do Modelo	62
III.10	- Estimação dos Parâmetros do Modelo	76
III.11	- Testes de Verificação	80
III.12	- Previsão Para o Modelo Misto	88
CAPÍTULO IV - APLICAÇÃO E DISCUSSÃO DOS RESULTADOS		99
IV.1	- Introdução	99
IV.2	- Previsão com Modelos de YEJVJEVICH	105
IV.3	- Previsão com Modelos de BOX e JENKINS	120
IV.4	- Previsão com Modelos Mistos YEJVJEVICH - BOX e JENKINS	133
IV.5	- Avaliação das Previsões Pelos Diferentes Modelos	152
CAPÍTULO V - CONCLUSÕES E RECOMENDAÇÕES		175
BIBLIOGRAFIA		178
APÊNDICES		184

I - INTRODUÇÃO

I.1 - IMPORTÂNCIA DO ESTUDO

A escolha de um específico modelo para descrever as futuras afluências sofre realce no momento em que se pensa em planejar a operação de um sistema hidrotérmico. As séries hidrológicas de tempo representam, na verdade, uma única realização e, ainda assim parcial, de um processo onde se superpõem componentes determinísticas e estocásticas, com alguma predominância desta última. Então, o caráter aleatório das afluências futuras implica em que se encare a energia hidroelétrica, em intervalos discretos de tempo, como uma variável aleatória. Em outras palavras, no planejamento da operação de sistemas mesmo que se tenha, no instante da decisão, informações sobre as afluências aos aproveitamentos serão, basicamente, as futuras afluências que determinarão a otimização do sistema como um todo. Por este motivo e, aliado ao fato de que no Brasil existe uma grande quantidade de recursos hidráulicos e escassez de combustíveis fósseis e físséis, é importante o desenvolvimento de técnicas como previsão e geração de sequências sintéticas.

Em relação à otimização dos custos de operação em sistemas mistos hidrotérmicos com reservatórios é necessário decidir, a cada intervalo de tempo, o grau de complementação térmica adequado. Nestas ocasiões fica-se entre duas alternativas: ou se produz uma maior quantidade de energia nas usinas térmicas, incorrendo-se assim,

em gastos imediatos de combustível ou utiliza-se a água armazenada nos reservatórios.

Na primeira alternativa, a energia potencial de origem hidráulica economizada poderá ser utilizada posteriormente para assegurar o atendimento da carga ou para substituir a geração de usinas térmicas mais caras, diminuindo, portanto, os riscos de ocorrência de déficits. Deve-se não desprezar, entretanto, quando se adota esta alternativa, a hipótese de que haja vertimento em um ou mais pontos do sistema. Esta quantidade de água vertida não poderá, logicamente, ser utilizada no atendimento da carga ou na substituição das térmicas mais onerosas no futuro.

A segunda alternativa atua no sentido de se assegurar uma economia imediata representada pela redução do consumo de combustível nas usinas térmicas. Por outro lado, a diminuição da reserva hidráulica limita sua capacidade de atender a carga em casos de situação crítica, o que poderá obrigar a operação de térmicas mais dispendiosas. Se o sistema não possuir uma complementação térmica capaz de contrabalançar as adversidades hidrológicas, então, maiores serão as possibilidades de ocorrência de déficits.

Nota-se, dessa forma, a necessidade do detalhado conhecimento do comportamento aleatório das afluências futuras. Com essa informação pode-se avaliar a probabilidade de ocorrências de déficits e, portanto, atribuir-se valores esperados às políticas de operação do sistema.

Os métodos de geração de séries desempenham

também um preponderante papel nos estudos concernentes à interligação, planejamento da expansão e operação de sistemas predominantemente hidroelétricos.

De uma forma geral, três diferentes métodos para investigação dos diversos problemas relacionados com a energia hidroelétrica podem ser utilizados: o uso direto de dados históricos, a utilização de um método analítico e a geração de séries a partir das informações extraídas dos dados históricos. Este último apresenta vantagens significantes, visto que as distribuições de amostragem de várias estatísticas podem ser aproximadas pelas distribuições de frequência obtidas pela geração de séries. As informações contidas nas séries históricas podem ser, então, extraídas completamente; riscos e incertezas determinados; hipóteses e teoremas, verificados. Como aplicação particular deste método, sobressai a otimização estocástica, pelo emprego de programação dinâmica, da operação de sistemas de potência.

Outro problema importante em estudos eletroenergéticos é a distribuição da reserva hidráulica em um sistema composto de um certo número de reservatórios. Nestas situações pode-se adotar um modelo composto onde um conjunto de reservatórios é substituído por um único reservatório com energia acumulada equivalente e, as aflúências às usinas, transformadas em aflúências de energia equivalente. Para este sistema é, então, determinado o nível ótimo de geração térmica, sendo o saldo da carga sob a responsabilidade de geração hidráulica distribuído

pelas usinas.

Neste ponto deve-se notar que as afluências a cada um dos reservatórios são variáveis aleatórias, sendo necessário, pois, uma função de densidade de probabilidade n dimensional para descrever probabilisticamente as afluências, onde n representa o número de usinas. Uma solução para este caso pode ser obtida transformando-se o problema de probabilístico em determinístico, através do emprego de um modelo de simulação do sistema a usinas individualizadas, para diversas hipóteses de hidraulicidade.

Uma outra alternativa consiste na utilização de um modelo multivariado de geração de séries que compreende de uma análise da estrutura univariada para cada posto e de uma análise da estrutura multivariada para todos os postos juntos. Com a utilização de uma técnica como componentes principais, por exemplo, pode-se gerar séries. Este modelo pode permitir que se faça, também, previsões multivariadas ou mesmo que se obtenha soluções para o problema da distribuição da reserva hidráulica.

Merece realce, também, o fato de que na operação dos sistemas brasileiros, tem sido utilizado o critério do risco de déficit, fornecido pela "Curva Limite", que fornece o limite abaixo do qual as térmicas são operadas na base. A Curva Limite é determinada utilizando as sequências de descargas registradas no passado e compõe-se dos requisitos mínimos de armazenamento mês a mês a fim de garantir o atendimento da carga em período crítico. Ela res-

tringe, implicitamente, o risco de déficit a situações piores que as ocorridas no histórico. Tais eventualidades são, num consenso geral, consideradas calamitosas, admitindo-se, nestes casos, o não atendimento da carga.

Considerando que é bastante improvável a repetição da sequência de vazões históricas e consequentemente do período crítico, existe pouca informação sobre as possibilidades do comportamento futuro do sistema. Por outro lado o período crítico é uma variável aleatória que possui características importantes, além de sua duração, comumente utilizada, tais como intensidade, volume de déficit e a forma como as vazões se sucedem. Nos estudos referentes a simulação da operação tem-se recorrido ao seu valor histórico. Uma melhor solução seria colocá-lo num contexto probabilístico, ou seja, estimar sua representatividade a partir das características estruturais das séries anuais históricas.

Por todos esses motivos, torna-se desejável que se utilize modelos estocásticos de séries hidrológicas temporais nos estudos referentes à operação e planejamento da expansão de sistemas mistos hidrotérmicos pois, estes modelos podem ser empregados em conjunto com técnicas de programação matemática, no sentido de se otimizar a operação do sistema.

I.2 - OBJETIVOS DOS ESTUDOS

Este estudo tem como objetivo a apresentação de um modelo determinístico-estocástico univariado pa-

ra sêries hidrológicas de tempo, pela combinação dos modelos apresentados por YEVJEVICH (1972) e BOX e JENKINS(1970). A procura da integração dos dois modelos se prende ao facto de que o primeiro aventa somente a hipótese de modelos de dependência autoregressivos, enquanto que o segundo, no caso sazonal, só consegue representar a variabilidade no momento ordinário de primeira ordem.

O estudo visa contribuir com o planeamento da operação de aproveitamentos hidroelétricos a médio prazo, através de previsões univariadas. É importante que se ressalte por outro lado, que um modelo confiável para sêries temporais univariadas se constitui num passo prévio e imprescindível para mais adiante fazer o modelamento multivariado, visto que este é capaz de levar em conta as condições de dependência entre sêries, para um conjunto básico de postos.

Foram desenvolvidos, na análise da estrutura de sêries hidrológicas, novos testes para a determinação do número de harmônicos significativos em parâmetros periódicos e constância de parâmetros, baseados na teoria dos intervalos de confiança.

O capítulo II deste trabalho relaciona, de forma sucinta, os diversos modelos constantes da literatura sobre o assunto.

O capítulo III versa, basicamente, sobre o modelo, dando-se realce aos novos testes propostos e a maneira pela qual os dois modelos podem ser integrados para

a obtenção de um terceiro.

As aplicações com seus dados e resultados constam do capítulo IV. A série temporal hidrológica utilizada é referente as vazões médias mensais medidas no posto fluviométrico de Salto Osório, situado na bacia do rio Iguaçu, no período de janeiro de 1941 à dezembro de 1971.

Finalmente, no capítulo V são apresentadas respectivamente, as conclusões do estudo e as recomendações consideradas oportunas para o desenvolvimento do modelo.

II - REVISÃO DA LITERATURA

II.1 - MODELOS PIONEIROS

O primeiro trabalho voltado para geração de sequências sintéticas de vazões que se tem notícia foi o desenvolvido por SUDLER (1927). Selecionando 50 valores de vazões anuais de uma distribuição com um especificado coeficiente de variação e, escrevendo cada uma em um cartão, ele obteve uma sequência de 1000 vazões sintéticas após em baralhar 20 vezes a massa de cartões. Esta sequência foi usada em estudo para a determinação de possíveis tamanhos de barragens. O trabalho de Sudler representa um significativo avanço sobre outros métodos pioneiros, visto que considerou diferentes amostras de vazões, em lugar de se limitar somente aos dados históricos. Em contrapartida, como sõ considerou a mesma sequência de 50 anos a estrutura de autocorrelação dos dados históricos foi destruída.

Como é sabido, as vazões em muitos rios parecem mostrar uma certa persistência, isto é, altas vazões seguem altas vazões e baixas vazões seguem baixas vazões. Em séries com curto intervalo de observação, por exemplo hora, a persistência é óbvia, se a vazão em um rio é alta em um determinado instante é bem provável que ela ainda seja alta uma hora após este instante. Quando do aparecimento do trabalho de Sudler não havia sido desenvolvido ainda nenhum estudo sistemático para se verificar se as vazões mostravam uma persistência sobre uma longa estrutura de tem

po. HORTON (1923) apresentou alguns testes simples associados com a moderna teoria do movimento e mostrou que as vazões não derivavam de um processo sem autorrelação. Entretanto, Horton não investigou a estrutura da correlação serial.

HURST e FELLER (1951) forneceram uma contribuição decisiva à Hidrologia ao investigarem a persistência hidrológica a longo termo.

Nesta ocasião estudaram a estatística chamada de amplitude dos desvios acumulados em relação à média. Considerando dados extensivos de fenômenos naturais tais como vazão, precipitação, etc. e, usando técnicas de regressão, determinaram os valores de k na equação,

$$(2.1) \quad \frac{R\hat{n}^*}{s} = \left(\frac{n}{2}\right)^k$$

onde,

$R\hat{n}^*$ - estimador das amplitudes dos desvios acumulados em relação à média da série;

s - estimador do desvio padrão da série;

n - número de termos da série.

O conjunto de resultados para k mostrou uma média igual a 0,72 e um desvio padrão de 0,09; os valores de k excederam significativamente o valor teórico 0,5.

CHOW (1951) usou mais dados de Hurst e uma segunda regressão mais geral, da forma,

$$(2.2) \quad \frac{R\hat{n}^*}{s} = k_1 n^{k_2}$$

e encontrou valores para o expoente k_2 que excediam, também, ao valor teórico 0,5. O fato de séries de longo período de observação fornecerem expoentes diferentes de 0,5 é chamado de fenômeno de Hurst.

Hurst e Feller sugeriram, então, que se abandonasse a hipótese de que as vazões derivassem de populações com ausência de correlação serial. Hurst sugeriu, em adição, que a possibilidade de tendência a longo termo em séries naturais merecia investigação.

THOMAS e FIERING (1962) desenvolveram um modelo estocástico linear para geração de sequências sintéticas de vazões. Tomando por base dados mensais, o modelo representa as médias, os desvios padrões e o coeficiente de correlação de ordem 1 entre vazões mensais sucessivas e utiliza uma regressão linear para relacionar a vazão $Q_{t+1}(j)$ no $(t+1)$ -ésimo mês do j -ésimo ano (t sendo estimado a partir do início da sequência gerada) com a vazão $Q_t(j)$ no t -ésimo mês do j -ésimo ano. Considera-se \bar{Q}_t e \bar{Q}_{t+1} , as descargas médias mensais durante os meses t e $t+1$, S_t e S_{t+1} , os desvios padrões do registro histórico para os meses t e $(t+1)$, r_t , o coeficiente de correlação estimado entre as vazões para os meses t e $(t+1)$, respectivamente e ξ_t uma sequência de variáveis aleatórias, independentes entre si e normalmente distribuídas com média zero e variância um. O modelo de Thomas - Fiering é dado por:

$$(2.3) \quad \frac{Q_{t+1}(j) - \bar{Q}_{t+1}}{S_{t+1}} = r_t \frac{Q_t(j) - \bar{Q}_t}{S_t} + \sqrt{1 - r_t^2} \xi_t$$

Por outro lado, lembrando-se que o coeficiente de regressão estimado entre as séries de vazões dos meses t e $(t+1)$, b_t , se relaciona com o coeficiente de correlação estimado, r_t , através de,

$$(2.4) \quad b_t = r_t \frac{s_{t+1}}{s_t}$$

vem que (2.3) pode ser escrito como,

$$(2.5) \quad Q_{t+1}(j) = \bar{Q}_{t+1} + b_t(Q_t(j) - \bar{Q}_t) + s_{t+1} \sqrt{1-r_t^2} \xi_t$$

Deve-se notar, também, que se o coeficiente de correlação mensal não é significativamente diferente de uma constante, um coeficiente representando a média entre os 12 valores pode ser usado e a equação (2.3) passa a representar um modelo de 1.^a ordem da família dos autoregressivos não sazonais.

Como dito anteriormente, a equação (2.5) pode ser utilizada para gerar eventos sintéticos que se assemelham aos históricos em termos da média, desvio padrão e coeficiente de correlação de ordem 1. Para considerar a assimetria, THOMAS e FIERING (1963) substituíram a componente aleatória ξ_t por ε_t , a qual é definida como,

$$(2.6) \quad \varepsilon_t = \frac{2}{\gamma_\varepsilon} \left[\left(1 + \frac{\gamma_\varepsilon \eta_t}{6} - \frac{\gamma_\varepsilon^2}{36} \right)^3 - 1 \right]$$

onde a assimetria de ε_t , γ_ε , é relacionada com a estimativa da assimetria de um processo markoviano de ordem 1, $\hat{\gamma}_x$, por,

$$(2.7) \quad \gamma_{\epsilon} = \frac{1 - \hat{\rho}_X^3(1)}{\{1 - \hat{\rho}_X^2(1)\}^{3/2}} \hat{\gamma}_X$$

Se η_t é uma variável aleatória normalmente distribuída com média zero e variância um, então ϵ_t é aproximadamente uma variável aleatória com distribuição gama com média zero, variância unitária e assimetria γ_{ϵ} . Com ϵ_t como componente aleatória, o processo de Markov de 1.^a ordem constitui um processo estacionário de 3.^a ordem que pode ser usado para gerar eventos sintéticos que se assemelhem aos eventos históricos em termos da média, desvio padrão, assimetria e coeficiente de correlação de ordem 1.

YEVJEVICH (1964) investigou, através de análise de correlação serial, o comportamento de séries anuais de vazão, precipitação e precipitação efetiva para quatro grandes amostras. As conclusões a que chegou foram:

- a) o fator físico básico para explicar a dependência no tempo entre as séries estudadas é o transporte da água de ano para ano;
- b) não há evidência, mesmo estatística, da existência de movimentos cíclicos nas sequências destas variáveis;
- c) as flutuações nas vazões e chuvas de anos úmidos para anos secos não é explicada, de forma significativa, pelos ciclos das manchas solares;
- d) a sequência de anos úmidos e secos pode ser considerada como um processo estocástico puro;

- e) a dependência no tempo para as séries estudadas é explicada por um modelo à médias móveis;
- f) modelos lineares autoregressivos de 1.^a e 2.^a ordem se ajustam bem quando há um grande transporte de água de ano para outro.

MATALAS (1967) a partir do esquema básico de Thomas - Fiering questionou a normalidade para dados hidrológicos tanto no campo da Física quanto da Estatística (não negatividade das vazões) e considerou, também, a assimetria pela utilização das distribuições gama e log-normal de três parâmetros no ajustamento da componente independente. Nesta ocasião propôs a utilização de técnicas de regionalização e estimativas de máxima verossimilhança para minimizar as tendências associadas com as estimativas dos parâmetros que caracterizam as sequências históricas.

FIERING (1967) analisando, sob a luz do fenômeno de HURST, as séries geradas a partir do esquema de Thomas - Fiering para amostras de pequeno comprimento, típicas em registros hidrológicos, obteve resultados empíricos para determinação do expoente na equação de regressão,

$$(2.8) \quad \frac{R_n}{\sigma} = k_1 n^{k_2}$$

encontrando valores maiores que 0,5 - valor assintótico para séries independentes - mas menores que 0,72 ou 0,87, obtidos por Hurst e Chow, respectivamente.

Partindo do fato de que os modelos de Markov não explicavam completamente o fenômeno de Hurst, Fiering considerou a possibilidade de que as vazões não dependessem somente das imediatamente anteriores mas também de uma longa sequência do histórico, obtendo o modelo geral múltiplo,

$$(2.9) \quad Q_t = \beta_0 + \beta_1 Q_{t-1} + \dots + \beta_k Q_{t-k} + \sigma \sqrt{1 - r^2} \xi_t$$

onde a vazão atual, Q_t , depende linearmente de k vazões anteriores $Q_{t-1} \dots Q_{t-k}$. Os coeficientes de regressão β_0, \dots, β_k são estimados a partir da amostra e r^2 é o coeficiente de determinação encontrado na análise de regressão. Outra vez Fiering utilizou o esquema de geração para obter estimativas empíricas do coeficiente de Hurst; para modelos múltiplos encontrou expoentes da mesma ordem que o obtido por Chow.

Alguns pesquisadores como FIERING (1964) e MATALAS (1967) progrediram com os métodos de geração ao considerarem várias sequências de valores sintéticos correlacionados; tais métodos são úteis em estudos que lidam com mais de um rio onde as vazões, para um rio particular, são autocorrelacionadas no tempo e correlacionadas no espaço com as vazões de outros rios.

ASKEW, YEH e HALL (1971) fizeram um estudo comparando os resultados obtidos por diversas técnicas de geração para processos normais, observando que nenhuma delas produziu um período crítico tão o mais severo que

aquele registrado no histórico. Para regiões áridas e, considerando vazões mensais, o processo Markoviano normal de primeira ordem se mostrou melhor ajustado que um de segunda.

SINGH e LONNQUIST (1974) mostraram que na adoção do modelo básico de Thomas e Fiering seria desejável que os dados históricos fossem transformados de tal forma que se ajustassem a uma distribuição normal tão bem quanto possível. Tomando os logarítmos das vazões mensais do rio Sangamon, no período 1914 - 1970, observaram que o ajustamento de duas distribuições apresentava resultados melhores que quando se utilizava distribuições como normal ou gama.

LENTON, RODRIGUEZ - ITURBE e SCHAAKE (1974) propuseram o método de Bayes para estimação do coeficiente de autocorrelação, ρ , no modelo autoregressivo normal de ordem 1. Três diferentes métodos determinaram as funções de densidade de probabilidade marginal posterior com os estimadores de Bayes sendo obtidos a partir de uma dada função de perda. Através da comparação, pela utilização da função de risco e do risco esperado, os estimadores Bayesianos mostraram-se mais vantajosos, especialmente para o caso de pequenas amostras, que os obtidos pelos métodos dos momentos e máxima verossimilhança.

II.2 - MODELOS BASEADOS NA ANÁLISE ESTRUTURAL PROPOSTA POR YEJVJEVICH

YEJVJEVICH (1972) mostra como pode ser le-

vada a termo uma análise estrutural em séries hidrológicas de tempo mensais onde, basicamente, os seguintes pontos são realçados:

- (i) Remoção da periodicidade da série através da análise harmônica em parâmetros periódicos;
- (ii) a inferência sobre o melhor modelo de dependência (de uma certa ordem de estacionaridade) para as componentes estocásticas;
- (iii) a inferência acerca da distribuição de probabilidade que melhor se ajuste à distribuição de frequência da componente estocástica estacionária independente.

O processo gerador de uma série hidrológica temporal, X_t , pode ser representado por três componentes,

$$(2.10) \quad X_t = T_t + S_t + \epsilon_t$$

$$(2.10.a) \quad X_t = T_t + S_t + S'_t + \epsilon_t$$

A primeira, T_t , representando a tendência do processo pode ser explicada pela não homogeneidade natural ou artificial das bacias ou pela heterogeneidade (erros grosseiros) e inconsistências (erros sistemáticos) dos dados. Sua remoção pode ser feita com auxílio de uma média móvel com muitos pontos ou pelo ajuste, através do método dos mínimos quadrados, de uma função polinomial.

A segunda, S_t e/ou S'_t , responsável pela sazonalidade do processo, pode ser identificada através de parâmetros periódicos como médias, variâncias, assimetrias mensais, etc.

Removidas as duas primeiras componentes ob-
têm-se a terceira, ϵ_t , que representa um processo aproxi-
madamente estacionário e é provocada por diferentes fato-
res tais como a turbulência do ar, transmissão de calor,
opacidade do ar quando à radiação de calor, etc.

- SEPARAÇÃO DAS COMPONENTES PERIÓDICA E ESTOCÁSTICA

Desde que exista a componente T_t e esta
já tenha sido removida da série, a separação das componen-
tes periódica e estocástica pode ser feita através de dois
métodos: não paramétrico e paramétrico. Em ambos, a re-
moção da componente periódica, visando uma componente
estocástica estacionária na média e na variância, consiste
em dividir a série, subtraída da média, pelo desvio padrão.
Os dois parâmetros podem, ou não, ser periódicos. Caso se-
jam, no primeiro método eles são obtidos da própria série
enquanto que no segundo são substituídos por funções pe-
riódicas ajustadas, com um número limitado de harmônicos.

De acordo com as diversas possibilidades,
os modelos de composição, para a média e variância, podem
ser representados por:

a) média periódica e variância constante

$$(2.11) \quad X_{p,\tau} = \bar{X}_\tau + S_x \epsilon_{p,\tau}$$

onde,

$X_{p,\tau}$ - série original no ano p ($p = \overline{1,n}$) e no instante
 τ ($\tau = \overline{1,\omega}$);

$\varepsilon_{p,\tau}$ - componente estocástica no ano p e instante τ ;

\bar{X}_τ e S_{X_τ} - respectivamente, médias no período básico ω e desvio padrão de $X_{p,\tau}$.

b) média constante e variância periódica

$$(2.12) \quad X_{p,\tau} = \bar{X} + S_\tau \varepsilon_{p,\tau}$$

onde,

\bar{X} - média da série $X_{p,\tau}$;

S_τ - desvios padrão no período básico ω .

c) média e variância periódicas.

Este caso pode ser dividido em dois tipos:

c.1) coeficiente de variação constante

$$(2.13) \quad X_{p,\tau} = \bar{X}_\tau (1 + \hat{\eta} \varepsilon_{p,\tau}) = \bar{X}_\tau \gamma_{p,\tau}$$

onde,

$$(2.14) \quad \hat{\eta} = \frac{1}{\omega} \sum_{\tau=1}^{\omega} \eta_\tau$$

e,

$$(2.15) \quad \eta_\tau = \frac{S_\tau}{\bar{X}_\tau}$$

Neste caso, desde que se tenha $X_{p,\tau} > 0, \bar{X}_\tau > 0$

e $\gamma_{p,\tau} > 0$, vem:

$$(2.16) \quad L_n X_{p,\tau} = L_n \bar{X}_\tau + L_n \gamma_{p,\tau}$$

ou,

$$(2.17) \quad X_{p,\tau}^* = \bar{X}_\tau^* + \gamma_{p,\tau}^*$$

o que indica um modelo de composição simplificado, similar ao modelo (a).

c.2) coeficiente de variação periódico

$$(2.18) \quad X_{p,\tau} = \bar{X}_\tau + S_\tau \varepsilon_{p,\tau}$$

Ressalte-se, como observação, que no caso da utilização do método paramétrico a componente estocástica dependente, $\varepsilon_{p,\tau}$ ou $\gamma_{p,\tau}^*$, deve ser padronizada em $\zeta_{p,\tau}$, visto que não possui exatamente média zero e variância um. Isto se deve ao fato de que este método consiste de funções periódicas ajustadas às médias e/ou às variâncias com um número limitado de harmônicos.

Em relação aos testes de significância de harmônicos em parâmetros periódicos, Yevjevich apresenta o método de Fisher, aplicável quando no modelo de composição, em que a variância de $X_{p,\tau}$ é constante, a componente estocástica for independente e métodos empíricos, sendo um deles gráfico, pela utilização do periodograma acumulado, para os casos mais complexos de composição como os da equação (2.18); propõe, também, dois métodos de adaptação ao método de Fisher para o caso de normalidade das séries estocásticas dependentes.

O passo seguinte consiste em estimar os coeficientes de autocorrelação $R_{k,\tau}$ da componente estocástica,

obtida pela remoção da componente periódica da série $X_{p,\tau}$, os quais podem ou não ser periódicos.

Um ponto importante a ser realçado consiste no fato de que Yevjevich recomenda, nos teste para constância dos coeficientes de autorrelação, que se faça uma dicotomia nos ω valores de $R_{k,\tau}$ da amostra segundo uma técnica que vise eliminar a correlação espúria proveniente da própria forma de estimação dos coeficientes.

De acordo com o método proposto, as duas sub-amostras são obtidas da seguinte forma para $k=1$:

- a) todos os $R_{1,\tau}$, $\tau = \text{ímpar}$;
- b) todos os $R_{1,\tau}$, $\tau = \text{par}$.

Já para $k = 2$, tem-se:

- a) todos os $R_{2,\tau}$, $\tau = 1,2,5,6,9,10,\dots$
- b) todos os $R_{2,\tau}$, $\tau = 3,4,7,8,11,12,\dots$

No caso de séries hidrológicas de tempo em que os coeficientes de autocorrelação mensais não são significante diferentes de uma constante, o coeficiente de autorrelação de ordem 1 da componente estocástica pode ser estimado por,

$$(2.19) \quad R_1 = \frac{1}{12} \sum_{\tau=1}^{12} R_{1,\tau}$$

- MODELOS DE DEPENDÊNCIA PARA A COMPONENTE ESTOCÁSTICA

Em geral, para o caso de vazões mensais, os valores da componente estocástica $\zeta_{p,\tau}$ constituem, segundo Yevjevich, uma série temporal dependente, a qual pode ser satisfatoriamente descrita por um modelo linear autoregressivo de ordem m . Caso os testes rejeitem a hipótese de periodicidade nos coeficientes de autocorrelação, o modelo pode ser escrito como,

$$(2.20) \quad \zeta_{p,\tau} = \sum_{k=1}^m \phi_k \zeta_{p,\tau-k} + \sigma \xi_{p,\tau}$$

onde, ϕ_k são coeficientes autoregressivos e σ é o desvio padrão, o qual possibilita que $\xi_{p,\tau}$ seja encarada como uma variável aleatória independente, padronizada e estacionária de 2ª ordem.

Para a escolha da ordem m do processo autoregressivo, YEVJEVICH(1972) mostra um método simplificado onde são estimados os coeficientes de determinação $R_i^2, i=1,2,3,\dots$, visto que eles medem a parte da variância total de $\zeta_{p,\tau}$ que é explicada por cada um dos termos de esquema autoregressivo, à exceção do termo $\sigma \xi_{p,\tau}$. A escolha do critério envolvendo os coeficientes de determinação se prende ao fato de que $R_m^2 > \dots > R_3^2 > R_2^2 > R_1^2$.

Se $R_2^2 - R_1^2 \leq 0,01$ e $R_3^2 - R_1^2 \leq 0,02$ o modelo de 1ª ordem é selecionado, se $R_2^2 - R_1^2 > 0,01$ mas $R_3^2 - R_2^2 \leq 0,01$ seleciona-se o de 2ª ordem. O de 3ª ordem é selecionado quando $R_2^2 - R_1^2 > 0,01$ e $R_3^2 - R_2^2 > 0,01$.

- AJUSTAMENTO DE UMA DISTRIBUIÇÃO DE PROBABILIDADE \tilde{A} COM-
 PONENTE ESTOCÁSTICA INDEPENDENTE, $\xi_{p,\tau}$

O comportamento aleatório da componente es
tocástica independente pode, muitas vezes, ser razoavel-
 mente bem explicado pelo ajustamento de uma distribuição
 teórica de probabilidade. Se $\xi_{p,\tau}$ é normalmente distri-
 buída, a estacionaridade de 2ª ordem implica, também, em
 uma estacionaridade de ordem mais alta. Por outro lado,
 $\xi_{p,\tau}$ pode apresentar uma simetria não normal ou mesmo ter
 uma distribuição assimétrica. Em vista disso Yevjevich
 propõe que se tente ajustar distribuições como a NORMAL, a
 LOGNORMAL de três parâmetros e a GAMA de três parâmetros
 \tilde{A} componente estocástica independente, $\xi_{p,\tau}$.

Para a escolha da distribuição de probabi-
 lidade a ser ajustada \tilde{A} componente $\xi_{p,\tau}$ pode-se recorrer
 a testes de aderência paramétricos como o χ^2 (QUI-QUADRA-
 DO) e não paramétricos como o de KOLMOGOROV-SMIRNOV. Caso
 as hipóteses nulas de bom ajustamento sejam consistentes \tilde{A}
 especificados níveis de significância, uma das alternati-
 vas para a seleção da distribuição de probabilidade seria
 a que indicasse um menor valor de χ_{OBS}^2 .

BARRETTO (1974) desenvolveu, em uma primei-
 ra fase, uma análise estrutural de vazões mensais baseada
 nos estudos apresentados por Yevjevich; com o modelo re-
 sultante gerou várias sequências de vazões e determinou,
 para cada uma, o período crítico do reservatório. A se-
 gunda fase consistiu de uma otimização determinística da
 operação do reservatório, em cada um dos períodos críticos,

visando a máxima produção de energia, sob a restrição de descarga efluente mínima para fins de navegação. A regra ótima de operação do reservatório foi obtida, em uma terceira fase, por meio de uma análise de regressão, considerando os períodos críticos de cada sequência gerada com iguais probabilidades de ocorrência.

CEPEL-ELETOBRÁS (1976) desenvolveram um modelo de programação dinâmica estocástica para operação mensal de um sistema hidrotérmico. A adoção de um modelo a sistema equivalente permitiu converter, em energias afluentes, as sequências de descargas observadas no passado nas diversas usinas hidráulicas, tendo-se em vista futuras configurações do sistema hidráulico. Um modelo estocástico simplificado, baseados nos estudos apresentados por YEVJEVICH (1972) e BARRETTO (1974) foi então utilizado no sentido de que se fizesse uma análise estrutural, para cada mudança de configuração, visando o emprego de programação dinâmica.

O objetivo do modelo resultante, é a determinação, em cada início de mês, do total de energia hidráulica a ser produzida durante o mês, de modo a garantir a minimização do custo esperado de combustível a longo prazo.

II.3 - MODELOS BASEADOS NA METODOLOGIA DE BOX E JENKINS

Os modelos paramétricos ARIMA (Autoregressive-Integrated-Moving Average) não sazonais e sazonais

propostos por BOX e JENKINS (1970) baseados no princípio da parcimônia, ou seja, representação da série de tempo por um número mínimo de parâmetros representa um enfoque moderno no tratamento - controle e previsão - de séries históricas. Um modelo individual não sazonal pode ser autoregressivo de ordem p e/ou médias móveis de ordem q e pode utilizar a d -ésima diferença simples dos dados. Caso se infira um período de sazonalidade, s , passa-se a ter um modelo multiplicativo, onde a parte sazonal pode ser autoregressiva de ordem P e/ou médias móveis de ordem Q , bem como pode operar na D -ésima diferença sazonal dos dados. Então a família ARIMA para série de tempo pode ser classificada por sete inteiros $(p,d,q) \times (P,D,Q)_s$.

Os modelos ARIMA podem ser melhor apresentados com a introdução de alguns operadores.

(i) OPERADOR ATRASO (B) - este operador é tal que aplicado à série no tempo t , Z_t , fornece a série no tempo $t-1$. Assim,

$$B Z_t = Z_{t-1}$$

e,

$$B^m Z_t = Z_{t-m}$$

(ii) OPERADOR ADIANTAMENTO (F) - representa a operação inversa do operador B, isto é, $F = B^{-1}$. Então,

$$F Z_t = Z_{t+1}$$

e,

$$F^m Z_t = Z_{t+m}$$

(iii) OPERADOR DIFERENÇA (∇) - definido por: $\nabla = 1-B$. Logo,

$$\nabla Z_t = Z_t - Z_{t-1}$$

e,

$$\nabla^d Z_t = (1-B)^d Z_t$$

e,

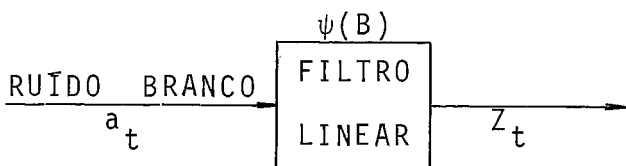
$$\nabla_s^D Z_t = (1-B^s)^D Z_t$$

(iv) OPERADOR SOMA (S) - definido como $S = \nabla^{-1}$, esse operador é tal que,

$$\begin{aligned} S Z_t &= \sum_{j=0}^{\infty} Z_{t-j} = Z_t + Z_{t-1} + Z_{t-2} + \dots \\ &= (1 + B + B^2 + \dots) Z_t \\ &= (1 - B)^{-1} Z_t \end{aligned}$$

- MODELOS LINEARES

Os modelos empregados baseiam-se na idéia de YULE (1927) de que uma série de tempo na qual sucessivos valores são altamente dependentes, pode ser considerada como gerada de uma série de "CHOQUES" a_t , aleatórios e com distribuição fixa, em geral, normal com média zero e variância finita. A sequência de variáveis aleatórias $a_t, a_{t-1}, a_{t-2}, \dots$ é chamada ruído branco. O processo de ruído branco, a_t , é suposto transformado no processo Z_t pelo que se chama de filtro linear.



O filtro linear simplesmente toma a soma ponderada de observações anteriores, isto é:

$$(2.21) \quad Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ = \mu + \psi(B) a_t$$

Em geral, μ é o parâmetro que determina o nível do processo e,

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$$

é o operador linear que transforma a_t em Z_t e, portanto, é chamado de função de transferência do filtro.

Seja,

$$(2.22) \quad \tilde{Z}_t = Z_t - \mu$$

Então, a forma inversa do modelo pode ser expressa por,

$$(2.23) \quad a_t = \tilde{Z}_t - \Pi_1 \tilde{Z}_{t-1} - \Pi_2 \tilde{Z}_{t-2} - \dots \\ = \Pi(B) \tilde{Z}_t$$

onde,

$$\Pi(B) = 1 - \Pi_1 B - \Pi_2 B^2 - \dots$$

As sequências formadas pelos pesos ψ 's e Π 's podem, teoricamente, ser finitas ou infinitas. Se a sequência ψ_1, ψ_2, \dots , é finita ou infinita convergente, o filtro é dito estável e o processo Z_t estacionário. O parâmetro μ é, então, a média sobre a qual o processo varia. De outra forma, Z_t não é estacionário e μ não tem significado específico, exceto como ponto de referência do nível

do processo. Se a sequência formada pelos pesos π_1, π_2, \dots , é finita ou infinita convergente então dizemos que o modelo é inversível.

A equação (2.21) pode ser reescrita como,

$$(2.24) \quad \tilde{z}_t = \psi(B) a_t$$

Multiplicando-se ambos os lados de (2.24) por $\Pi(B)$ e comparando com (2.23), obtêm-se:

$$(2.25) \quad a_t = \psi(B) \Pi(B) a_t$$

então a relação entre os pesos ψ^1 's e Π^1 's vale,

$$(2.26) \quad \psi(B) \Pi(B) = 1$$

ou

$$(2.26 a) \quad \Pi(B) = \psi^{-1}(B)$$

resultando que o conhecimento de $\psi(B)$ implica no conhecimento de $\Pi(B)$ e vice-versa.

- MODELOS AUTOREGRESSIVOS (AR)

Nos modelos autoregressivos o valor atual do processo é expresso como agregado linear de valores anteriores do processo e do choque a_t , ou seja,

$$(2.27) \quad \tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t$$

é chamado de processo autoregressivo (AR) de ordem p . A razão para este nome é que (2.27) é uma regressão linear sobre os valores passados da própria variável independente.

Representando o operador autoregressivo de ordem p por,

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

vem que (2.27) pode ser representado compactamente por,

$$(2.28) \quad \phi_p(B) \tilde{Z}_t = a_t$$

O modelo contém $(p+2)$ parâmetros desconhecidos: μ , ϕ_1 , ϕ_2, \dots, ϕ_p , σ_a^2 , os quais devem ser estimados a partir dos dados da série. O parâmetro adicional, σ_a^2 é a variância da série de ruído branco $\{a_t\}$.

Fazendo

$$(2.29) \quad \tilde{Z}_t = \phi_p^{-1}(B) a_t$$

a série $\{\tilde{Z}_t\}$ pode ser tratada como a saída de um filtro linear com função de transferência $\phi_p^{-1}(B)$, quando a entrada é o ruído branco a_t .

- MODELOS DE MÉDIAS MÓVEIS (MA)

Outro tipo de modelo de grande importância na representação de séries de tempo é chamado de médias móveis finita. Considerando \tilde{Z}_t linearmente dependente de um número finito, q , de ruídos brancos a_t defasados no tempo, isto é:

$$(2.30) \quad \tilde{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

é chamado de processo de médias móveis (MA) de ordem q .

Na realidade o nome médias móveis não é de todo correto visto que os pesos $1, -\theta_1, -\theta_2, \dots, -\theta_q$ não somam 1 nem precisam ser positivos.

Representando o operador média móvel de ordem q por,

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

o modelo de médias móveis pode ser representado por,

$$(2.31) \quad \tilde{Z}_t = \theta_q(B) a_t$$

Este modelo contém $(q + 2)$ parâmetros desconhecidos: $\mu, \theta, \theta_2, \dots, \theta_q, \sigma_a^2$, que devem ser estimados a partir dos dados da série.

Deve-se observar que um processo de médias móveis pode ser tratado como a saída, \tilde{Z}_t , de um filtro linear com função de transferência $\theta_q(B)$, quando a entrada é o ruído branco a_t .

- MODELOS MISTOS: AUTOREGRESSIVOS - MÉDIAS MÓVEIS (ARMA)

Uma maior flexibilidade no tratamento de séries de tempo reais, apresentando como principal vantagem a parcimônia pode ser obtida pela mistura de modelos autoregressivos e de médias móveis, isto é:

$$(2.32) \quad \tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

ou, equivalente,

$$(2.32a) \quad \phi_p(B) \tilde{Z}_t = \theta_q(B) a_t$$

O modelo contém $(p+q+2)$ parâmetros desconhecidos: μ , $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q, \sigma_a^2$, que devem ser estimados a partir dos dados da série.

Desde que se faça,

$$(2.33) \quad \tilde{Z}_t = \frac{\theta_q(B)}{\phi_p(B)} a_t = \frac{1 - \theta_1 B - \dots - \theta_q B^q}{1 - \phi_1 B - \dots - \phi_p B^p} a_t$$

um processo misto autoregressivo médias móveis pode ser tratado a saída, \tilde{Z}_t , de um filtro linear cuja função de transferência é a relação entre dois polinômios, $\theta_q(B)$ e $\phi_p(B)$, quando a entrada é o ruído branco a_t .

- MODELOS NÃO ESTACIONÁRIOS (ARIMA)

Na realidade alguma das séries encontradas exibem um comportamento não estacionário e, em particular, não variam sobre uma média fixa. Pode-se mostrar que tal comportamento pode ser representado por um operador autoregressivo generalizado, $\Psi(B)$, no qual um ou mais zeros do polinômio $\Psi(B)$ - existe uma ou mais raízes na equação $\Psi(B) = 0$ - é a unidade.

Então, o operador $\Psi(B)$ pode ser escrito,

$$\Psi(B) = \phi_p(B) (1 - B)^d$$

onde,

$\phi_p(B)$ é o operador estacionário.

Logo, o modelo geral, representando um comportamento homogêneo estacionário é da forma,

$$(2.34) \quad \Psi(B) \tilde{Z}_t = \phi_p(B) (1-B)^d \tilde{Z}_t = \theta_q(B) a_t$$

ou,

$$(2.34 a) \quad \phi_p(B) \omega_t = \theta_q(B) a_t$$

onde,

$$(2.35) \quad \omega_t = (1-B)^d \tilde{Z}_t = \nabla^d \tilde{Z}_t$$

Assim, o comportamento não estacionário pode ser representado por um modelo onde a d -ésima diferença do processo é estacionária. Na prática d usualmente é menor ou igual a 2.

O processo descrito fornece um poderoso modelo para descrever séries de tempo estacionárias e não estacionárias e é chamado de ARIMA (Autoregressive Integrated Moving Average Process) de ordem (p,d,q) .

O processo é representado por,

$$(2.36) \quad \omega_t = \phi_1 \omega_{t-1} + \dots + \phi_p \omega_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

com,

$$\omega_t = \nabla^d \tilde{Z}_t$$

A denominação "INTEGRATED" (ou somação) em ARIMA pode ser facilmente entendida pelo que se segue,

$$\omega_t = \nabla^d \tilde{Z}_t$$

ou,

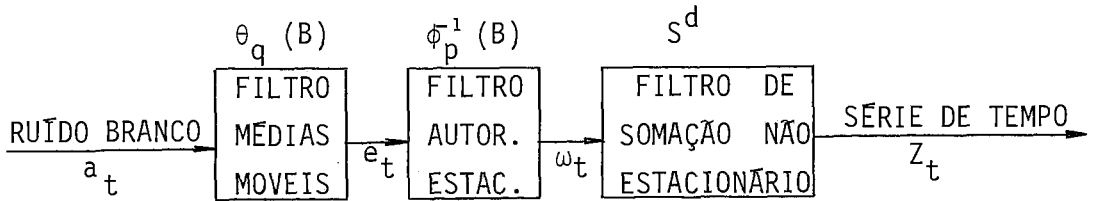
$$\tilde{Z}_t = S^d \omega_t$$

onde,

$$S \omega_t = \sum_{j=0}^{\infty} \omega_{t-j} = \omega_t + \omega_{t-1} + \omega_{t-2} + \dots$$

ou seja, $\{\tilde{z}_t\}$ é uma soma infinita de valores de $\{\omega_t\}$.

Então, o processo ARIMA pode ser gerado de um ruído branco, a_t , por meio de três operações:



O primeiro filtro possui entrada a_t , função de transferência $\theta_q(B)$ e saída e_t , onde,

$$(2.37) \quad e_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

$$= \theta_q(B) a_t$$

O segundo filtro possui entrada e_t , função de transferência $\phi_p^{-1}(B)$ e saída ω_t onde,

$$(2.38) \quad \omega_t = \phi_1 \omega_{t-1} + \dots + \phi_p \omega_{t-p} + e_t$$

$$= \phi_p^{-1}(B) e_t$$

O terceiro filtro possui entrada ω_t , função de transferência S^d e saída \tilde{z}_t .

- MODELO ARIMA SAZONAL MULTIPLICATIVO

De uma forma geral, pode-se dizer que uma série apresenta um comportamento periódico com período s , quando similaridades ocorrerem após s unidades de tempo. Normalmente, na maioria das séries históricas, o intervalo

básico de tempo é tomado como sendo igual a um mês, o que conduz a um período $s=12$ meses. No caso específico de séries hidrológicas de tempo, diárias, semanais ou mensais, sempre está presente uma componente periódica com um ciclo fundamental de um ano. Isto se deve, basicamente, à rotação da terra em torno do sol (ciclos astronômicos).

Desde que se tenha uma série exibindo um comportamento sazonal com período s , conhecido, pode-se arrumar os dados referentes à série em uma tabela contendo s colunas. Arrumada desta forma, a tabela enfatiza o fato de que não existe um, mas dois intervalos de tempo importantes: um correspondendo aos meses e o outro aos anos. Espera-se, dessa forma, que exista uma relação entre observações para sucessivos meses em um particular ano e entre observações para um mesmo mês em anos sucessivos.

Para fixação de idéias, admita-se que Z_t seja uma observação referente a setembro de um particular ano em uma série de tempo mensal. Pode-se esperar que este valor tenha alguma relação com as observações tomadas nos meses de setembro dos anos anteriores, no consenso de que, em todas observações, espera-se uma divergência na mesma direção em relação à tendência a longo termo. Tal relação pode ser representada por,

$$(2.39) \quad \Phi_p (B^s) \nabla_s^D \tilde{Z}_t = \Theta_Q (B^s) \alpha_t$$

onde,

$s=12$ - período de sazonalidade;

$\nabla_S^D = (1 - B^S)^D$ - operador diferença sazonal de ordem D capaz de produzir estacionaridade nos dados;

$\Phi_P (B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_P B^{PS}$ - operador autorregressivo sazonal;

$\Theta_Q (B^S) = 1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS}$ - operador médias móveis sazonal e,

$\{\alpha_t\}$ é uma sequência de variáveis aleatórias que será melhor entendida adiante.

É razoável esperar-se que o mesmo tipo de relação sirva para as observações de outros meses; dados de agosto, por exemplo, podem ser relacionados por,

$$(2.40) \quad \Phi_P (B^S) \nabla_S^D \tilde{z}_{t-1} = \Theta_Q (B^S) \alpha_{t-1}$$

Além disso, pode-se também supor que os parâmetros de $\Phi_P (B^S)$ e $\Theta_Q (B^S)$, contidos nestes modelos mensais, sejam aproximadamente os mesmos para cada mês. Entretanto, α_t pode não ser independente de α_{t-1} . Dessa forma, para se levar em conta tais relações, um segundo modelo é introduzido,

$$(2.41) \quad \phi_p (B) \nabla^d \alpha_t = \theta_q (B) a_t$$

onde,

$\phi_p (B)$ e $\theta_q (B)$ são polinômios em B nos graus p e q , respectivamente, e $\nabla^d = (1 - B)^d$ é o operador diferença não sazonal de ordem d , capaz de produzir uma estacionaridade nas d -ésimas diferenças de $\{\alpha_t\}$.

Substituindo-se (2.41) em (2.39), obtêm-se,

$$(2.42) \quad \phi_p(B) \Phi_P(B^S) \nabla^d \nabla_S^D \tilde{Z}_t = \theta_q(B) \Theta_Q(B^S) a_t$$

Esta equação representa o MODELO ARIMA GERAL da família de modelos multiplicativos sazonais de ordem $(p,d,q) \times (P,D,Q)_S$.

Finalmente cabe ressaltar que quando se aplica os modelos ARIMA de Box e Jenkins a uma série particular de dados o primeiro passo consiste na IDENTIFICAÇÃO do modelo. A seguir passa-se a ESTIMAÇÃO dos parâmetros do modelo e, após esta fase, faz-se uma VERIFICAÇÃO, visando possíveis inadequabilidades.

Entre os diversos trabalhos empregando a metodologia proposta por BOX e JENKINS (1970) pode-se destacar, inicialmente, o desenvolvimento por CARLSON (1970) para quatro séries de vazões anuais, onde identificou modelos ARIMA não sazonais. Para duas das séries ajustou um modelo autoregressivo de 1.^a ordem. Para as outras duas, um autoregressivo de 2.^a ordem e um misto, autoregressivo - médias móveis de 1.^a ordem.

O'CONNELL (1971) sugeriu que a adoção de modelos ARIMA em geração de séries de vazões poderia explicar o fenômeno de Hurst. Mostrou, também, que o modelo $(1,0,1)$, com $0,80 \leq \phi_1 \leq 0,99$ e $0,5 \leq \theta_1 \leq 0,95$, comparado a um processo autoregressivo era capaz, em certos casos, de preservar o expoente de Hurst suficientemente maior que o valor assintótico 0,5.

MC MICHEL e HUNTER (1972) mostraram, também, uma aplicação da metodologia proposta por Box e Jenkins, em hidrologia. Em particular, estudaram o fenômeno da temperatura diária dos rios. Com dados diários de temperatura do rio Ohio no período de 1 de janeiro de 1963 a 31 de dezembro de 1968, identificaram um modelo ARIMA $(1,0,1) \times (1,0,1)_{365}$.

MC KERCHAR e DELLEUR (1974) mostraram o procedimento a ser adotado para se ajustar um modelo ARIMA sazonal multiplicativo aos logaritmos de vazões mensais. Estabeleceram, também, neste trabalho, comparações entre as previsões fornecidas pelo modelo $(2,0,0) \times (0,1,1)_{12}$ com aquelas obtidas a partir do modelo convencional de Thomas-Fiering. Estas comparações foram, de certa forma, prejudicadas, porque o modelo ARIMA empregado não considerava a periodicidade dos desvios padrões mensais. Logo, os erros padrões das previsões não eram fisicamente corretos e, portanto, os logaritmos não poderiam ser corretamente transformados em vazões naturais.

ARAUJO (1974) e SOUZA (1974) apresentaram trabalhos baseados na teoria de modelagem de séries temporais, proposta por Box e Jenkins. Araujo dá ênfase ao problema de transformação e estimação de parâmetros, enquanto que Souza ressalta a parte relacionada à identificação e testes estatísticos. Os trabalhos apresentam um conjunto de programas visando, como aplicação prática das técnicas desenvolvidas o caso particular da série hidrológica de vazões mensais do rio Grande, no período 1931-1970.

Em continuidade aos trabalhos de Araujo e Souza, FURNAS (1975) desenvolveu um trabalho visando previsões pelo método de Box e Jenkins calcado na série de vazões afluentes ao reservatório de FURNAS, de janeiro de 1931 à dezembro de 1970. O desempenho do modelo resultante foi testado, então, no período de janeiro de 1971/ fevereiro de 1975.

FURNAS (1976) dá seguimento ao trabalho anteriormente publicado, fazendo uma avaliação das previsões no histórico de afluências ao mesmo reservatório. Nesta ocasião, foram testadas as performances do modelo em anos "médios", "secos", "de cheias" e "de período crítico". Notou-se maus resultados apenas quando a série mudava bruscamente de inclinação.

No mais extenso e completo trabalho até hoje publicado sobre a modelagem de séries hidrológicas de tempo pelo método Box-Jenkins, HIPEL (1975) desenvolveu a modelagem ARIMA, fenômeno de Hurst, preservação de estatísticas amostrais, análise de intervenção e projeto de reservatório através de simulação.

No primeiro tópico mostrou, em adição aos procedimentos tradicionais, como se poderia utilizar as funções inversas de autocorrelação e autocorrelação parcial para uma correta identificação do modelo. No segundo, apresentou o fenômeno de Hurst como resultado de uma interpretação incorreta. No terceiro, ressaltou o fato de que os modelos ARIMA preservaram importantes estatísticas do

histórico, como o coeficiente de Hurst, a função de autocorrelação e as propriedades espectrais. No quarto, procedeu a uma análise das mudanças dos níveis médios das séries de tempo, devidas as intervenções como construções de reservatórios e incêndios nas florestas. Finalmente, no último ítem, utilizou dados sintéticos, de diferentes tipos de modelos estocásticos, para fazer análises econômicas de projetos de reservatórios.

SILVA e SALES (1975) apresentaram uma aplicação de modelos estocásticos lineares para séries hidrológicas de vazões mensais. Pode-se dizer, na verdade, que uma parte deste trabalho se constituiu na primeira tentativa de se integrar os modelos de Yevjevich, Box e Jenkins, com o objetivo de se fazer previsões univariadas.

III - INTEGRAÇÃO DO MODELO PROPOSTO POR YEVJEVICH COM OS MODELOS ARIMA DE BOX E JENKINS

III.1 - INTRODUÇÃO

Na maioria das séries hidrológicas de tempo sempre está presente uma componente periódica com um ciclo fundamental de um ano. Logo, uma alternativa para se modelar tais séries consiste em se adotar um modelo misto, composto de uma parte determinística e outra estocástica. O problema passa a ser, então, o da escolha de um determinado procedimento que transforme as séries ciclicamente não estacionárias em estacionárias. De acordo com o critério da parcimônia um destes procedimentos pode ser o de se fazer uma análise harmônica nos parâmetros periódicos e representá-los por uma função matemática com um número limitado de harmônicos.

Desde que os efeitos dos ciclos determinísticos tenham sido removidos da série, a metodologia ARIMA de identificação, estimação e testes de verificação pode ser aplicada no sentido de se obter um modelo linear para a componente estocástica.

Neste capítulo, a seção III.2 trata da transformação de parâmetros proposta por Box e Cox. As seções de III.3 à III.7, da análise na estrutura periódica da série transformada. Um novo método para computação dos harmônicos significativos em parâmetros periódicos é mostrado na seção III.6. Na seção III.7 é feita uma nova proposta

para se testar a constância de parâmetros.

A seção III.8 trata da integração entre as duas metodologias. As de III.9 a III.11 estão relacionadas com a metodologia ARIMA de Box e Jenkins.

Finalmente, na seção III.12 enfatiza-se a utilização do modelo misto, no sentido de se obter previsões univariadas.

III.2 - TRANSFORMAÇÃO DE PARÂMETROS

Frequentemente pode acontecer que a série particular de dados $\{Z_t, t = \overline{1, N}\}$, que se queira analisar, possua uma distribuição assimétrica ou mesmo uma simetria não normal. Dentro do escopo de Box e Jenkins, hipóteses como homocedasticidade, independência e normalidade são postuladas à série de resíduos $\{a_t\}$. Como as variáveis Z_t e a_t se relacionam linearmente, uma adequada transformação pode ser feita à primeira, para que as hipóteses anteriormente citadas, sejam inerentes à segunda.

A normalidade dos resíduos é uma hipótese útil para aplicação de propósitos. Se o modelo se ajusta bem aos dados, passando por todos os testes de verificação e possui uma transformação que satisfaça a condição de normalidade, então as estatísticas dos dados históricos são preservadas nas séries simuladas a partir do modelo. Por outro lado, sem a hipótese de normalidade, o cálculo dos intervalos de confiança para dados previstos, poderia se tornar não adequado para propósitos práticos.

BOX e COX (1964) sugeriram, entre outras, as seguintes famílias de transformações:

$$(3.1) \quad Z_t^{(\lambda)} = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda} & , \lambda \neq 0 \\ \text{Ln } Z_t & , \lambda = 0 \end{cases}$$

onde $Z_t > 0$ e $|\lambda| \leq 1$

e,

$$(3.2) \quad Z_t^{(\lambda)} = \begin{cases} \frac{(Z_t + \lambda_2)^{\lambda_1} - 1}{\lambda_1} & , \lambda_1 \neq 0 \\ \text{Ln } (Z_t + \lambda_2) & , \lambda_1 = 0 \end{cases}$$

onde, $(Z_t + \lambda_2) > 0$ e $|\lambda_1| \leq 1$

No presente trabalho a família de transformações a ser aplicada é aquela dada pela equação (3.2).

Sem perda de generalidade, pode-se fazer $\lambda_2=0$ na segunda família de transformações, recaindo-se, dessa forma, na primeira família. Então, em princípio, para

$\lambda=0$ a expressão $\frac{Z_t^\lambda - 1}{\lambda}$ parece ser indeterminada. Entre

tanto, deve-se notar que qualquer número positivo, Z, pode ser escrito como,

$$(3.3) \quad Z = e^{\text{Ln } Z}$$

mas,

$$(3.4) \quad e^{\text{Ln } Z} = 1 + \text{Ln } Z + \frac{1}{2!} (\text{Ln } Z)^2 + \frac{1}{3!} (\text{Ln } Z)^3 + \dots$$

Logo,

$$(3.5) \quad \frac{Z_t^\lambda - 1}{\lambda} = \text{Ln } Z_t + \frac{\lambda}{2} (\text{Ln } Z_t)^2 + \frac{\lambda^2}{3} (\text{Ln } Z_t)^3 + \dots$$

Então, para $\lambda = 0$, tem-se:

$$\frac{Z_t^\lambda - 1}{\lambda} = \text{Ln } Z_t$$

A figura a seguir ilustra a família de transformações da equação (3.1). Nela, Z_t representa uma variável cuja distribuição de probabilidade é assimétrica positiva e, $Z_t^{(\lambda)}$, a variável transformada normalmente distribuída.

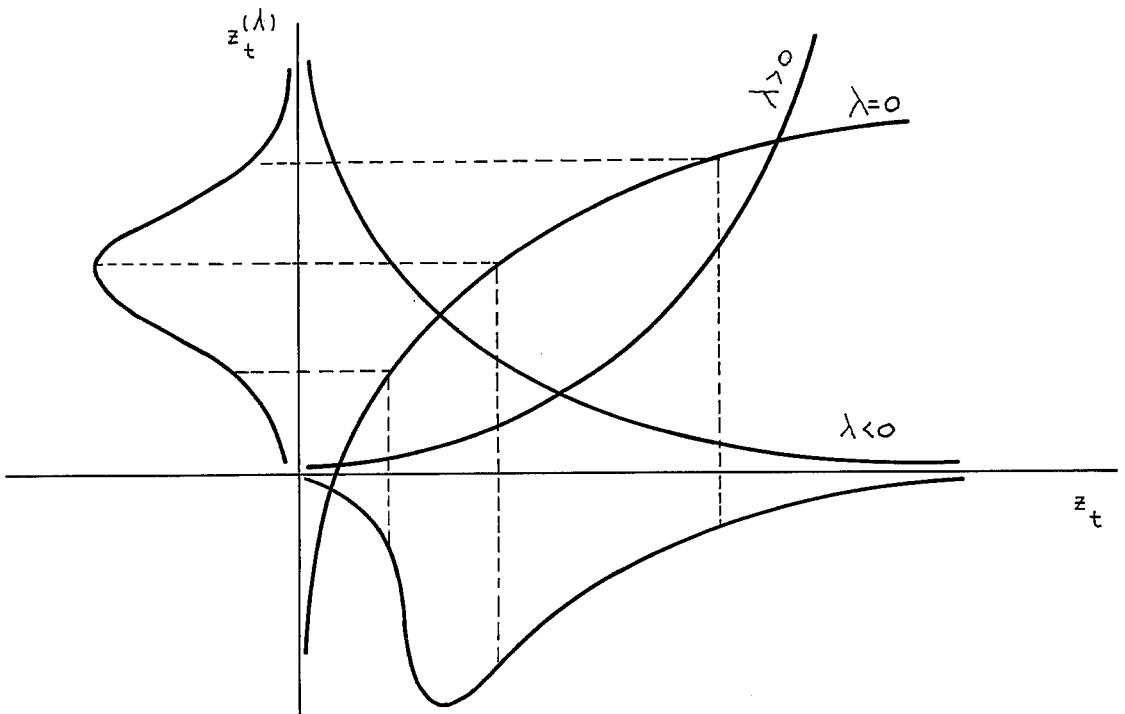


FIG. 3.1 - FAMÍLIA DE TRANSFORMAÇÕES DA EQUAÇÃO (3.1)

Finalmente, pode-se dizer, que o objetivo de se aplicar a família de transformações da equação (3.2), à série original, Z_t , foi o de se obter uma distribuição normal para os resíduos. Isto porque, na prática, quando

este objetivo é alcançado, em geral se está caminhando no sentido de que a variância da série $\{a_t\}$ seja constante. Em relação à independência dos resíduos, testes suplementares, desenvolvidos em III.11, confirmam ou não esta hipótese.

Considerando $Z_t^{(\lambda)}$ a variável transformada, então, um teste de aderência como o qui-quadrado, permite obter um par de valores ótimos $(\lambda_1^* ; \lambda_2^*)$ de tal forma que a distribuição desta variável se aproxime o mais possível da distribuição normal.

III.3 - HIPÓTESES BÁSICAS

YEVJEVICH (1972) relaciona um conjunto de hipóteses básicas a serem feitas antes de se proceder a uma análise na estrutura de uma série hidrológica de tempo. Quanto a estas hipóteses pode-se destacar em relação ao modelo integrado:

- a) uma série hidrológica de tempo é constituída de componentes determinísticas (na forma de parâmetros periódicos) e uma componente estocástica. A hipótese básica é que a série pode ser separada nestas componentes;
- b) as não homogeneidades (erros grosseiros) e inconsistências (erros sistemáticos) devem ser detetados e removidas antes que se faça a análise estrutural da série de tempo;
- c) a remoção das componentes periódicas nos vários parâmetros da série ou a remoção da ciclicidade conduz a

- hipótese de que a componente estocástica da série \bar{e} , aproximadamente, uma variável aleatória estacionária de certa ordem;
- d) um mínimo de coeficientes e/ou parâmetros devem ser estimados;
- e) a componente estocástica estacionária dependente ajusta-se um modelo matemático de dependência, objetivando-se a determinação da componente estocástica estacionária independente.

III.4 - SEPARAÇÃO DAS COMPONENTES PERIÓDICA E ESTOCÁSTICA

Antes de se proceder, efetivamente, à separação das componentes periódica (S_t) e estocástica (ϵ_t), deve-se pesquisar a existência ou não de tendências na série original. A remoção da componente de tendência (T_t), caso exista, pode ser feita de acordo com um dos procedimentos apresentados no capítulo anterior.

Para separar as componentes S_t e ϵ_t foi adotado o método paramétrico, descrito em II.2, em virtude da hipótese d.

Testes de constância - aqui baseados na teoria dos intervalos de confiança - em parâmetros como média, variância e coeficiente de variação se constituem no fator preponderante para seleção do melhor modelo de composição. As equações de (2.11) a (2.18) são representativas dos modelos alternativos de composição para a média e variância.

A componente estocástica $\bar{\epsilon}$ é então obtida a partir de uma destas equações.

III.5 - ANÁLISE HARMÔNICA EM PARÂMETROS PERIÓDICOS

A componente periódica de um parâmetro π pode ser aproximada por m harmônicos de um período básico ω , da seguinte forma,

$$(3.6) \quad \pi_{\tau} = \pi_x + \sum_{j=1}^m \{ A_j \cos(\lambda_j \cdot \tau) + B_j \sin(\lambda_j \cdot \tau) \}, \tau = \overline{1, \omega}$$

onde,

π_{τ} - função periódica ajustada aos ω valores do parâmetro p_{τ} , da amostra;

π_x - médias dos π_{τ} valores ajustados aos ω valores de p_{τ} da amostra ou média destes ω valores;

m - número de harmônicos significativos no período básico ω , $m = 1, 2, \dots, \omega/2$ ou $(\omega-1)/2$, se ω é ímpar;

$\lambda_j = \frac{2\pi j}{\omega}$ - frequência angular;

$C_j = \sqrt{A_j^2 + B_j^2}$ - amplitude do harmônico j ;

$\phi_j = \text{TAN}^{-1} \frac{B_j}{A_j}$ - fase do harmônico j ;

A_j e B_j - coeficientes de Fourier, estimados a partir dos valores de p_{τ} pelas fórmulas,

$$(3.7) \quad A_j = \frac{2}{\omega} \sum_{\tau=1}^{\omega} p_{\tau} \cos \left(\frac{2\pi j \cdot \tau}{\omega} \right)$$

$$(3.8) \quad B_j = \frac{2}{\omega} \sum_{\tau=1}^{\omega} p_{\tau} \sin \left(\frac{2\pi j \cdot \tau}{\omega} \right)$$

para o último harmônico, $j = \omega/2$ ou $j = (\omega-1)/2$, os coeficientes são estimados por,

$$(3.9) \quad A_{\omega/2} = \frac{A_j}{2}$$

$$(3.10) \quad B_{\omega/2} = 0$$

III.6 - SIGNIFICÂNCIA DE HARMÔNICOS EM PARÂMETROS PERIÓDICOS

A importância de um método para determinar o número de harmônicos significativos em parâmetros periódicos está diretamente relacionada com a hipótese de parcimônia.

O desenvolvimento de um novo método baseado na teoria dos intervalos de confiança, se prendeu ao fato de que a literatura especializada ou apresenta métodos empíricos ou métodos que são podem ser utilizados nos modelos mais simples de composição, como foi explicado anteriormente.

Suponha-se, então, que um parâmetro π seja periódico num período básico ω e que, portanto, se tente representar este parâmetro por uma função periódica. O procedimento tradicional consiste em se ajustar uma série de

Fourier aos ω valores de π da amostra, sendo o número máximo de harmônicos igual a $\omega/2$ ou $(\omega-1)/2$. No entanto, como as séries de Fourier são rapidamente convergentes, os harmônicos de ordem superior a quatro, para séries mensais podem ser, na maioria das vezes, considerados não significativos. O teste ora proposto visa determinar a ordem, a partir da qual, os harmônicos podem ser desprezados.

A filosofia do teste é bastante singela. Baseia-se no fato de que as ω estimativas de π sejam encaradas como uma amostra efetiva. Desde que o estimador de π é uma variável aleatória então, cada uma das ω observações representa, na verdade, uma realização dessa variável e pode-se, dessa forma, determinar intervalos que, com uma probabilidade especificada, contenham os verdadeiros e desconhecidos valores do parâmetro π . Obtidos os intervalos, fica determinado, também, uma faixa de confiança para o parâmetro no período fundamental ω . Então, o parâmetro π é estimado com o número máximo de harmônicos e, com os harmônicos classificados em ordem decrescente de variância, vai-se computando o parâmetro sucessivamente, pela subtração de um harmônico de cada vez, no sentido de menos para o mais significativo, até uma certa ordem, a partir da qual a função harmônica ajustada não esteja integralmente contida na faixa de confiança determinada. Em consequência, para a ordem anterior a esta assim determinada, a função periódica ajustada estará contida na faixa de confiança. Esta ordem será considerada aqui como limi

trofe, sendo, então, o número de harmônicos para o parâmetro, dado pela ordem anterior a esta.

Com relação a classificação dos harmônicos em ordem de importância basta notar que o valor $\frac{C_j^2}{2}$, onde $C_j = \sqrt{A_j^2 + B_j^2}$ representa a amplitude do harmônico j , mede a contribuição do harmônico de ordem j para a variância total. Entretanto, o problema maior reside em se obter os intervalos de confiança de $(1-\lambda)\%$, que se apresenta a seguir.

Seja E um evento aleatório com probabilidade λ de ocorrer numa dada experiência e admita-se que esta probabilidade se mantenha constante para n experiências independentes, realizadas nas mesmas condições. A probabilidade para que o evento E ocorra, pelo menos uma vez, é dado por,

$$(3.11) \quad \alpha = 1 - (1-\lambda)^n$$

Suponha-se, agora, o seguinte teste de hipóteses:

H_0 : o parâmetro π é periódico

H_1 : o parâmetro π não é periódico

Admitindo-se a hipótese nula como verdadeira e fixando-se a probabilidade α em digamos 1%, 5%, etc, tem-se o nível de significância do teste, ou seja, a probabilidade de se cometer um erro do tipo I, no sentido de que a função periódica deva estar contida na faixa de con

fiança do parâmetro. O valor α será, então, a probabilidade de que, por mero acaso, sendo o parâmetro periódico, a função harmônica ajustada caia, pelo menos uma vez, fora da faixa de confiança determinada. Desde que se especifique α , obtêm-se λ e, portanto, os intervalos de confiança desejados.

Caso o intervalo de confiança seja mensal e o nível de significância seja fixado em 5%, tem-se:

$$0,05 = 1 - (1-\lambda)^{12}$$

$$\lambda = 0,004265 \cong 0,4\%$$

Um procedimento análogo consiste em se fixar λ (ou $1-\lambda$) e obter o correspondente valor de α . Pode-se, então, construir a curva característica de operação ou curva OC do teste (figura 3.2).

Deve-se notar, entretanto, que no caso de o parâmetro π ser o coeficiente de correlação mensal, uma leve mudança deve ser introduzida no teste. Para evitar a correlação espúria explicada no capítulo anterior, o intervalo de confiança, λ , em função do nível de significância, α , é dado por,

$$(3.12) \quad \lambda = 1 - (1-\alpha)^{1/6}$$

A curva característica de operação, para este caso, é representada pela figura 3.3.

Fixado α (ou λ), passa-se a obtenção dos intervalos de confiança para as médias, variâncias (d. pa-

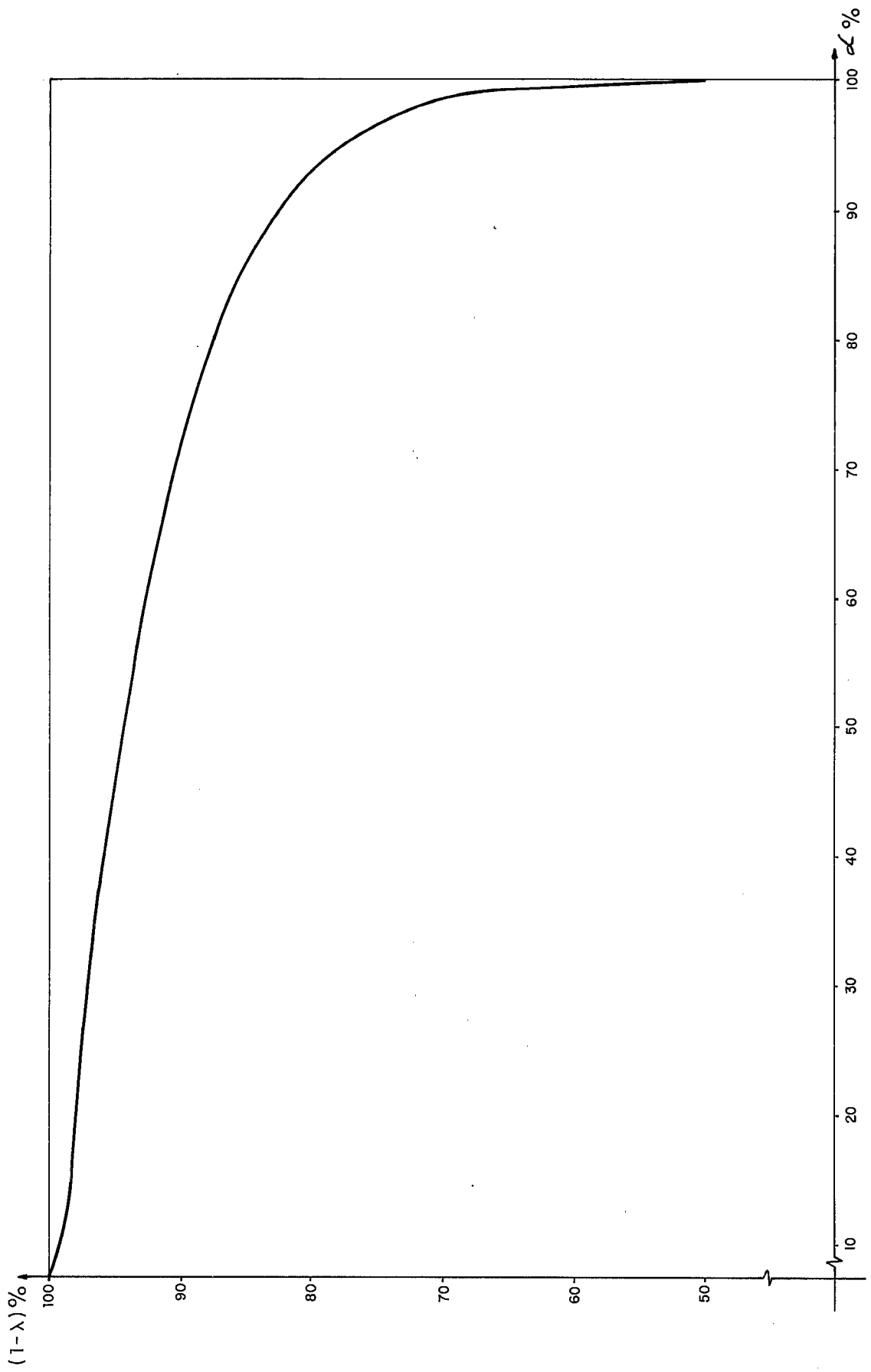


FIG. 3.2.- CURVA CARACTERÍSTICA DE OPERAÇÃO PARA O TESTE DE SIGNIFICÂNCIA DE HARMÔNICOS

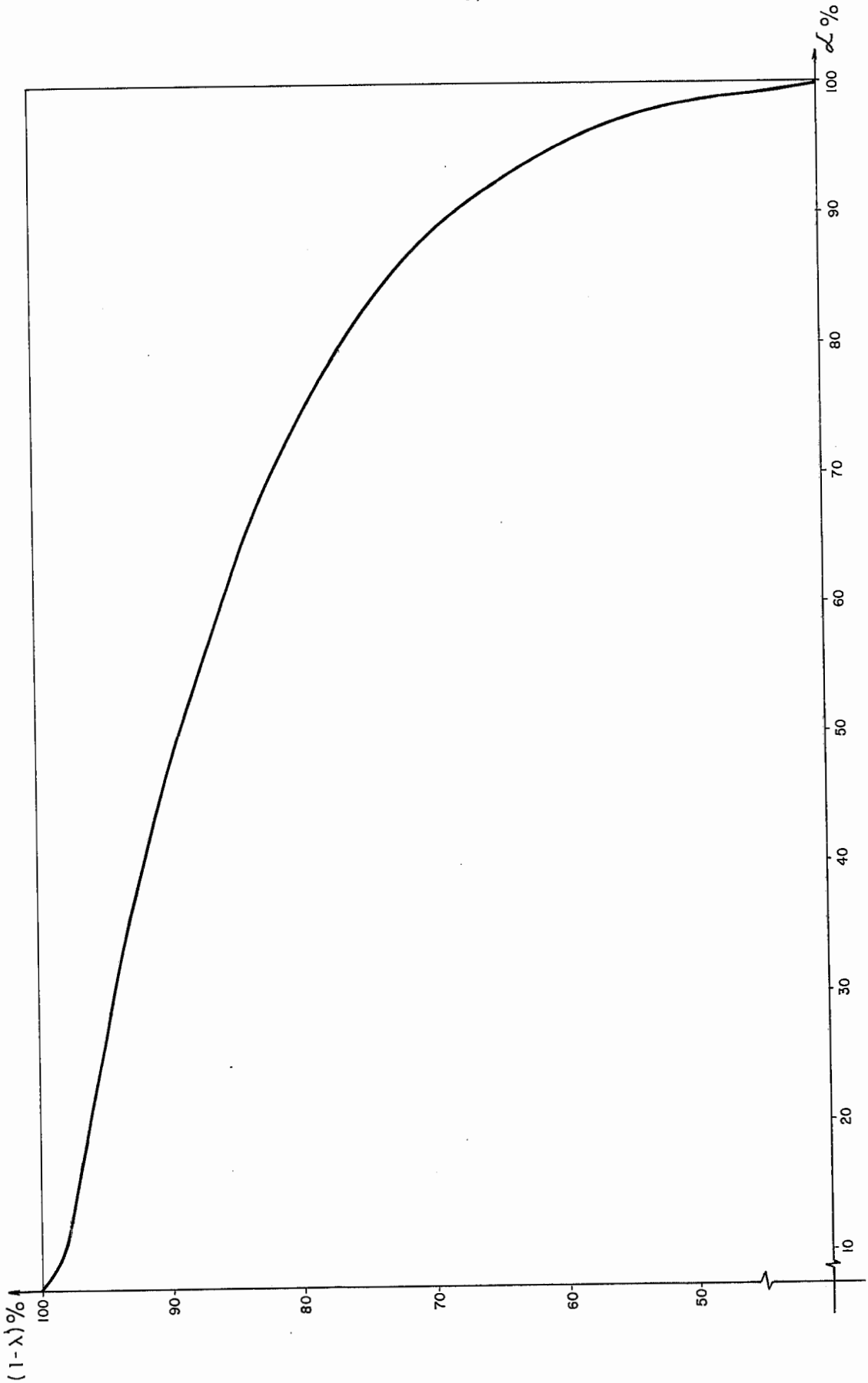


FIG. 3.3.- CURVA CARACTERÍSTICA DE OPERAÇÃO PARA O TESTE DE SIGNIFICÂNCIA DE HARMÔNICOS NO COEFICIENTE DE CORRELAÇÃO MENSAL

drões), coeficientes de correlação e coeficientes de variação mensais.

Assim,

a) INTERVALOS DE CONFIANÇA PARA AS MÉDIAS MENSAIS

Seja X_τ ($\tau=1,12$) uma variável aleatória $N(\mu_\tau, \sigma_\tau^2)$. A distribuição de amostragem de \bar{X}_τ (média da amostra na posição τ) é normal com média μ_τ e variância σ_τ^2/n , onde n é o tamanho da amostra. Se X_τ não é normal, a estatística \bar{X}_τ é assintoticamente normal $(\mu_\tau; \frac{\sigma_\tau^2}{n})$. Por outro lado, pode-se mostrar também que,

$$\bar{X}_\tau \xrightarrow{p} \mu_\tau$$

No caso mais geral, se σ_τ é desconhecido, tem-se que o intervalo de confiança para μ_τ é dado por,

$$(3.13) \quad P \left\{ \bar{X}_\tau - t_{\lambda/2} \frac{S_\tau}{\sqrt{n}} \leq \mu_\tau \leq \bar{X}_\tau + t_{\lambda/2} \frac{S_\tau}{\sqrt{n}} \right\} = 1-\lambda$$

onde, $t_{\lambda/2}$ é um valor da distribuição t de Student com $v = n-1$ graus de liberdade, S_τ uma estimativa de σ_τ e λ o nível de significância.

Logo, o intervalo de confiança de $(1-\lambda)\%$ pode ser representado por,

$$(3.14) \quad \text{CONF} \left\{ C_{1_\tau} \leq \mu_\tau \leq C_{2_\tau} \right\}, \quad \tau = 1, 2, \dots, 12$$

b) INTERVALOS DE CONFIANÇA PARA AS VARIÂNCIAS (D.PADRÕES) MENSAIS

O estimador não tendencioso para a variância da população, σ_{τ}^2 ($\tau=1,12$) é dado por,

$$(3.15) \quad \hat{S}_{\tau}^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x}_{\tau})^2$$

Pode-se mostrar (KMENTA, 1971) que a distribuição de amostragem da estatística \hat{S}_{τ}^2 é um caso especial da distribuição gama de parâmetros $\alpha=(n-1)/2$ e $\beta=2\sigma_{\tau}^2/(n-1)$. A forma exata da distribuição de \hat{S}_{τ}^2 depende, então, de dois parâmetros : da variância da população, σ_{τ}^2 e do tamanho n da amostra. A distribuição é sempre assimétrica positiva para pequenas amostras e vai se tornando cada vez mais simétrica a medida que o tamanho da amostra cresce. Consequentemente a média da distribuição, de \hat{S}_{τ}^2 vale σ_{τ}^2 e a variância é dada por,

$$(3.16) \quad \text{VAR} (\hat{S}_{\tau}^2) = \frac{2 \sigma_{\tau}^4}{n-1}$$

Para se determinar a probabilidade de que \hat{S}_{τ}^2 caia num específico intervalo, tem-se que obter áreas sob correspondentes curvas gama. Ter-se-ã, então, diferentes curvas para diferentes combinações de σ_{τ}^2 e n e a determinação da área apropriada passa a ser muito trabalhosa.

Com relação à determinação dos intervalos de confiança, o problema pode ser contornado de forma ra-

zoável bastando para tanto que se assuma a normalidade da distribuição.

Lembrando que a variável aleatória $\frac{n S_{\tau}^2}{\sigma_{\tau}^2}$

tem distribuição χ^2 assintótica com $\nu=n$ graus de liberdade, onde,

$$(3.17) \quad S_{\tau}^2 = \frac{1}{n} \sum_i (x_i - \mu_{\tau})^2$$

vem que,

$$(3.18) \quad P \left\{ C_{1\tau} \leq \frac{n S_{\tau}^2}{\sigma_{\tau}^2} \leq C_{2\tau} \right\} = 1 - \lambda$$

com $\lambda =$ nível de significância.

Através das tabelas da distribuição χ^2 pode-se obter dois números $C_{1\tau}$ e $C_{2\tau}$, tais que:

$$P \left\{ C_{1\tau} \leq \chi_{OBS}^2 \right\} = 1 - \lambda/2 \quad \text{e} \quad P \left\{ \chi_{OBS}^2 \leq C_{2\tau} \right\} = 1 - \lambda/2$$

ou seja, os limites $C_{1\tau}$ e $C_{2\tau}$ são tais que a probabilidade de que um valor observado exceda $C_{1\tau}$ seja igual a $1 - \lambda/2$ e de que não exceda $C_{2\tau}$ seja também igual a $1 - \lambda/2$.

Então, o intervalo de confiança de $(1 - \lambda) \%$ para o parâmetro σ_{τ}^2 vale,

$$(3.19) \quad \text{CONF} \left\{ \frac{\sum_i (x_i - \mu_{\tau})^2}{C_{2\tau}} \leq \sigma_{\tau}^2 \leq \frac{\sum_i (x_i - \mu_{\tau})^2}{C_{1\tau}} \right\}$$

Se a média μ_τ é desconhecida, então a va-

riável aleatória $\frac{\sum_i (x_i - \bar{x}_\tau)^2}{\sigma_\tau^2}$ é assintoticamente $\chi^2_{(n-1)}$.

Valores $C'_{1\tau}$ e $C'_{2\tau}$ são obtidos de forma

análoga, considerando apenas que $\nu = n-1$ graus de liberdade, sendo o intervalo de confiança de $(1-\lambda)\%$ igual a:

$$(3.20) \quad \text{CONF} \left\{ \frac{\sum_i (x_i - \bar{x}_\tau)^2}{C'_{2\tau}} \leq \sigma_\tau^2 \leq \frac{\sum_i (x_i - \bar{x}_\tau)^2}{C'_{1\tau}} \right\}$$

Com relação aos desvios padrões mensais, o intervalo de confiança de $(1-\lambda)\%$ vale,

$$(3.21) \quad \text{CONF} \left\{ \sqrt{\frac{\sum_i (x_i - \bar{x}_\tau)^2}{C'_{2\tau}}} \leq \sigma_\tau \leq \sqrt{\frac{\sum_i (x_i - \bar{x}_\tau)^2}{C'_{1\tau}}} \right\}$$

c) INTERVALOS DE CONFIANÇA PARA OS COEFICIENTES DE CORRELAÇÃO MENSAIS

Suponha-se que $(X_\tau, X_{\tau+1})$ seja uma variável aleatória normal bi-dimensional e ρ_τ ($\tau = 1, 12$) o coeficiente de correlação entre as variáveis X_τ e $X_{\tau+1}$.

Através do coeficiente de correlação da amostra bi-dimensional, R_τ , seja a estatística

$$(3.22) \quad Z_{0\tau} = \frac{1}{2} \ln \frac{1 + R_\tau}{1 - R_\tau}$$

FISHER (1921) mostrou que esta estatística é uma estimativa da média,

$$(3.23) \quad \mu_{\tau}^* = \frac{1}{2} \operatorname{Ln} \frac{1+\rho_{\tau}}{1-\rho_{\tau}}$$

de uma variável aleatória Z_{τ}^* , assintoticamente normal com variância dada por

$$(3.24) \quad \sigma_{\tau}^{*2} = \frac{1}{n-3}$$

onde, n é o tamanho da amostra.

Por analogia com intervalo de confiança para as médias, pode-se mostrar que,

$$(3.25) \quad P \left\{ Z_{0_{\tau}} - K \leq \mu_{\tau}^* \leq Z_{0_{\tau}} + K \right\} = 1-\lambda$$

onde, $K = Z_{\lambda/2} \frac{1}{\sqrt{n-3}}$ e λ é o nível de significância.

Consequentemente os intervalos de confiança para ρ_{τ} é dado por:

$$(3.26) \quad \text{CONF} \left\{ C_{1_{\tau}} \leq \rho_{\tau} \leq C_{2_{\tau}} \right\}$$

com, $C_{1_{\tau}} = \operatorname{TANH} (Z_0 - K)$ e $C_{2_{\tau}} = \operatorname{TANH} (Z_0 + K)$

d) INTERVALOS DE CONFIANÇA PARA OS COEFICIENTES DE VARIAÇÃO MENSAIS

O coeficiente de variação mensal é definido como sendo a relação entre o desvio padrão mensal, S_{τ} e a média mensal, \bar{X}_{τ} , ou seja,

$$(3.27) \quad V_{\tau} = \frac{S_{\tau}}{\bar{X}_{\tau}} \quad (\tau = \overline{1,12})$$

Pode-se mostrar (CRAMÉR, 1970) que a distribuição de amostragem dos V_{τ} , cada um computado para os n anos de observação, é assintoticamente normal com média

$$(3.28) \quad \eta_{\tau} = \frac{\sigma_{\tau}}{\mu_{\tau}} \quad (\tau = \overline{1,12})$$

e variância dada por,

$$(3.29) \quad \sigma_{V_{\tau}}^2 = \frac{\mu_{\tau}^2 (\mu_{4\tau} - \mu_{2\tau}^2) - 4 \mu_{\tau} \mu_{2\tau} \mu_{3\tau} + 4 \mu_{2\tau}^3}{4 \mu_{\tau}^4 \mu_{2\tau} n}$$

onde, μ_{τ} é média mensal e $\mu_{2\tau}$, $\mu_{3\tau}$ e $\mu_{4\tau}$ são os momentos centrais de ordem 2, 3 e 4, respectivamente.

Então, a variável aleatória,

$$(3.30) \quad Z = \frac{V_{\tau} - \eta_{\tau}}{\sigma_{V_{\tau}}} \sim N(0,1)$$

Se, $\sigma_{V_{\tau}}$ é desconhecido pode-se recorrer ao fato de que a variável aleatória,

$$(3.31) \quad t = \frac{V_{\tau} - \eta_{\tau}}{S_{V_{\tau}}}$$

possui distribuição t de Student assintótica com $\nu = n-1$ graus de liberdade, onde $S_{V_{\tau}}$ é obtido através a utilização da média e dos respectivos momentos centrais de ordem 2, 3 e 4 da amostra.

Então, os intervalos de confiança de $(1-\lambda)\%$ para os coeficientes de variação η_τ ($\tau=1,12$) são obtidos por,

$$(3.32) \quad P \left\{ V_\tau - t_{\lambda/2} S_{V_\tau} \leq \eta_\tau \leq V_\tau + t_{\lambda/2} S_{V_\tau} \right\} = 1-\lambda$$

ou seja,

$$(3.33) \quad \text{CONF} \left\{ C_{1\tau} \leq \eta_\tau \leq C_{2\tau} \right\}, \quad \tau = 1, 2, \dots, 12$$

III.7 - TESTE PARA CONSTÂNCIA DE PARÂMETROS

A partir do instante de que um dos principais objetivos deste trabalho é a obtenção de um modelo econômico, quanto ao número de parâmetros e coeficientes, torna-se imprescindível que se verifique a hipótese de que alguns parâmetros não são significativamente diferentes de uma constante. As equações (2.11), (2.12) e (2.17) ressaltam esta afirmação.

Em relação ao novo teste aqui proposto, pode-se dizer que como o anterior, descrito em III.6, é de uma filosofia bastante simples e se constitui numa alternativa para procedimentos tradicionais como dividir a amostra em duas sub-amostras e testar a igualdade de suas médias e variâncias. Para tanto, basta que se obtenha, para o parâmetro a ser testado, os respectivos intervalos de confiança.

Ponha-se, então, como hipótese nula do teste,

$$H_0 : \pi = c$$

e, como alternativa,

$$H_1 : \pi \neq c$$

onde, π é o parâmetro que se deseja testar e c é uma constante.

O teste consiste em se comparar o mínimo dos limites superiores com o máximo dos inferiores e rejeitar H_0 quando,

$$\text{MAX} \{ \text{LIM INF} \} \geq \text{MIN} \{ \text{LIM SUP} \}$$

ou seja, rejeitar a hipótese de que o parâmetro π é significativamente igual a uma constante se a linha poligonal formada pelos limites inferiores (superiores) interceptar a formada pelos superiores (inferiores). Em outras palavras, isto significa que a componente (π_x) da análise harmônica não está totalmente contida na faixa de confiança do parâmetro. Os gráficos a seguir ilustram a idéia do teste, onde a hipótese nula é rejeitada em (a) e aceita em (b).

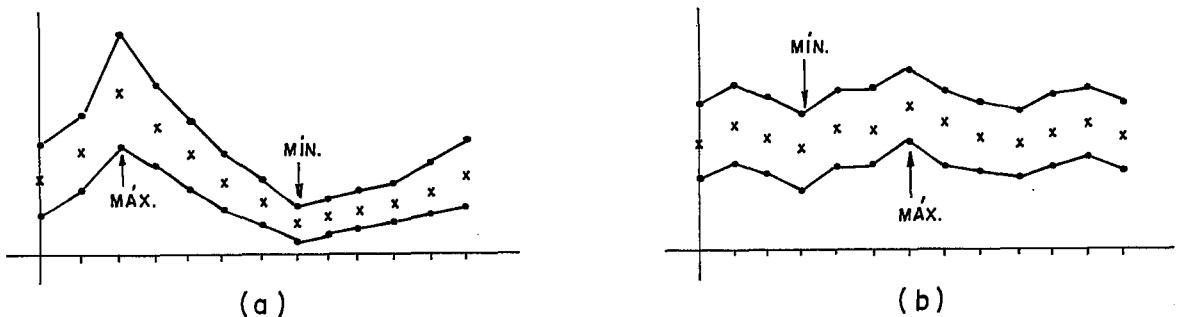


FIG. 3.4 - INTERVALOS DE CONFIANÇA DE $(1 - \lambda) \%$ PARA π

Para reforçar a idéia da importância deste teste, note-se, apenas como exemplo, que caso se aceite a hipótese de que o coeficiente de variação mensal, η_{τ} , é um parâmetro não periódico, uma economia imediata é obtida na estimação dos parâmetros, visto isto implicar em que os harmônicos das funções periódicas ajustadas às médias e aos desvios padrões mensais, respectivamente, \bar{X}_{τ} e S_{τ} , possuem as mesmas frequências e fases e amplitudes proporcionais.

Finalmente, se o parâmetro a ser testado é o coeficiente de correlação mensal, uma dicotomia semelhante àquela descrita no capítulo anterior, reunindo os meses pares e ímpares, deve ser realizada, sendo o teste aplicado, então, nas duas sub-amostras assim constituídas.

III.8 - MODELOS DE DEPENDÊNCIA PARA A COMPONENTE ESTOCÁSTICA

Uma maior generalização das duas metodologias, a de Yevjevich e a de Box e Jenkins, pode ser obtida neste estágio. As equações (2.11) a (2.18) fornecem um meio de se obter a componente estocástica do modelo, com ausência de periodicidade na média e variância. Esta variável deve ser padronizada em vista de se ter adotado o método paramétrico para separar as componentes periódica e estocástica. Seja, então, $\zeta_{\rho, \tau}$ a componente estocástica dependente padronizada do modelo.

Certamente, dependendo da particular aplicação, esta componente pode ser representada por um modelo ARIMA sazonal

Basta, por exemplo, que a análise harmônica não tenha retirado toda a periodicidade da média e/ou variância. O que se quer dizer com isso é que $\zeta_{\rho, \tau}$ não deve ser considerada automaticamente um modelo Markoviano ou outro qualquer tipo de modelo mas sim, um apropriado ARIMA, sazonal ou não, a partir dos estágios de identificação, estimação e testes de verificação da modelagem Box e Jenkins. Por outro lado, é possível que em alguns casos, o correto modelo de dependência para a componente $\zeta_{\rho, \tau}$ seja o de Markov.

Seja, então, $\{\zeta_t, t=1, N, N= n\omega\}$ a série discreta, medida a intervalos equiespaçados no tempo. Admita-se que ζ_t possa ser modelada por um modelo multiplicativo ARIMA sazonal da forma,

$$(3.34) \quad \phi_p(B) \phi_p(B^S) \nabla^d \nabla_S^D \zeta_t = \theta_q(B) \theta_Q(B^S) a_t$$

onde,

t - discretização no tempo;

s - período de sazonalidade;

B, ∇ - operadores, atraso e diferença, definidos em II.3;

$\phi_p(B), \phi_p(B^S), \theta_q(B), \theta_Q(B^S)$ - respectivamente, operador autoregressivo, autoregressivo sazonal, médias móveis, e médias móveis sazonal definidos, também, em II.3;

$\{a_t, t=1, N-d-sD\}$ - sequência de variáveis aleatórias independentes e normalmente distribuídas.

A equação (3.34) pode ser representada, de forma compacta, por $(p,d,q) \times (P,D,Q)_S$ com (p,d,q) representando a ordem dos operadores não sazonais e $(P,D,Q)_S$, a ordem dos operadores sazonais. Evidentemente, se o modelo é não sazonal, somente a notação (p,d,q) é necessária.

III.9 - IDENTIFICAÇÃO DO MODELO

O propósito do estágio de IDENTIFICAÇÃO é determinar os graus das diferenças simples (d) e sazonal (D), requeridas para produzir estacionaridade, bem como a ordem não sazonal e sazonal dos operadores autoregressivos (AR) e médias móveis (MA) da série adequadamente diferenciada. Isto pode ser conseguido após o exame cuidadoso do gráfico da série e das funções de autocorrelação e autocorrelação parcial estimadas.

Para modelos não estacionários sem componente sazonal, a função de autocorrelação decresce linearmente; a ocorrência de picos em intervalos que sejam múltiplos inteiros do período de sazonalidade (s) indica, em geral, que uma diferença sazonal é requerida. Para a determinação do grau adequado de diferenciação, um cuidado especial deve ser tomado, visto que a variância da série decresce com o grau da diferença até que se obtenha estacionaridade e cresce novamente com a sobrediferenciação.

III.9.1 - FUNÇÃO DE AUTOCORRELAÇÃO DA AMOSTRA

O coeficiente de correlação linear simples entre duas variáveis X e Y (R_{xy}) pode ser encarado como uma medida de ajustamento de uma equação da forma,

$$(3.35) \quad Y = \alpha + \beta X$$

R_{xy} pode ser estimado por,

$$(3.36) \quad R_{xy} = \frac{N \sum x_i y_i - (\sum x_i) (\sum y_i)}{\sqrt{N \sum x_i^2 - (\sum x_i)^2} \sqrt{N \sum y_i^2 - (\sum y_i)^2}}$$

onde os somat6rios se estendem de $i=1$ a $i=N$. R_{xy} pode variar no intervalo $(-1, 1)$. A um valor de $R_{xy} = 0$ corresponde a uma aus6ncia de correla63o; para $R_{xy} = 1$ tem-se uma correla63o linear positiva perfeita, enquanto que $R_{xy} = -1$ mostra uma associa63o linear perfeitamente negativa. O sinal indica se y cresce ou decresce com o aumento de x . Em geral se sup6e que os valores da vari6vel dependente, y , s3o normalmente distribu6dos em torno da reta de regress3o ajustada a6s dados e, os desvios das observa63es em rela63o a linha ajustada, independentes.

Ao quadrado de R_{xy} denomina-se coeficiente de determina63o. Fisicamente representa a percentagem da varia63o de y que 6 associada a uma dada varia63o em x .

Se, por outro lado, o conjunto de dados 6 relativo a s6ries de tempo estacion6rias, pode-se aplicar (3.36) para se estimar os coeficientes de autocorrela63o entre os valores da vari6vel x_i e a mesma vari6vel defasada de k unidades. Os coeficientes de autocorrela63o para um particular intervalo k , s3o estimados por,

$$(3.37) \quad R_k = \frac{(N-k) (\sum x_i x_{i+k}) - (\sum x_i) (\sum x_{i+k})}{\sqrt{(N-k) \sum x_i^2 - (\sum x_i)^2} \sqrt{(N-k) \sum x_{i+k}^2 - (\sum x_{i+k})^2}}$$

neste caso, os somat6rios se estendem de $i=1$ a $i=N-k$.

À R_k , expresso como função de k , dá-se o nome de função de autocorrelação da amostra. O exame desta função indica os intervalos ou períodos, em relação aos quais, os dados das séries de tempo podem ser correlacionados.

Se previamente os dados são padronizados, a expressão anterior, para N suficientemente grande, se reduz a,

$$(3.38) \quad R_k = \frac{\sum x_i x_{i+k}}{\sum x_i^2}, \quad i = \overline{1, N-k}$$

A função de autocovariância da amostra para séries temporais com média zero é, então, estimada por,

$$(3.39) \quad C_k = \frac{1}{N-k} \sum_{i=1}^{N-k} x_i x_{i+k}$$

Segue-se que,

$$(3.40) \quad R_k = \frac{C_k}{C_0} = \frac{C_k}{S_x^2}$$

Um ponto importante a ser realçado é o erro padrão das estimativas das autocorrelações. A variância de R_k , para um processo normal estacionário, fornecida por BARTLETT (1946) é dada por,

$$(3.41) \quad \text{VAR}(R_k) \cong \frac{1}{N} \sum_{v=-\infty}^{\infty} \left[\rho_v^2 + \rho_{v+k} \rho_{v-k} - 4 \rho_k \rho_v \rho_{v-k} + 2\rho_v^2 \rho_k^2 \right]$$

onde, ρ_v é a autocorrelação teórica defasada de v unidades, definida por,

$$(3.42) \quad \rho_v = \frac{\gamma_v}{\gamma_0}$$

com, γ_v sendo a autocovariância teórica de ordem v .

Para um processo, no qual todas as autocorrelações ρ_v são nulas para $v > q$, a aproximação de Bartlett fornece,

$$(3.43) \quad \text{VAR} (R_k) \cong \frac{1}{N} \left[1 + 2 \sum_{v=1}^q \rho_v^2 \right], \quad k > q$$

Expressões semelhantes foram obtidas por Bartlett para a covariância entre autocorrelações estimadas, R_k e R_{k+s} . Em particular, para grandes valores de k , tem-se,

$$(3.44) \quad \text{COV} \left[R_k, R_{k+s} \right] \cong \frac{1}{N} \sum_{v=-\infty}^{\infty} \rho_v \rho_{v+s}$$

O resultado de Bartlett mostra que um cuidado especial deve ser tomado ao se analisar as autocorrelações estimadas, visto que para valores próximos pode existir uma alta correlação.

BOX e JENKINS (1970) sugerem que o máximo valor de k , no cálculo da função de autocorrelação da amostra, não ultrapasse a $N/4$, onde N é o comprimento da série.

III.9.2 - FUNÇÃO DE AUTOCORRELAÇÃO TEÓRICA PARA MODELOS AUTOREGRESSIVOS (AR), MÉDIAS MÓVEIS (MA) E MISTOS (ARMA)

III.9.2.1 - MODELOS AUTOREGRESSIVOS

Seja um processo autoregressivo estacionário de ordem p , definido pela equação (2.27).

Multiplicando (2.27) por Z_{t-k} e tomando os valores esperados obtêm-se,

$$(3.45) \quad \gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}, \quad k > 0$$

Dividindo (3.45) por γ_0 vem,

$$(3.46) \quad \rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k > 0$$

Substituindo $k = 1, 2, \dots, p$ em (3.46), obtêm-se um conjunto de equações lineares para $\phi_1, \phi_2, \dots, \phi_p$ em termos de $\rho_1, \rho_2, \dots, \rho_p$. A esse conjunto dá-se o nome de equações de Yule-Walker que, em termos matriciais são representadas por,

$$(3.47) \quad \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \vdots \\ \rho_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \vdots \\ \phi_p \end{bmatrix}$$

A equação (3.47) pode, também, ser representada por,

$$(3.48) \quad \phi_p(B) \rho_k = 0, \quad k = 1, 2, \dots, p$$

onde, $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ e B opera agora em k e não em t .

Escrevendo $\phi_p(B)$ como um produto de monômios, tem-se:

$$(3.49) \quad \phi_p(B) = \prod_{i=1}^p (1 - G_i B)$$

Então, a solução geral de (3.46) é dada por,

$$(3.50) \quad \rho_k = A_1 G_1^k + A_2 G_2^k + \dots + A_p G_p^k$$

onde, $G_1^{-1}, G_2^{-1}, \dots, G_p^{-1}$ são as raízes da equação característica $\phi_p(B) = 0$.

A condição de estacionaridade requer que $|G_i| < 1, \forall_i = \overline{1, p}$. Então se as raízes G_i são distintas, pode-se ter dois casos:

(a) a raiz G_i é real $\Rightarrow A_i G_i^k \rightarrow 0$ quando $k \rightarrow \infty$. A função de autocorrelação assume um aspecto exponencial decrescente;

(b) o par de raízes complexas conjugadas G_i, G_j contribui com o termo $A d^k \text{sen}(2\pi f k + \phi)$ onde d é o fator de amortecimento, f , a frequência e ϕ , a fase. A função de autocorrelação tem o aspecto de uma senóide amortecida.

Para o caso de um processo autoregressivo de primeira ordem, tem-se:

$$(3.51) \quad \rho_k = \phi_1 \rho_{k-1}, \quad k > 0$$

$$(3.51.a) \quad \rho_0 = 1 \Rightarrow \rho_k = \phi_1^k, \quad k \geq 0$$

Logo, para um processo AR (1) a função de autocorrelação decresce exponencialmente a zero quando ϕ_1 é positivo; para ϕ_1 negativo, o decréscimo é acompanhado de uma oscilação.

Para um processo AR (2):

$$(3.52) \quad \rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k > 0$$

Expressando ρ_1 e ρ_2 em termos de ϕ_1 e ϕ_2 , obtêm-se para valores iniciais:

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}$$

e,

$$\rho_2 = \phi_2 + \frac{\phi_1^2}{1 - \phi_2}$$

A solução geral para a equação em diferenças (3.52) é dada por,

$$(3.53) \quad \rho_k = A_1 G_1^k + A_2 G_2^k = \frac{G_1(1-G_2^2)G_1^k - G_2(1-G_1^2)G_2^k}{(G_1 - G_2)(1 + G_1 G_2)}$$

onde, G_1^{-1} e G_2^{-1} são as raízes da equação característica $\phi_2(B) = 0$.

Se as raízes são reais ($\phi_1^2 + 4\phi_2 \geq 0$), a função de autocorrelação consiste de uma mistura de exponenciais amortecidas. Se as raízes forem complexas conjugadas ($\phi_1^2 + 4\phi_2 < 0$) o processo autoregressivo de segunda ordem apresenta um comportamento pseudo-periódico. A função de autocorrelação tem o aspecto de uma senóide amortecida.

III.9.2.2 - MODELOS MÉDIAS MÔVEIS

Seja um processo médias móveis inversível-inversibilidade é uma condição necessária para garantir a unicidade - de ordem q , definido pela equação (2.30).

Usando (2.30) a função de autocovariância para um processo médias móveis de ordem q é dada por,

$$(3.54) \quad \gamma_k = E \{ (a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}) (a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q}) \}$$

Por conseguinte, a variância do processo vale:

$$(3.55) \quad \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2$$

e,

$$(3.56) \quad \gamma_k = \begin{cases} (-\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q) \sigma_a^2, & k = \overline{1, q} \\ 0 & , k > q \end{cases}$$

Então, a função de autocorrelação $\bar{\rho}$ é dada por,

$$(3.57) \quad \rho_k = \frac{\gamma_k}{\gamma_0} = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1 + \dots + \theta_q} & , k = \overline{1, q} \\ 0 & , k > q \end{cases}$$

Logo, a função de autocorrelação para um processo médias móveis de ordem q , assume o valor zero para todo $k > q$.

III.9.2.3 - MODELOS MISTOS

Para processos mistos, autoregressivo de ordem p e médias móveis de ordem q satisfazendo, respectivamente, as condições de estacionaridade - a equação característica $\phi_p(B) = 0$ possuindo todas as raízes fora do círculo unitário - e inversibilidade - as raízes de $\theta_q(B) = 0$, caindo fora do círculo unitário - a função de autocorrelação tem o aspecto de exponenciais e/ou senóides amortecidas, para $k > q - p$.

III.9.3 - FUNÇÃO DE AUTOCORRELAÇÃO PARCIAL DA AMOSTRA

Da mesma forma que a função de autocorrelação, a finalidade precípua da função de autocorrelação parcial é a ajuda na identificação da ordem e classe do modelo que possa melhor se ajustar a uma série de tempo observada.

Representando por ϕ_{kj} , o j -ésimo coeficiente da equação (3.46), tem-se:

$$(3.58) \quad \rho_j = \phi_{k1} \rho_{j-1} + \phi_{k2} \rho_{j-2} + \dots + \phi_{kk} \rho_{j-k}, \quad j=\overline{1, k}$$

Sob a forma das equações de Yule-Walker, vem:

$$(3.59) \quad \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix}$$

Resolvendo-se estas equações para $k=1,2,3,\dots$, obtêm-se:

$$\phi_{11} = \rho_1 \quad ;$$

$$\phi_{22} = \frac{\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \quad ;$$

$$\phi_{33} = \frac{\begin{bmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{bmatrix}} \quad \text{e, assim por diante.}$$

À ϕ_{kk} expresso como função de k dá-se o nome de função de autocorrelação parcial.

A estimação da função de autocorrelação parcial pode ser obtida através da substituição de ρ_j por seu estimador, R_j , isto é,

$$(3.60) \quad R_j = \hat{\phi}_{k1} R_{j-1} + \hat{\phi}_{k2} R_{j-2} + \dots + \hat{\phi}_{kk} R_{j-k}, \quad j = \overline{1, k}$$

obtendo-se, desta forma, um conjunto de $j = 1, 2, \dots, k$ equações que devem ser resolvidas para $k = 1, 2, \dots$. Como observação, cabe ressaltar que este procedimento não deve ser empregado quando os parâmetros assumem valores próximos a faixas de não estacionaridade. A função de autocorrelação parcial pode também ser estimada pelo coeficiente de mais alto grau de modelos autoregressivos de ordem $1, 2, 3, \dots$, respectivamente, utilizando-se estimadores de mínimos quadrados.

Por outro lado, QUENOUILLE (1949) mostrou que para um processo autoregressivo de ordem p as autocorrelações parciais estimadas para ordens maiores ou iguais a $(p+1)$ são aproximadamente independentes e normalmente distribuídas com média zero e variância dada por,

$$(3.61) \quad \text{VAR} (\hat{\phi}_{kk}) \cong \frac{1}{N}, \quad k \geq p+1$$

III.9.4 - FUNÇÃO DE AUTOCORRELAÇÃO PARCIAL TEÓRICA PARA MODELOS AUTOREGRESSIVOS (AR), MÉDIAS MÔVEIS (MA) E MISTOS (ARMA)

III.9.4.1 - MODELOS AUTOREGRESSIVOS

Para processos autoregressivos estacionários de ordem p , a função de autocorrelação parcial assume valores não nulos para $k > p$. Este resultado é uma consequência natural da definição de função de autocorrelação parcial.

III.9.4.2 - MODELOS MÉDIAS MÔVEIS

A função de autocorrelação parcial para processos inversíveis médias móveis de ordem q tem uma extensão infinita sendo seu aspecto dominado por exponenciais e/ou senóides amortecidas (BOX e JENKINS, 1970).

A partir da equação (3.57), para um processo médias móveis de primeira ordem, tem-se que:

$$(3.62) \quad \rho_1 = \frac{-\theta_1}{1 + \theta_1^2} \quad \text{e} \quad \rho_k = 0, \quad k > 1$$

Levando estes resultados em (3.59), vem:

$$(3.63) \quad \phi_{kk} = \frac{-\theta_1^k (1 - \theta_1^2)}{1 - \theta_1^{2(k+1)}}$$

Então, se $|\phi_{kk}| < \theta_1^k$, a função de autocorrelação parcial é dominada por uma exponencial amortecida. Se ρ_1 é positivo, então θ_1 é negativo e vice-versa; no primeiro caso a função alterna o sinal; no segundo é negativa (BOX e JENKINS, 1970).

A função de autocorrelação parcial para processos médias móveis de segunda ordem é denominada por uma soma de exponenciais - raízes de $\theta_2(B) = 0$, reais - ou de uma senóide amortecida - raízes de $\theta_2(B) = 0$, complexas conjugadas. Seu comportamento pode ser comparado ao da função de autocorrelação para um processo AR (2).

III.9.4.3 - MODELOS MISTOS

Para um processo misto ARMA (p,q), satisfazendo as condições de estacionaridade e inversibilidade, tem-se que:

$$a_t = \theta_q^{-1}(B) \phi_p(B) Z_t$$

onde, $\theta_q^{-1}(B)$ é uma série infinita em B. Então, a função

de autocorrelação parcial para um processo misto tem extensão infinita. Seu aspecto é dominado por uma mistura de exponenciais e/ou senóides amortecidas para $k > p - q$ (BOX e JENKINS, 1970).

III.9.5 - MODELOS SAZONAIS

O processo sugerido para identificação do modelo referente a uma série de tempo com componente sazonal é semelhante ao de uma série não sazonal.

Então, para um processo médias móveis sazonal, a função de autocorrelação será composta de valores nulos para $k > q + SQ$, sendo válida, também, a aproximação de BARTLETT (1946),

$$(3.64) \quad \text{VAR}(R_k) \approx \frac{1}{N} \left(1 + 2 \sum_{v=1}^{q+sQ} \rho_v^2 \right), \quad k > q + sQ$$

De forma análoga, para processos autoregressivos sazonais, a função de autocorrelação parcial assume valores nulos para $k > p + SP$. Neste caso, é válida, também, a expressão fornecida por QUENOUILLE (1949),

$$(3.65) \quad \text{VAR}(\hat{\phi}_{kk}) \approx \frac{1}{N}, \quad k > p + SP$$

Em adição, BOX e JENKINS (1970) apresentam uma tabela de funções de autocovariâncias (autocorrelações) teóricas para modelos com diferentes valores de p, q, P e Q .

III.10 - ESTIMAÇÃO DOS PARÂMETROS DO MODELO

Suponha-se que o processo gerador de uma série $\{ Z_t, t = \overline{1, N} \}$ seja um modelo ARIMA (p, d, q) . Ponha-se como objetivo a estimação dos parâmetros ϕ_1, \dots, ϕ_p e $\theta_1, \dots, \theta_q$ deste modelo. O problema é equivalente a estimar os parâmetros ϕ 's e θ 's do modelo estacionário ARMA (p, q) , associado, o qual pode ser escrito por,

$$(3.66) \quad a_t = \omega_t - \phi_1 \omega_{t-1} - \dots - \phi_p \omega_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

onde,

$$\omega_t = \nabla^d Z_t$$

Ao se minimizar a soma dos quadrados dos resíduos, a_t , obtêm-se estimativas aproximadas de máxima verossimilhança para os parâmetros (BOX e JENKINS, 1970).

De acordo com BOX e JENKINS (1970) estas estimativas podem ser obtidas pela utilização de dois procedimentos gerais:

- (i) cálculo condicional da soma dos quadrados dos resíduos a_t ou,
- (ii) cálculo não condicional da soma dos quadrados dos resíduos a_t .

Observando a equação (3.66) nota-se que o problema maior reside em não se dispor dos valores iniciais.

O contorno desta situação, no caso de se adotar o procedimento (i) pode ser feito de duas maneiras:

(ia) colocando todos os valores iniciais iguais a zero. Neste caso, se $n = N-d$ é suficientemente grande, obtêm-se bons resultados, a menos que alguns zeros do polinômio autoregressivo estejam próximos do círculo unitário. Se isto acontecer, longos e demorados transientes podem ser introduzidos;

(ib) fazendo os p primeiros valores de a_t iguais a zero, isto é, calcula-se os a_t a partir de $t = p+1$. Se o tamanho N da série for grande (ou, $n = N-p-d$, grande), a perda de informação será desprezível.

Caso o modelo seja um IMA (d,q) , (ia) e (ib) são idênticos.

Para o procedimento (ii) um cálculo preliminar em sentido inverso é feito, tomando-se as esperanças condicionais da equação (3.66), isto é $[\omega_j]$, $j=0,1,2,\dots,p$. Estes valores condicionais são as previsões inversas de ω_j , necessárias ao início da recursão. Esta técnica é chamada de "back forecasting" e é descrita sucintamente a seguir.

A equação (3.66) pode, alternativamente, ser escrita por,

$$(3.67) \quad \phi_p(B) \omega_t = \theta_q(B) a_t$$

ou por,

$$(3.68) \quad \phi_p (F) \omega_t = \theta_q (F) e_t$$

com e_t sendo uma seqüência de variáveis aleatórias independentes com média zero e variância $\sigma_e^2 = \sigma_a^2$.

As equações (3.67) e (3.68) dão-se, respectivamente, o nome modelo adiantamento e modelo atraso.

Tomando-se o valor esperado condicional de (3.68), obtêm-se:

$$(3.69) \quad \phi_p (F) [\omega_t] = \theta_q (F) [e_t]$$

onde,

$$[\omega_t] = \omega_t, \text{ para } t = \overline{1, n} \text{ e igual a previsão inversa para } t \leq 0,$$

$$[e_t] = E [e_t | \underset{\sim}{\phi}, \underset{\sim}{\theta}, \underset{\sim}{\omega}] \text{ , para } t = \overline{1, n} \text{ e igual a zero para } t \geq n+1 .$$

A equação precedente é utilizada para gerar previsões inversas.

A expectância condicional de (3.67) conduz a:

$$(3.70) \quad \phi_p (B) [\omega_t] = \theta_q (B) [a_t]$$

onde,

$$[a_t] = E [a_t | \underset{\sim}{\phi}, \underset{\sim}{\theta}, \underset{\sim}{\omega}]$$

A equação (3.70) representa a condição necessária para se gerar os $[a_t]'$ s requeridos.

No caso de um modelo multiplicativo sazonal os mesmos procedimentos podem ser utilizados. A equação (3.34) mostra como os resíduos a_t dependem dos parâmetros. Então, ao se adotar o procedimento (ii), os processos atraso e adiantamento podem ser escritos como:

$$(3.71) \quad \phi_p(F) \Phi_P(F^S) \omega_t = \theta_q(F) \Theta_Q(F^S) e_t$$

e,

$$(3.72) \quad \phi_p(B) \Phi_P(B^S) \omega_t = \theta_q(B) \Theta_Q(B^S) a_t$$

onde,

$$\omega_t = \nabla^d \nabla_S^D \zeta_t$$

Tomando-se os valores esperados condicionais das equações anteriores, vem:

$$(3.73) \quad \phi_p(F) \Phi_P(F^S) [\omega_t] = \theta_q(F) \Theta_Q(F^S) [e_t]$$

e,

$$(3.74) \quad \phi_p(B) \Phi_P(B^S) [\omega_t] = \theta_q(B) \Theta_Q(B^S) [a_t]$$

A equação (3.73) é utilizada para gerar os "back-forecasts" ($[\omega_{-j}]$, $j = 0, 1, 2, \dots$) necessários a obtenção dos $[a_t]'$ s da equação (3.74).

Deve-se ressaltar que na estimação dos parâmetros pelo método dos mínimos quadrados podem ser obtidos também os respectivos erros padrão de estimativa, através da matriz de variância-covariância dos estimadores.

III.11 - TESTES DE VERIFICAÇÃO

Concomitantemente com os testes de verificação, chega-se ao último estágio da modelagem Box e Jenkins ou, em última análise, ao último estágio do ciclo ora proposto. Os testes, além de fornecerem uma medida da aderência do (s) modelo (s) filtrado (s) nos estágios anteriores, devem ser capazes de indicar uma maneira pela qual um modelo inadequado deve ser modificado. Tem-se, então, para cada modificação, uma repetição total ou parcial do ciclo Yevjevich-Box e Jenkins.

O sobreajustamento, assunto da seção III.11.1, é o primeiro dos testes considerados. Objetiva verificar se a identificação do modelo foi correta ou, se incorreu-se em um subajustamento, isto é, na adoção de um modelo com menos parâmetros que o necessário.

As seções III.11.2 a III.11.4 se relacionam com as importantes hipóteses postuladas aos resíduos: independência e normalidade.

A seção III.11.5 mostra a maneira pela qual os resíduos podem ser utilizados, para se efetuar modificações em um modelo inadequado.

III.11.1 - TESTE DE SOBREAJUSTAMENTO

Tendo-se identificado o que se acredita ser o modelo correto, pode-se colocá-lo em suspeição pelo ajustamento de um modelo mais elaborado. A aceitação de um modelo com um número adicional de parâmetros está condicionada aos estágios de estimação e verificação. Os parâmetros adicionais devem ser estimados para o modelo mais complexo somente onde se teme que o modelo mais simples seja fraco. Por exemplo, se a função de autocorrelação possui valores significantes nos "lags" 1 e 3 e houve razões para que um modelo MA (1) fosse previamente estimado, então, o modelo para verificar, através do sobreajustamento, o modelo estimado, seria um MA (3) com $\theta_2 = 0$.

Com auxílio do erro padrão da estimativa de θ_3 , pode-se decidir pela aceitação ou não do modelo mais completo. No entanto, a aceitação final fica condicionada aos testes de verificação.

Deve-se frisar que um cuidado especial deve ser tomado visto que uma redundância pode ocorrer quando se aumenta simultaneamente as componentes AR e MA de um modelo.

III.11.2 - ANÁLISE DAS AUTOCORRELAÇÕES DOS RESÍDUOS

Suponha-se que se tenha ajustado a uma série particular de dados $\{ Z_t, t = \overline{1, N} \}$ um modelo definido pela equação (2.34).

Com os estimadores de máxima verossimilhança, $\hat{\phi}$ e $\hat{\theta}$, pode-se obter uma estimativa dos resíduos através de,

$$(3.76) \quad \hat{a}_t = \hat{\theta}_q^{-1} (B) \hat{\phi}_p (B) \omega_t$$

É possível mostrar (BOX e JENKINS, 1970) que se o modelo \bar{e} é adequado, então,

$$(3.77) \quad \hat{a}_t = a_t + O\left(\frac{1}{\sqrt{n}}\right)$$

onde,

$$n = N-d,$$

ou seja, a menos de um infinitésimo de mesma ordem que $(n)^{-1/2}$, os resíduos estimados são iguais aos verdadeiros e desconhecidos resíduos a_t . Então, com o crescimento de n , a série de resíduos estimados se aproxima de um ruído branco.

Em particular, se o modelo \bar{e} é correto, suas autocorrelações, $R_k(\hat{a})$, serão para valores moderados e grandes de k , não correlacionadas e tenderão a uma distribuição normal assintótica, isto é,

$$(3.78) \quad R_k(\hat{a}) \sim N\left(0; \frac{1}{n}\right)$$

Na realidade, este resultado é sugerido por um outro devido ANDERSON (1942). Para as autocorrelações

estimadas dos resíduos verdadeiros, a_t , tem-se,

$$(3.79) \quad R_k(a) \sim N\left(0; \frac{1}{n}\right)$$

Entretanto, BOX e PIERCE (1970) mostraram que quando os resíduos são substituídos por suas estimativas, obtidas após se ajustar a forma correta do modelo, este resultado é distorcido para pequenos valores de k . Então, $R_k(\hat{a}_t)$ pode ser altamente correlacionada e suas variâncias podem ser muito menores que $\frac{1}{n}$.

III.11.3 - TESTE DE PORTMANTEAU (BOX e PIERCE)

O teste anterior considerava as autocorrelações $R_k(\hat{a})$ individualmente. Uma alternativa consiste em se considerar as L primeiras autocorrelações como um todo. Nisto baseia-se o teste de "portmanteau".

Suponha, então, que o modelo ajustado seja aquele dado pela equação (2.34). Então, se o modelo é apropriado, a estatística Q , definida por,

$$(3.80) \quad Q = n \sum_{k=1}^L (R_k(\hat{a}))^2$$

tem uma distribuição qui-quadrado assintótica com $\nu=L-p-q$ graus de liberdade. Se o modelo é inadequado, os valores de Q serão elevados.

Por intermédio da estatística Q pode-se formular um teste de hipóteses sobre a adequabilidade, ou não, do modelo.

Em relação ao número de autocorrelações a serem estimadas, deve-se tomar o valor de L suficientemente grande, de tal forma a garantir que os pesos ψ_j , do modelo escrito sob a forma,

$$(3.81) \quad \omega_t = \phi_p^{-1}(B) \theta_q(B) a_t = \psi(B) a_t$$

possam ser desprezados para $j > L$.

Para a situação mais geral, ou seja, quando o modelo apresenta componentes sazonais, tem-se que,

$$(3.82) \quad Q \sim \chi^2 (L-p-q-P-Q)$$

neste caso, deve-se ter $L \approx 4s$ (HIPEL, 1975).

III.11.4 - TESTE DO PERIODOGRAMA ACUMULADO

Um outro teste a ser aplicado consiste no periodograma normalizado acumulado da série de resíduos, $\{a_t, t = \overline{1, n}\}$, definido por,

$$(3.83) \quad C(f_j) = \frac{\sum_{i=1}^j I(f_i)}{n \hat{\sigma}_a^2}$$

onde,

$$(3.84) \quad I(f_i) = \frac{2}{n} \left[\left(\sum_{t=1}^n a_t \cos 2\pi f_i t \right)^2 + \left(\sum_{t=1}^n a_t \sin 2\pi f_i t \right)^2 \right]$$

é o periodograma da série $\{a_t\}$,

$f_i = \frac{i}{n}$, a i -ésima frequência e,

$$(3.85) \quad \hat{\sigma}_a^2 = \frac{1}{n} \sum_{i=1}^L I(f_i)$$

a estimativa da variância da série de resíduos, σ_a^2 , com $L = \frac{n}{2}$, se n é par ou, $L = \frac{n-1}{2}$ se n é ímpar.

A grande vantagem de se utilizar o periodograma acumulado se prende ao fato de que ele ajuda à identificação de não aleatoriedades periódicas remanescentes. Melhor dizendo, em algumas situações, particularmente quando se trabalha com séries de tempo sazonais, pode ocorrer que não se tenha adequadamente considerado as características periódicas da série.

Como, $C(f_j)$ representa uma estimativa de $P(f_j) / \sigma_a^2$ onde, $P(f_j)$ é o espectro acumulado, tem-se que o gráfico $C(f_j) \times f_j$, para um ruído branco, apresenta um comportamento linear entre os pontos $(0;0)$ e $(0,5;1)$. Em contrapartida, o periodograma acumulado para séries não aleatórias apresenta desvios significantes em relação à reta que passa por estes pontos.

Na prática, entretanto, não se dispõe dos a 's mas somente se seus valores estimados \hat{a} 's. No entanto, para grandes amostras o periodograma para os \hat{a} 's possui propriedades semelhantes ao dos a 's (BOX e JENKINS, 1970).

Por outro lado, o periodograma acumulado es

tã para o espectro acumulado da mesma forma que a função de frequência acumulada empírica está para a função de distribuição acumulada. Então, desde que se considere os \hat{a} 's como ruído branco, pode-se probabilizar os afastamentos do periodograma acumulado em torno do valor esperado (reta unindo os pontos (0;0) e (0,5;1)). Em particular, um teste como o de Kolmogorov-Smirnov pode ser utilizado (HIPEL, 1975)

A probabilidade assintótica de que,

$$(3.86) \quad \text{MAX}_j | C(f_j) - 2 f_j | > \frac{k_\lambda}{\sqrt{\left[\frac{n-1}{2} \right]}}$$

onde, $\left[\frac{n-1}{2} \right]$ representa a parte inteira do número dentro dos colchetes, $\bar{\epsilon}$ é o nível de significância,

$$(3.87) \quad \lambda = \sum_{u=-\infty}^{\infty} (-1)^u e^{-2k_\lambda^2 u^2}$$

Limites de confiança para o ruído branco podem ser desenhados no gráfico do periodograma acumulado através de linhas paralelas distando, para a direita e para

a esquerda, $k_\lambda / \sqrt{\left[\frac{n-1}{2} \right]}$ da reta que une os pontos (0;0) e (0,5;1).

Valores aproximados de k_λ em função do nível de significância λ são fornecidos na tabela que se segue.

λ	0,01	0,05	0,10	0,25
k_λ	1,63	1,36	1,22	1,02

TABELA 3.1 - VALORES DE k PARA O TESTE DE KOLMOGOROV-SMIRNOV.

III.11.5 - MODIFICAÇÃO DO MODELO ATRAVÉS DOS RESÍDUOS

Suponha que os resíduos b_t de um modelo in correto,

$$(3.88) \quad \phi_{p_0}(B) \nabla^{d_0} z_t = \theta_{q_0}(B) b_t$$

sejam não aleatórios. Admita, então, que com auxílio das funções de autocorrelação e autocorrelação parcial seja identificado, para a série $\{b_t, t = \overline{1, n}, n = N - d_0\}$, um modelo da forma,

$$(3.89) \quad \phi_{p_1}(B) \nabla^{d_1} b_t = \theta_{q_1}(B) a_t$$

Eliminando-se b_t nas equações (3.88) e (3.89) obtêm-se o novo modelo,

$$(3.90) \quad \phi_{p_0}(B) \phi_{p_1}(B) \nabla^{d_0} \nabla^{d_1} z_t = \theta_{q_0}(B) \theta_{q_1}(B) a_t$$

O modelo dado pela equação (3.90) deve, em

seguida, passar pelos estágios de estimação e testes de verificação.

Um desenvolvimento similar pode ser feito para o caso de uma série com componentes sazonais.

III.12 - PREVISÃO PARA O MODELO MISTO

III.12.1 - FORMAS EXPLÍCITAS DO MODELO PARA A COMPONENTE ESTOCÁSTICA DEPENDENTE PADRONIZADA ζ_t

Suponha que \tilde{a} componente estocástica dependente padronizada ζ_t seja ajustado um modelo multiplicativo ARIMA sazonal conforme a equação (3.34). Esta equação pode ser reescrita como,

$$(3.91) \quad \phi'(B) \zeta_t = \theta'(B) a_t$$

onde,

$$\begin{aligned} \phi'(B) &= \phi_p(B) \phi_P(B^S) \nabla^d \nabla_S^D \\ &= 1 - \phi'_1 B - \phi'_2 B^2 - \dots - \phi'_{p+sP+d+sD} B^{p+sP+d+sD} \end{aligned}$$

\tilde{e} o operador autoregressivo generalizado de ordem $p+sP+d+sD$ tal que $d+sD$ das raízes da equação característica $\phi'(B)=0$ estejam sobre e as demais raízes fora do círculo unitário.

$$\theta'(B) = \theta_q(B) \theta_Q(B^S) = 1 - \theta'_1 B - \theta'_2 B^2 - \dots - \theta'_{q+sQ} B^{q+sQ}$$

é o operador médias móveis generalizado de ordem $q + sQ$ tal que as raízes da equação característica $\theta'(B) = 0$ caiam fora do círculo unitário para garantir a unicidade (invertibilidade) do modelo.

A equação (3.91) pode, alternativamente, ser escrita de três formas diferentes:

(i) equação em diferenças

$$(3.92) \quad \zeta_t = \phi_1' \zeta_{t-1} + \phi_2' \zeta_{t-2} + \dots + \phi_{p+sP+d+sD}' \zeta_{t-p-sP-d-sD} + \\ + a_t - \theta_1' a_{t-1} - \theta_2' a_{t-2} - \dots - \theta_{q+sQ}' a_{t-q-sQ}$$

(ii) soma infinita de resíduos - de forma análoga à equação (2.21), tem-se,

$$(3.93) \quad \zeta_t = \psi'(B) a_t$$

ou

$$(3.93.a) \quad \zeta_t = a_t + \psi_1' a_{t-1} + \psi_2' a_{t-2} + \dots$$

onde,

$$\psi'(B) = \frac{\theta_q(B) \theta_Q(B^S)}{\nabla^d \nabla_S^D \phi_p(B) \Phi_P(B^S)} = 1 + \psi_1' B + \psi_2' B^2 + \dots$$

Os pesos ψ_j' ($j = 1, 2, \dots$) são obtidos igualando-se os coeficientes em B na equação,

$$(3.94) \quad \phi'(B) \psi'(B) = \theta'(B)$$

(iii) soma infinita dos valores passados da sêrie, somados a um resíduo - por analogia a equação (2.23), vem,

$$(3.95) \quad \pi' (B) \zeta_t = a_t$$

ou

$$(3.95.a) \quad \zeta_t = \pi'_1 \zeta_{t-1} + \pi'_2 \zeta_{t-2} + \dots + a_t$$

onde,

$$\pi'(B) = \frac{\phi_p(B) \phi_p(B^S) \nabla^d \nabla_S^D}{\theta_q(B) \theta_Q(B^S)} = 1 - \pi'_1 B - \pi'_2 B^2 - \dots$$

Os pesos π'_j ($j = 1, 2, \dots$) são obtidos igualando-se os coeficientes em B na equação,

$$(3.96) \quad \phi'(B) = \theta'(B) \pi'(B)$$

III.12.2 - DERIVAÇÃO DA PREVISÃO DE ERRO MÉDIO QUADRÁTICO MÍNIMO

Seja a sêrie $\{\zeta_t\}$, onde se conhece os valores $\zeta_1, \zeta_2, \dots, \zeta_t$. Deseja-se fazer uma previsão de valores futuros da sêrie, ζ_{t+L} ($L = 1, 2, 3, \dots$). Seja $\hat{\zeta}_t(L)$ a previsão do valor ζ_{t+L} . Suponha que $\hat{\zeta}_t(L)$ seja uma função linear de valores observados da sêrie,

$$(3.97) \quad \hat{\zeta}_t(L) = C_0(L) \zeta_t + C_1(L) \zeta_{t-1} + C_2(L) \zeta_{t+2} + \dots$$

Como,

$$\zeta_t = a_t + \psi_1' a_{t-1} + \psi_2' a_{t-2} + \dots$$

$$\zeta_{t-1} = a_{t-1} + \psi_1' a_{t-2} + \psi_2' a_{t-3} + \dots$$

⋮

$$\zeta_{t-j} = a_{t-j} + \psi_1' a_{t-j-1} + \psi_2' a_{t-j-2} + \dots$$

Então, pode-se substituir (3.97) e obter:

$$(3.98) \quad \hat{\zeta}_t(L) = \psi_L^* a_t + \psi_{L+1}^* a_{t-1} + \psi_{L+2}^* a_{t-2} + \dots$$

Deseja-se calcular ψ_L^* , ψ_{L+1}^* , A previsão de erro médio quadrático mínimo, $\hat{\zeta}_t(L)$, é aquela que minimiza a expressão:

$$(3.99) \quad E \{ \zeta_{t+L} - \hat{\zeta}_t(L) \}^2$$

Mas, fazendo $t = t+L$ na equação (3.93.a), vem:

$$(3.100) \quad \zeta_{t+L} = a_{t+L} + \psi_1' a_{t+L-1} + \dots + \psi_{L-1}' a_{t+1} + \psi_L' a_t + \psi_{L+1}' a_{t-1} + \dots$$

Então,

$$(3.101) \quad \{ \zeta_{t+L} - \hat{\zeta}_t(L) \}^2 = a_{t+L}^2 + \psi_1'^2 a_{t+L-1}^2 + \dots + \psi_{L-1}'^2 a_{t+1}^2 + \\ + (\psi_L' - \psi_L^*)^2 a_t^2 + (\psi_{L+1}' - \psi_{L+1}^*)^2 a_{t-1}^2 + \dots + \\ + \text{soma de produtos do tipo}$$

$$C_{ij} (a_i \times a_j ; i, j = 1, 2, \dots).$$

Tomando-se o valor esperado,

$$(3.102) \quad E \{ \zeta_{t+L} - \hat{\zeta}_t(L) \}^2 = \sigma_a^2 + \psi_1'^2 \sigma_a^2 + \dots + \psi_{L-1}'^2 \sigma_a^2 + \\ + (\psi_L' - \psi_L^*)^2 \sigma_a^2 + (\psi_{L+1}')^2 \sigma_a^2 + \dots + \\ + \text{zero} .$$

Como a primeira linha é uma constante positiva, o mínimo da expressão acima ocorre quando,

$$\psi_L^* = \psi_L'$$

$$\psi_{L+1}^* = \psi_{L+1}'$$

$$\psi_{L+2}^* = \psi_{L+2}'$$

·
·
·

E, portanto,

$$(3.103) \quad \hat{\zeta}_t(L) = \psi_L' a_t + \psi_{L+1}' a_{t-1} + \psi_{L+2}' a_{t-2} + \dots$$

Note-se que a equação (3.100) pode ser expressa como,

$$(3.104) \quad \zeta_{t+L} = e_t(L) + \hat{\zeta}_t(L)$$

onde,

$$(3.105) \quad e_t(L) = a_{t+L} + \psi'_1 a_{t+L-1} + \dots + \psi'_{L-1} a_{t+1}$$

representa o erro da previsão.

Seja, $\left[\zeta_{t+L} \right]$ a esperança condicional de ζ_{t+L} dado o conhecimento de todos os ζ 's até o instante t . Então,

$$(3.106) \quad \hat{\zeta}_t(L) = \psi'_L a_t + \psi'_{L+1} a_{t-1} + \psi'_{L+2} a_{t-2} + \dots = \left[\zeta_{t+L} \right]$$

ou seja, a previsão de erro médio quadrático mínimo, na origem t , para um período de tempo \tilde{a} frente $t+L$ é a expectativa condicional de ζ_{t+L} no instante t .

Tomando-se o valor esperado da equação (3.105), obtêm-se:

$$(3.107) \quad E \{ e_t(L) \} = 0$$

A equação (3.107) indica que a previsão é não tendenciosa. A variância do erro de previsão, vale:

$$(3.108) \quad \text{VAR} \{ e_t(L) \} = (1 + \psi_1'^2 + \psi_2'^2 + \dots + \psi_{L-1}'^2) \sigma_a^2$$

III.12.3- TRÊS FORMAS BÁSICAS PARA PREVISÃO

Sejam, respectivamente, $\left[a_{t+L} \right]$ e $\left[\zeta_{t+L} \right]$ as expectativas condicionais de a_{t+L} e ζ_{t+L} , tomadas no instante t .

Para $L > 0$, as três diferentes maneiras de expressar as previsões, são:

(i) previsões a partir da equação em diferenças - substituindo t por $t+L$ na equação (3.92) e tomando as esperanças condicionais, vem:

$$(3.109) \quad \left[\zeta_{t+L} \right] = \hat{\zeta}_t(L) = \phi'_1 \left[\zeta_{t+L-1} \right] + \dots + \phi'_{p+sP+d+sD} \left[\zeta_{t+L-p-sP-d-sD} \right] + \left[a_{t+L} \right] - \theta'_1 \left[a_{t+L-1} \right] - \dots - \theta'_{q+sQ} \left[a_{t+L-q-sQ} \right]$$

(ii) previsões a partir de uma soma infinita de resíduos tomando-se os valores esperados condicionais da equação (3.100), obtêm-se:

$$(3.110) \quad \zeta_{t+L} = \hat{\zeta}_t(L) = \left[a_{t+L} \right] + \psi'_1 \left[a_{t+L-1} \right] + \dots + \psi'_{L-1} \left[a_{t+1} \right] + \psi'_L \left[a_t \right] + \psi'_{L+1} \left[a_{t-1} \right] + \dots$$

(iii) previsões a partir de uma soma infinita dos valores passados da série, somados a um resíduo - fazendo $t = t + L$ na equação (3.95.a) e tomando-se as expectâncias condicionais, encontra-se:

$$(3.111) \quad \left[\zeta_{t+L} \right] = \hat{\zeta}_t(L) = \pi'_1 \left[\zeta_{t+L-1} \right] + \pi'_2 \left[\zeta_{t+L-2} \right] + \dots + \left[a_{t+L} \right]$$

Com relação ao cálculo dos valores esperados condicionais das equações (3.109), (3.110) e (3.111), note-se que se j é um inteiro não negativo, então,

$$(3.112) \left\{ \begin{array}{ll} \left[\zeta_{t-j} \right] & = \zeta_{t-j} \quad , j= 0,1,2,\dots \\ \left[\zeta_{t+j} \right] & = \hat{\zeta}_t(j) \quad , j= 1,2,3,\dots \\ \left[a_{t-j} \right] & = a_{t-j} = \zeta_{t-j} - \hat{\zeta}_{t-j-1}(1) \quad , j= 0,1,2,\dots \\ \left[a_{t+j} \right] & = 0 \quad , j= 1,2,3,\dots \end{array} \right.$$

No presente trabalho, as previsões para a componente ζ_t foram obtidas a partir do procedimento (i).

III.12.4 - DISTRIBUIÇÃO DAS PREVISÕES

Se a série ruído branco é normal e identicamente distribuída com média zero e variância σ_a^2 e, levando-se em conta as equações de (3.104) a (3.108), um futuro valor ζ_{t+L} do processo possui distribuição $N\left[\hat{\zeta}_t(L); \sigma_a^2 \sum_{j=0}^{L-1} \psi_j'\right]$, onde $\psi_0 = 1$.

Por conseguinte, pode-se facilmente obter os limites de confiança para as previsões.

Por outro lado, examinando-se a equação (3.108), nota-se que se $d+D > 0$, a sequência formada pelos pesos ψ_j' ($j= 1,2,\dots$) não converge, acarretando que $\text{VAR}\{e_t(L)\} \rightarrow \infty$

quando $L \rightarrow \infty$. De outra forma, se o modelo é estacionário ($d+D=0$), $\text{VAR} \{e_t(L)\} \rightarrow \gamma_0$ (variância da série de tempo) quando $L \rightarrow \infty$ e as previsões tendem ao nível médio da série.

Previsões feitas a partir da mesma origem t mas para diferentes períodos à frente, L e $L+m$, são correlacionadas (ANDERSON, 1975):

$$(3.113) \quad \text{COV} \left[e_t(L), e_t(L+m) \right] = \sigma_a^2 \sum_{j=0}^{L-1} \psi_j \psi_{j+m}$$

Isto significa que as previsões feitas a partir de uma origem fixa, para períodos de tempo à frente, tendem a superestimar ou a subestimar os valores reais. As covariâncias das previsões tomadas a partir de diferentes origens mas para o mesmo período de tempo à frente valem (ANDERSON, 1975):

$$(3.114) \quad \text{COV} \left[e_t(L), e_{t+m}(L) \right] = \begin{cases} 0 & , |m| \geq L \\ \sigma_a^2 \sum_{j=0}^{L-1-|m|} \psi_j \psi_{j+|m|} & , |m| < L \end{cases}$$

Para este caso, as previsões são correlacionadas somente entre as origens.

III.12.5 - ATUALIZAÇÃO DAS PREVISÕES

Utilizando as equações (3.110) e (3.112), a previsão na origem $t+1$, para um período de tempo L passos à frente, é dada por,

$$(3.115) \quad \hat{\zeta}_{t+1}(L) = \psi'_L a_{t+1} + \psi'_{L+1} a_t + \psi'_{L+2} a_{t-1} + \dots$$

A previsão para $L+1$ períodos de tempo à frente da origem t , vale:

$$(3.116) \quad \hat{\zeta}_t(L+1) = \psi'_{L+1} a_t + \psi'_{L+2} a_{t-1} + \psi'_{L+3} a_{t-2} + \dots$$

Subtraindo (3.116) de (3.115), vem:

$$(3.117) \quad \hat{\zeta}_{t+1}(L) = \hat{\zeta}_t(L+1) + \psi'_L a_{t+1}$$

A equação (3.117) é utilizada para se atualizar as previsões. Se as previsões foram determinadas na origem t para períodos de tempo $1, 2, \dots, L$, à frente, então, quando ζ_{t+1} estiver disponível, calcula-se $a_{t+1} = \zeta_{t+1} - \hat{\zeta}_t(1)$ e as previsões na nova origem $t+1$, para períodos de tempo $1, 2, \dots, L-1$, à frente, pela utilização da equação (3.117). A última previsão, $\hat{\zeta}_{t+1}(L)$, pode ser obtida a partir da equação em diferenças (3.109), em conjunção com a equação (3.112).

III.12.6 - PREVISÃO PARA A SÉRIE ORIGINAL

As previsões com os intervalos de confiança são calculadas inicialmente para a série transformada $\{Z_t^{(\lambda)}\}$. Isto é obtido pela integração da componente determinística do modelo com as previsões e intervalos de confiança para a componente estocástica dependente padronizada ζ_t (a partir de uma das equações de (2.11) a (2.18)). Em seguida, uma transformação no sentido inverso é feita para que se possa obter os resultados desejados.

Se, por exemplo, a transformação adotada for a logarítmica e, o modelo de composição para as médias e variâncias mensais for aquele dado pela equação (2.18), tem-se que a previsão para a variável transformada, L passos à frente da origem t, $\hat{Z}_t^{(\lambda)}(L)$, é dada por,

$$(3.118) \quad \hat{Z}_t^{(\lambda)}(L) = \bar{X}_T + S_T (\bar{\varepsilon} + S_\varepsilon \hat{\zeta}_t(L))$$

A previsão para a variável original é, neste caso, obtida como,

$$(3.119) \quad Z_t(L) = e^{\hat{Z}_t^{(\lambda)}(L)}$$

Com respeito aos limites dos intervalos de confiança de $(1 - \lambda)\%$ basta que na equação (3.118) se substitua a previsão $\hat{\zeta}_t(L)$ pelos valores correspondentes aos referidos limites e se resolva, a partir daí, a equação (3.119).

IV - APLICAÇÃO E DISCUSSÃO DOS RESULTADOS

IV.1 - INTRODUÇÃO

Este capítulo tem como objetivo a aplicação dos modelos apresentados anteriormente, com vistas a se obter previsões univariadas. Os dados da série hidrológica temporal utilizada foram fornecidos pelo Departamento de Engenharia de Geração da ELETROSUL. Referem-se às vazões médias mensais no período de janeiro de 1941 a dezembro de 1971, obtidas no posto pluviométrico de Salto Osório (tabela 4.1 e figura 4.2). Este posto está situado na bacia do Rio Iguaçu, na Região Sul do Brasil, cuja localização é mostrada na figura 4.1. A revisão bem como a consistência dos dados foi feita, segundo a ELETROSUL, pela Universidade Federal do Paraná, em 1973, tendo sido adotada, portanto, a hipótese de consistência e homogeneidade dos mesmos. A avaliação dos desempenhos dos modelos mostrados nas seções II.2 e II.3 com os desenvolvidos no capítulo III foi feita para o ano de 1971 (não pertencente a amostra utilizada na estimação dos parâmetros dos modelos) bem como para os anos de 1968 (ano seco), 1970 (ano médio) e 1957 (ano úmido).

O computador empregado para processamento dos programas, utilizados nas seções posteriores deste capítulo, foi um IBM/370 modelo 145 com 1024 K bytes de memória real e 16 M bytes de memória virtual do Centro de Processamento de Dados da ELETROBRÁS.

A seção IV.2 apresenta os resultados, quando da aplicação da metodologia de Yevjevich, em duas etapas: análise estrutural e previsão. Para o programa da primeira fase foram gastos 48,14 seg. de CPU e para o da segunda, 1,27 seg.

A seção IV.3 relaciona os resultados obtidos, pela aplicação da metodologia de Box e Jenkins, em quatro etapas: transformação, identificação, estimação e testes e previsão. O tempo de CPU gasto pelo primeiro dos programas foi de 2 min. e 26,78 seg.; para o segundo 1 min. e 41,26 seg.; o terceiro gastou 7 min. e 30,98 seg. e o quarto 46,08 seg.

Na seção IV.4 tem-se os resultados concernentes à nova metodologia proposta: integração entre a análise estrutural apresentada por Yevjevich com os modelos ARIMA de Box e Jenkins. As etapas foram: transformação, análise estrutural, identificação, estimação e testes e previsão. Para a segunda, um novo programa foi desenvolvido, gastando 51,40 seg. de CPU.

Finalmente, na seção IV.5, três critérios foram utilizados no sentido de se avaliar os resultados obtidos pelos três métodos diferentes: soma dos quadrados dos desvios das previsões em relação aos valores observados, desvio máximo em valor absoluto e, soma dos quadrados dos afastamentos dos limites superiores dos intervalos de 95% em relação aos valores observados.

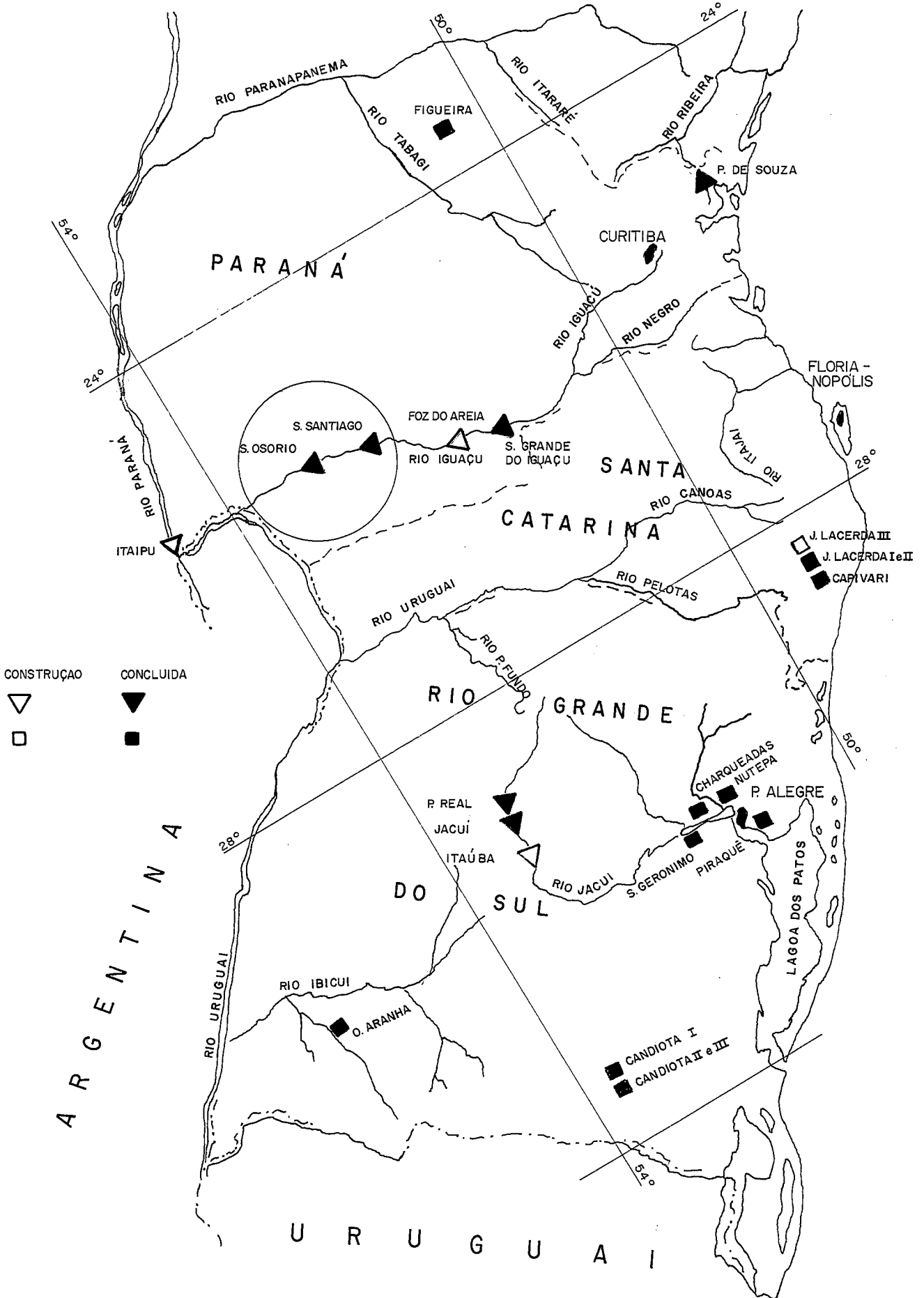
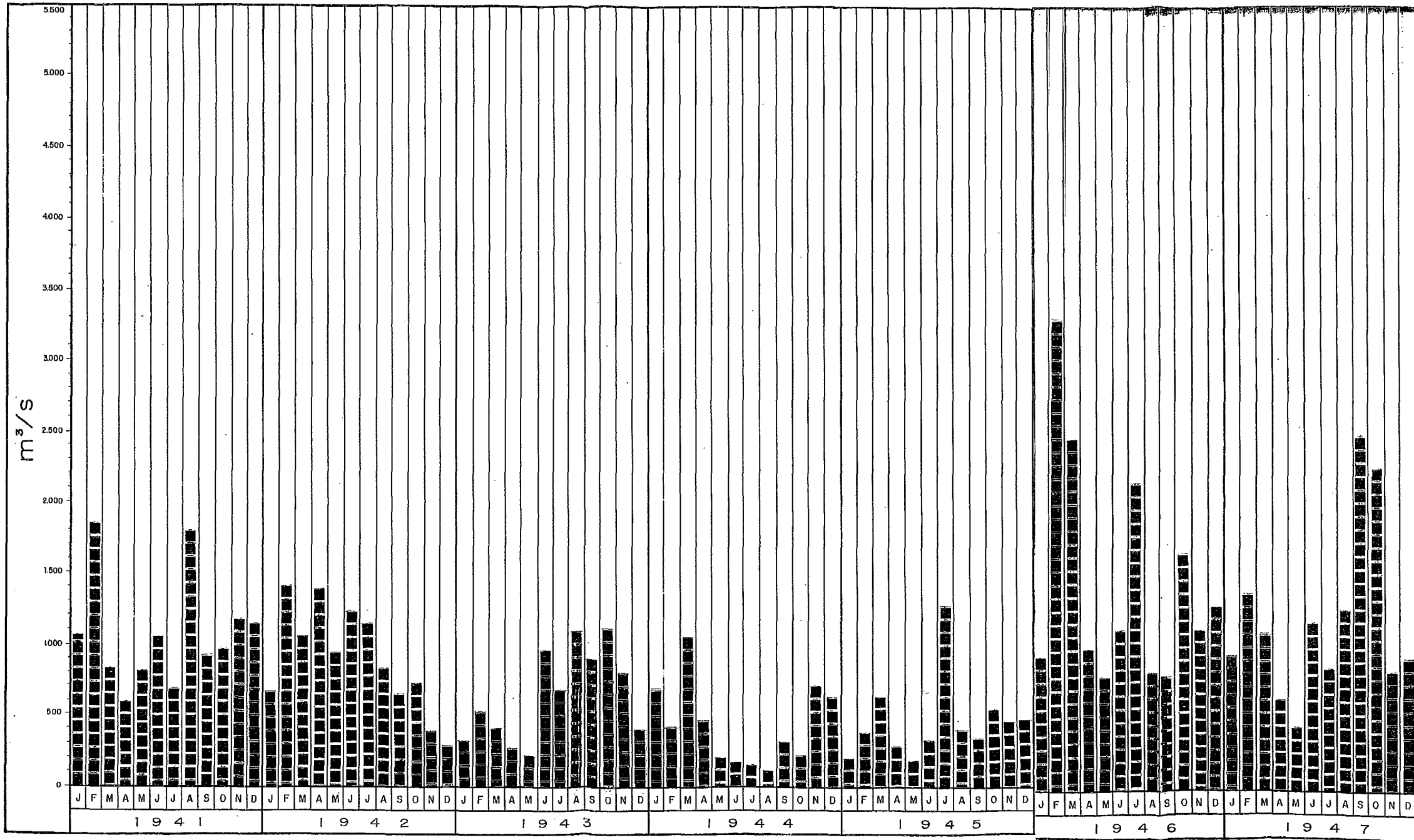
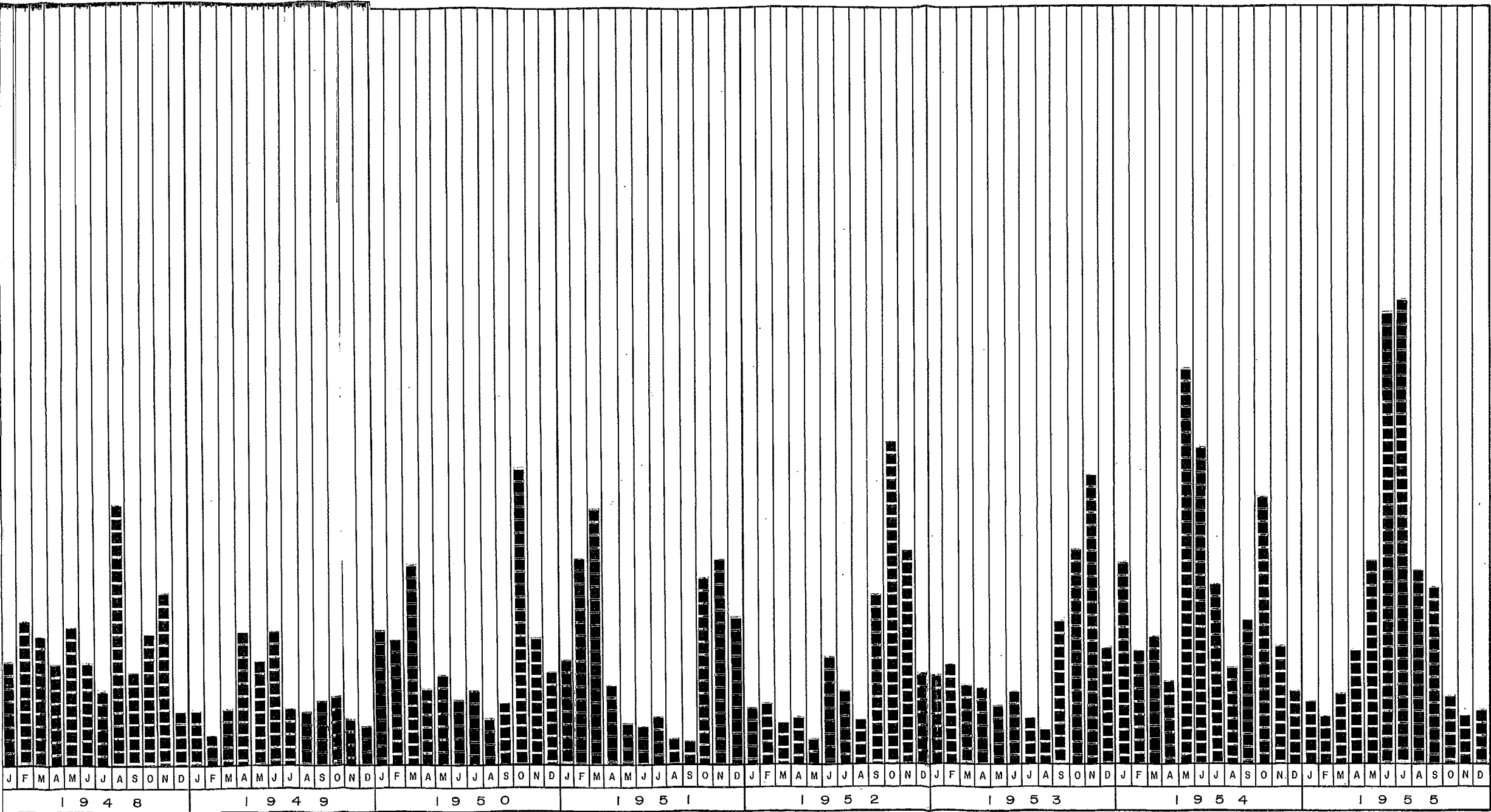


FIG. 4-1 - PRINCIPAIS USINAS GERADORAS DA REGIÃO SUL

A N O	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
1941	1075	1860	840	601	824	1069	688	1819	920	974	1183	1157
1942	686	1431	1074	1406	954	1256	1161	836	659	723	387	286
1943	335	531	420	264	219	972	690	1108	913	1119	804	416
1944	689	419	1070	468	211	175	157	118	315	220	692	627
1945	202	380	638	279	191	323	1288	405	345	552	466	470
1946	929	3301	2465	989	769	1125	2162	826	796	1648	1129	1295
1947	957	1393	1105	637	450	1193	863	1273	2485	2240	834	939
1948	745	1038	919	707	966	733	531	1872	670	918	1236	363
1949	370	211	405	943	745	944	404	363	430	468	308	259
1950	975	906	1455	536	646	465	536	306	445	2135	906	662
1951	760	1497	1824	562	285	268	332	173	159	1350	1480	1057
1952	408	439	301	320	150	762	506	308	1224	2344	1537	643
1953	631	716	553	544	421	526	332	245	1031	1561	2086	847
1954	1457	806	914	597	2872	2294	1296	674	1024	1925	844	518
1955	450	342	505	806	1456	3279	3365	1400	1270	464	342	385
1956	585	822	434	1371	1800	994	820	1323	1017	663	348	219
1957	432	1115	534	433	421	1090	3262	4796	5206	1887	1498	823
1958	538	387	928	478	259	517	485	865	1905	1031	936	980
1959	724	1024	518	475	603	764	482	587	868	556	331	238
1960	273	515	429	551	491	706	497	1423	1447	1451	1694	661
1961	579	511	1772	996	926	871	495	258	1294	1704	2240	1032
1962	638	1057	1247	553	352	512	369	271	1045	2192	978	477
1963	580	1135	924	842	328	280	178	159	192	1697	2742	983
1964	455	653	512	973	971	938	1106	1547	1506	946	512	588
1965	483	765	760	475	2456	893	2119	1115	1063	2487	1732	1992
1966	1180	2186	1417	524	425	647	956	465	1151	1710	1532	1238
1967	1475	1190	1588	780	332	771	594	608	764	542	674	980
1968	616	548	327	387	320	244	284	217	224	407	744	542
1969	1251	845	821	1722	1036	2116	1489	567	551	1333	1335	772
1970	736	625	517	456	651	1575	2121	541	693	1303	561	1147

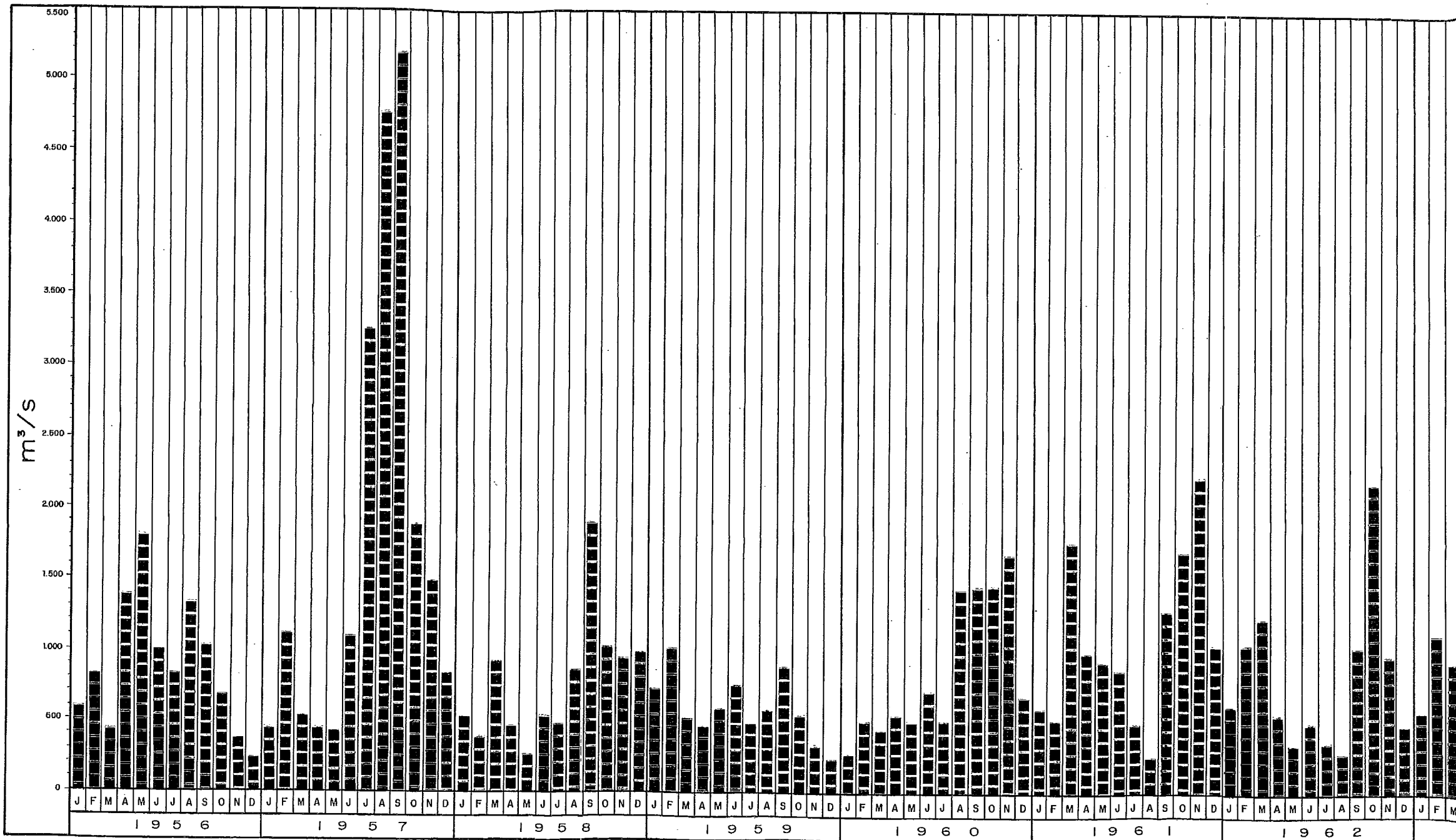
TABELA 4.1 - SÉRIE DE VAZÕES MÉDIAS MENSAIS DO POSTO DE SALTO OSÓRIO (m^3/s)





SALTO OSÓRIO - SÉRIE ORIGINAL

Paigun 10' - Part 01



IV.2 - PREVISÃO COM MODELOS DE YEVJEVICH

A análise estrutural proposta por Yevjevich foi realizada de duas formas diferentes: na série original e no logarítmo neperiano da série. Da primeira resultou o modelo A-1, onde as previsões foram obtidas através de uma distribuição log-normal de três parâmetros ajustada à série independente. O modelo A-2, resultante da transformação logarítma considerou uma distribuição normal ajustada à componente estocástica independente. Neste caso, as previsões para a série original foram tomadas como sendo os antilogaritmos das previsões para a série logaritmada.

As fases da análise estrutural para os modelos A-1 e A-2 foram:

1) COMPUTAÇÃO DAS ESTATÍSTICAS MENSAS DA SÉRIE

As estatísticas média, variância, desvio padrão, coeficientes de variação, assimetria e excesso, mensais, para os dois modelos, constam da tabela 4.2.

2) SEPARAÇÃO DAS COMPONENTES PERIÓDICA E ESTOCÁSTICA

Adotou-se, no presente trabalho, o método paramétrico para separar as componentes periódica e estocástica. Foi feita, também, a hipótese da existência de componentes periódicas de 12 meses nas médias e desvios padrões mensais. Para os dois modelos, a análise harmônica nas médias mensais forneceu os resultados mostrados na tabela 4.3. Os valores desta tabela foram estimados a partir das expressões (3.7) e (3.8).

MÊS	MÉDIA		VARIÂNCIA		D. PADRÃO		C. VARIACÃO		C. ASSIMETRIA		C. EXCESSO	
	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2
JAN	707	6,456	104400,88	0,218	323,11	0,467	0,457	0,072	0,822	-0,249	0,029	0,438
FEV	954	6,682	396847,81	0,352	629,96	0,593	0,660	0,089	1,891	0,121	4,285	-0,040
MAR	907	6,662	262904,25	0,294	512,74	0,542	0,565	0,081	1,144	0,167	0,931	-0,783
ABR	689	6,431	114884,19	0,201	338,95	0,449	0,492	0,070	1,348	0,311	1,428	-0,015
MAI	751	6,338	406323,25	0,540	637,43	0,735	0,849	0,116	1,891	0,323	3,225	-0,453
JUN	943	6,641	423650,63	0,435	650,88	0,659	0,690	0,099	1,824	-0,187	3,770	0,041
JUL	985	6,589	689551,94	0,600	830,39	0,775	0,843	0,118	1,553	0,188	1,655	-0,507
AGO	882	6,413	787053,69	0,734	887,16	0,857	1,006	0,134	2,839	0,112	9,845	-0,494
SET	1053	6,695	848908,00	0,541	921,36	0,736	0,874	0,110	3,061	-0,161	11,094	0,481
OUT	1285	6,997	419074,00	0,377	647,36	0,614	0,504	0,088	0,147	0,682	-1,151	-0,279
NOV	1069	6,804	371259,00	0,365	609,31	0,605	0,570	0,089	0,823	-0,198	0,153	-0,806
DEZ	753	6,486	149966,63	0,296	387,26	0,544	0,514	0,084	0,952	0,313	1,303	-0,364

TABELA 4.2 - ESTATÍSTICAS MENSIS PARA OS MODELOS A-1 E A-2 DA SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO

HARMÔNICO j	MODELO A-1			MODELO A-2		
	A _j	B _j	C _j	A _j	B _j	C _j
1	29,815	-149,699	152,639	0,092	-0,098	0,134
2	- 68,884	- 38,240	78,786	-0,066	-0,062	0,091
3	-143,118	- 75,106	161,629	-0,165	-0,095	0,190
4	- 0,811	1,300	1,532	0,021	0,034	0,040
5	18,200	1,326	18,248	-0,004	-0,012	0,013
6	2,800	0,000	2,800	0,009	0,000	0,009
COMP.CONTÍNUA	915,194			6,599		

TABELA 4.3 - ESTIMATIVAS DOS COEFICIENTES DE FOURIER E DAS AMPLITUDES DOS HARMÔNICOS DA FUNÇÃO PERIÔDICA CORRESPONDENTE ÀS MÉDIAS MENSAIS DA SÉRIE DE VAZÕES MENSAIS DE SALTO OSÓRIO.

A significância dos harmônicos foi testada segundo o procedimento empírico da seleção de dois valores críticos, P_{MIN} e P_{MAX} , para a proporção da variância do parâmetro explicada pelos seis harmônicos (YEVJEVICH, 1972). Para o modelo A-1, resultaram três harmônicos significativos e para o A-2, quatro. Em relação à classificação dos harmônicos em ordem decrescente de significância obteve-se, respectivamente: 3,1,2,5,6,4 e 3,1,2,4,5,6.

As médias mensais foram, então, ajustadas pelas seguintes funções periódicas:

MODELO A-1

$$\bar{X}_\tau = 915,194 - 143,118 \cos\left(\frac{\pi}{2} \cdot \tau\right) - 75,106 \sin\left(\frac{\pi}{2} \cdot \tau\right) + 29,815$$

$$\cos\left(\frac{\pi}{6} \cdot \tau\right) - 149,699 \sin\left(\frac{\pi}{6} \cdot \tau\right) - 68,884 \cos\left(\frac{\pi}{3} \cdot \tau\right) - 38,240$$

$$\sin\left(\frac{\pi}{3} \cdot \tau\right)$$

MODELO A-2

$$\bar{X}_\tau = 6,599 - 0,165 \cos\left(\frac{\pi}{2} \cdot \tau\right) - 0,095 \sin\left(\frac{\pi}{2} \cdot \tau\right) + 0,092 \cos\left(\frac{\pi}{6} \cdot \tau\right) +$$

$$-0,098 \sin\left(\frac{\pi}{6} \cdot \tau\right) - 0,066 \cos\left(\frac{\pi}{3} \cdot \tau\right) - 0,062 \sin\left(\frac{\pi}{3} \cdot \tau\right) + 0,021$$

$$\cos\left(\frac{2\pi}{3} \cdot \tau\right) + 0,034 \sin\left(\frac{2\pi}{3} \cdot \tau\right)$$

A tabela 4.4 apresenta, para os dois modelos, as médias mensais ajustadas.

MÊS	MODELO A-1	MODELO A-2
JAN	723,000	6,468
FEV	944,902	6,665
MAR	909,485	6,683
ABR	695,083	6,409
MAI	738,092	6,357
JUN	959,613	6,627
JUL	971,770	6,595
AGO	888,135	6,412
SET	1058,670	6,690
OUT	1270,421	7,001
NOV	1089,646	6,803
DEZ	733,007	6,482

TABELA 4.4 - MÉDIAS AJUSTADAS ÀS MÉDIAS MENSIS DA SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO.

A análise harmônica nas variâncias mensais, empregando as mesmas expressões forneceu, para cada modelo, os parâmetros A_j , B_j e C_j constantes da tabela 4.5.

HARMÔNICO j	MODELO A-1			MODELO A-2		
	A_j	B_j	C_j	A_j	B_j	C_j
1	-142302,25	-243735,00	282235,13	-0,121	-0,139	0,184
2	-101892,81	96174,00	140112,75	-0,013	0,061	0,062
3	- 31278,49	5985,99	31846,13	0,011	0,007	0,013
4	6788,38	- 91448,31	91699,88	-0,022	-0,086	0,088
5	36737,43	- 43280,64	56770,17	0,040	0,022	0,046
6	- 32655,86	0,00	- 32655,86	-0,014	0,000	-0,014
COMP.CONTÍNUA	414568,375			0,413		

TABELA 4.5 - ESTIMATIVAS DOS COEFICIENTES DE FOURIER E DAS AMPLITUDES DOS HARMÔNICOS DA FUNÇÃO PERIÓDICA CORRESPONDENTE ÀS VARIÂNCIAS DA SÉRIE DE VAZÕES MENSAIS DE SALTO OSÓRIO.

Para o primeiro modelo a ordem dos harmônicos foi: 1,2,4,5,6,3, sendo todos os seis considerados significativos. O segundo modelo mostrou a seguinte ordem: 1,4,2,5,6,3, sendo somente os quatro primeiros considera

dos significantes.

As funções harmônicas ajustadas às variâncias mensais foram:

MODELO A-1

$$\begin{aligned}
 S_{\tau}^2 = & 414568,375 - 142302,25 \cos\left(\frac{\pi}{6} \cdot \tau\right) - 243735,00 \sin\left(\frac{\pi}{6} \cdot \tau\right) + \\
 & -101892,81 \cos\left(\frac{\pi}{3} \cdot \tau\right) + 96174,00 \sin\left(\frac{\pi}{3} \cdot \tau\right) + 6788,38 \cos\left(\frac{2\pi}{3} \cdot \tau\right) + \\
 & -91448,31 \sin\left(\frac{2\pi}{3} \cdot \tau\right) + 36737,43 \cos\left(\frac{5\pi}{6} \cdot \tau\right) - 43280,64 \\
 & \sin\left(\frac{5\pi}{6} \cdot \tau\right) - 32655,86 \cos(\pi \cdot \tau) - 31278,49 \cos\left(-\frac{\pi}{2} \cdot \tau\right) + \\
 & + 5985,99 \sin\left(-\frac{\pi}{2} \cdot \tau\right)
 \end{aligned}$$

MODELO A-2

$$\begin{aligned}
 S_{\tau}^2 = & 0,413 - 0,12 \cos\left(-\frac{\pi}{6} \cdot \tau\right) - 0,14 \sin\left(-\frac{\pi}{6} \cdot \tau\right) - 0,02 \cos\left(-\frac{2\pi}{3} \cdot \tau\right) \\
 & - 0,09 \sin\left(-\frac{2\pi}{3} \cdot \tau\right) - 0,01 \cos\left(-\frac{\pi}{3} \cdot \tau\right) + 0,06 \sin\left(-\frac{\pi}{3} \cdot \tau\right) \\
 & + 0,04 \cos\left(-\frac{5\pi}{6} \cdot \tau\right) + 0,02 \sin\left(-\frac{5\pi}{6} \cdot \tau\right)
 \end{aligned}$$

Para os dois modelos, a função periódica ajustada aos desvios padrão mensais foi calculada como sendo a raiz quadrada de S_{τ}^2 .

As variâncias e os desvios padrão mensais ajustados, para os dois modelos, estão na tabela 4.6.

MÊS	MODELO A-1		MODELO A-2	
	VARIÂNCIAS	D. PADRÕES	VARIÂNCIAS	D. PADRÕES
JAN	104401,63	323,112	0,198	0,445
FEV	396847,88	629,959	0,377	0,614
MAR	262904,19	512,742	0,287	0,536
ABR	114884,75	338,947	0,204	0,451
MAI	406322,00	637,434	0,520	0,721
JUN	423651,19	650,885	0,459	0,678
JUL	689550,75	830,392	0,593	0,770
AGO	787052,50	887,160	0,737	0,858
SET	848907,38	921,362	0,521	0,722
OUT	419073,69	647,359	0,402	0,634
NOV	371260,25	609,311	0,358	0,599
DEZ	149965,69	387,254	0,298	0,546

TABELA 4.6 - VARIÂNCIAS E DESVIOS PADRÃO AJUSTADOS ÀS VARIÂNCIAS E DESVIOS PADRÃO MENSIS DA SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO.

Como para ambos os modelos, A-1 e A-2, a hipótese nula de que os coeficientes de variação mensais não diferissem de uma constante foi rejeitada, o modelo de composição, para a média e variância, ficou sendo aquele dado pela equação (2.18).

A média e o desvio padrão da componente estocástica dependente, $\varepsilon_{p,\tau}$, para os modelos A-1 e A-2, constam da tabela 4.8.

As componentes estocásticas dependentes padronizadas para os dois modelos foram computadas como sendo:

MODELO A-1:

$$\zeta_{p,\tau} = \varepsilon_{p,\tau} + 0,001$$

MODELO A-2

$$\zeta_{p,\tau} = \frac{\varepsilon_{p,\tau}}{1,002}$$

3) AJUSTAMENTO DE UM MODELO DE DEPENDÊNCIA À COMPONENTE ES TOCÁSTICA DEPENDENTE PADRONIZADA.

Os coeficientes de correlação mensais, estimados pela equação (3.36), estão na tabela 4.7. O teste de constância para estes coeficientes, explicado na seção II.2, indicou, para ambos os modelos, que eles poderiam ser considerados constantes, respectivamente iguais a 0,5239 e ã 0,5938.

MÊS	MODELO A-1	MODELO A-2
JAN	0,4970	0,6328
FEV	0,6872	0,6110
MAR	0,1875	0,3006
ABR	0,3310	0,5878
MAI	0,5803	0,6604
JUN	0,6863	0,7608
JUL	0,5808	0,6780
AGO	0,8337	0,7132
SET	0,3340	0,3808
OUT	0,5878	0,6727
NOV	0,5427	0,7020
DEZ	0,4395	0,4307
MÉDIA	0,5239	0,5938

TABELA 4.7 - COEFICIENTES DE CORRELAÇÃO MENSAIS DA SÉRIE DE VAZÕES MENSAIS DE SALTO OSÓRIO.

A escolha da ordem dos modelos autoregressivos foi feita de acordo com o método dos coeficientes de determinação, apresentado na seção II.2. Para o primeiro modelo, encontrou-se $R_1^2 = 0,2740$; $R_2^2 = 0,2741$ e $R_3^2 = 0,2799$. O modelo A-2, forneceu: $R_1^2 = 0,3516$; $R_2^2 = 0,3516$ e $R_3^2 = 0,3528$. Então, para ambos os casos, selecionou-se um modelo de dependência autoregressivo de 1ª ordem. Uma inspeção no correlograma da série dependente padronizada confirmou esta seleção.

4) AJUSTAMENTO DE UMA DISTRIBUIÇÃO DE PROBABILIDADE \bar{A} COM PONENTE ESTOCÁSTICA INDEPENDENTE.

\bar{A} componente estocástica independente, estimada a partir da equação (2.20) para $m=1$ e $\phi_1 = \bar{R}_{1,\tau}$, ajustou-se, para o modelo A-1, uma distribuição log-normal de três parâmetros, os quais foram estimados pelo método dos momentos (YEVJEVICH, 1970). As estimativas dos três parâmetros da distribuição log-normal constam da tabela 4.8.

Os testes de ajustamento apresentaram os seguintes resultados: QUI-QUADRADO: 19,250 e KOLMOGOROV-SMIRNOV: 0,055.

$$\text{Como, } \chi_{\text{OBS}}^2 = 19,250 < \chi_{\text{CRIT}(95\%,11)}^2 = 19,675 \text{ e}$$

$$\Delta_{\text{OBS}} = 0,055 < \Delta_{\text{CRIT}(95\%,360)} = 0,072, \text{ aceita-se, a um nível}$$

de significância de 5%, a hipótese de bom ajustamento.

Em relação ao modelo A-2, a distribuição de probabilidade ajustada foi a normal, cujos os parâmetros estimados, também, pelo método dos momentos, constam na ta bela 4.8.

Os resultados dos testes de ajustamento foram: QUI-QUADRADO: 23,917 e KOLMOGOROV-SMIRNOV: 0,062.

Como, $\chi^2_{OBS} = 23,917 < \chi^2_{CRIT(99\%,12)} = 26,2$ e

$\Delta_{OBS} = 0,062 < \Delta_{CRIT(95\%,360)} = 0,072$, aceita-se, a níveis

de significância de 1% e 5%, a hipótese de bom ajustamento.

Os parâmetros utilizados nos dois modelos encontram-se, de forma resumida, na tabela 4.8.

5) PREVISÃO

A maneira como se obtêm as previsões, L passos a frente de uma origem t, pré-fixada, a partir da metodologia de Yevjevich está descrita no APÊNDICE I.

As previsões e os intervalos de confiança de 50%, 90% e 95% pelos dois modelos para 1971 (ano não pertencente ao histórico), 1968 (ano seco), 1970 (ano médio) e 1957 (ano úmido) estão respectivamente, nas tabelas 4.9, 4.10, 4.11 e 4.12.

HARMÔNICO j	MODELO A-1			MODELO A-2		
	A _j	B _j	C _j	A _j	B _j	C _j
AN. HARMÔNICA NAS VARIÂNCIAS MENSASIS	29,815 - 68,884 - 143,118 - - -	- 149,699 - 38,240 - 75,106 - - -	152,639 78,786 161,629 - - -	0,092 - 0,066 - 0,165 0,021 - -	- 0,098 - 0,062 - 0,095 0,034 - -	0,134 0,091 0,190 0,040 - -
COMP. CONTÍNUA	915,194			6,599		
AN. HARMÔNICA NAS VARIÂNCIAS MENSASIS	- 142302,25 - 101892,81 - 31278,49 6788,38 36737,43 - 32655,86	-243735,00 96174,00 5985,99 - 91448,31 - 43280,64 0,00	282235,13 140112,75 31846,13 91699,88 56770,17 - 32655,86	- 0,121 - 0,013 - - 0,022 0,040 -	- 0,139 0,061 - - 0,086 0,022 -	0,184 0,062 - 0,088 0,046 -
COMP. CONTÍNUA	414568,375			0,413		
COMP. ESTOCÁSTICA DEPENDENTE	MÉDIA = - 0,001 D. PADRÃO = 1,000			MÉDIA = 0,000 D. PADRÃO = 1,002		
MODELO DE DEPENDÊNCIA	AUTOREGRESSIVO DE 1ª ORDEM $\phi_1 = \bar{R}_{1,T} = 0,5239$			AUTOREGRESSIVO DE 1ª ORDEM $\phi_1 = \bar{R}_{1,T} = 0,5938$		
COMP. ESTOCÁSTICA INDEPENDENTE	LOG-NORMAL DE 3 PARÂMETROS MÉDIA = 0,699 D. PADRÃO = 0,431 DESLOCAMENTO = - 2,210			NORMAL MÉDIA = 0,002 D. PADRÃO = 0,996		

TABELA 4.8 - PARÂMETROS DOS MODELOS A-1 E A-2 PARA A SÉRIE DE VAZÕES MENSASIS DE SALTO OSÓRIO.

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2						
JAN	3323	849	845	710	1035	665	1074	567	1424	468	1524	533	1585	419	1703
FEV	1212	967	978	666	1383	666	1437	370	2296	380	2521	303	2689	318	3012
MAR	1166	832	895	581	1182	630	1270	338	1960	377	2121	282	2297	321	2495
ABR	1238	611	642	444	844	476	866	283	1365	307	1342	246	1591	268	1541
MAI	2006	547	607	233	988	375	982	- 71	1970	186	1984	-140	2397	149	2478
JUN	2537	747	777	427	1197	494	1221	116	2201	255	2369	46	2638	206	2921
JUL	1818	689	744	280	1264	445	1245	-116	2545	210	2644	-206	3102	165	3355
AGO	822	580	615	143	1193	347	1092	-281	2563	150	2527	-377	3157	115	3295
SET	754	735	807	281	1372	498	1309	-159	2794	246	2651	-258	3412	197	3315
OUT	1003	1042	1099	723	1489	719	1679	414	2489	387	3121	344	2922	318	3798
NOV	437	874	900	574	1295	603	1344	283	2236	336	2414	217	2644	279	2905
DEZ	398	596	653	405	863	453	940	220	1461	266	1604	178	1721	224	1899

TABELA 4.9 - PREVISÕES DE VAZÕES MÉDIAS MENSIS PARA 1971 EM SALTO OSÓRIO.

METODOLOGIA DE YEVJEVICH.

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2						
JAN	616	777	783	638	962	616	995	495	1351	434	1413	461	1512	389	1578
FEV	548	892	919	592	1308	626	1350	296	2222	357	2369	228	2614	299	2830
MAR	327	800	866	550	1150	610	1230	306	1929	365	2054	251	2266	311	2416
ABR	387	600	632	433	833	469	852	272	1354	302	1321	235	1580	263	1517
MAI	320	536	598	223	977	370	967	- 82	1959	183	1954	-151	2386	147	2441
JUN	244	742	770	421	1192	490	1211	110	2196	252	2349	40	2632	205	2897
JUL	284	686	740	276	1260	442	1238	-120	2541	208	2629	-210	3098	164	3336
AGO	217	578	613	141	1191	345	1088	-283	2560	149	2518	-379	3155	114	3283
SET	224	734	806	280	1371	497	1306	-160	2793	245	2646	-259	3411	196	3308
OUT	407	1041	1098	722	1489	718	1677	413	2488	386	3118	343	2922	318	3794
NOV	744	874	900	573	1295	603	1343	282	2235	336	2412	217	2644	279	2904
DEZ	542	596	652	405	863	453	940	220	1461	265	1603	178	1721	224	1898

TABELA 4.10 - PREVISÕES DE VAZÕES MÉDIAS MENSIS PARA 1968 (ANO SECO) EM SALTO OSÓRIO.
METODOLOGIA DE YEVJEVICH.

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%			INT. DE 90%			INT. DE 95%					
		MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2				
												MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2
JAN	736	686	698	547	871	549	887	404	1260	387	1259	370	1421	346	1406
FEV	625	800	837	499	1216	570	1229	204	2130	325	2156	136	2522	272	2575
MAR	517	761	825	510	1111	581	1171	267	1889	348	1956	211	2226	296	2300
ABR	456	586	617	420	820	457	832	258	1341	295	1289	221	1567	257	1480
MAI	651	523	584	209	963	361	945	- 95	1946	179	1909	- 164	2373	143	2385
JUN	1575	734	760	414	1184	483	1196	103	2189	249	2319	33	2625	202	2859
JUL	2121	681	734	272	1255	439	1228	- 125	2536	207	2606	- 214	3093	163	3307
AGO	541	575	610	138	1189	343	1082	- 286	2558	148	2503	- 381	3152	114	3264
SET	693	733	804	279	1370	496	1302	- 161	2792	245	2639	- 261	3409	196	3299
OUT	1303	1041	1096	722	1488	717	1675	413	2488	386	3114	343	2921	317	3788
NOV	561	873	899	573	1295	602	1342	282	2235	335	2410	216	2643	279	2901
DEZ	1147	595	652	405	863	453	939	220	1461	265	1602	178	1720	224	1897

TABELA 4.11 - PREVISÕES DE VAZÕES MÉDIAS MENSIS PARA 1970 (ANO MÉDIO) EM SALTO OSÓRIO.

METODOLOGIA DE YEVJEVICH.

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2	MOD. A-1	MOD. A-2						
JAN	432	444	380	305	630	299	482	162	1019	210	685	128	1180	188	765
FEV	1115	554	508	253	970	346	747	- 42	1884	197	1310	- 110	2276	165	1565
MAR	534	656	638	406	1006	449	905	162	1785	269	1511	107	2122	229	1778
ABR	433	550	542	383	784	402	731	222	1304	260	1133	185	1530	226	1301
MAI	421	487	517	174	928	320	837	- 131	1910	158	1690	- 200	2337	127	2111
JUN	1090	715	710	395	1165	451	1117	84	2170	233	2166	14	2606	189	2671
JUL	3262	668	701	259	1242	419	1173	- 137	2524	197	2489	- 227	3080	155	3159
AGO	4796	568	591	131	1182	333	1049	- 293	2551	144	2429	- 388	3145	110	3167
SET	5206	729	792	275	1366	488	1283	- 165	2788	241	2599	- 265	3405	193	3250
OUT	1887	1039	1088	720	1487	712	1662	411	2486	383	3089	341	2920	315	3759
NOV	1498	873	895	572	1294	600	1336	281	2234	334	2400	216	2643	277	2888
DEZ	823	595	651	404	863	452	937	220	1461	265	1599	178	1720	224	1893

TABELA 4.12 - PREVISÕES DE VAZÕES MÉDIAS MENSIS PARA 1957 (ANO ÚMIDO) EM SALTO QSÓRIO.

METODOLOGIA DE YEVJEVICH.

IV.3 - PREVISÃO COM MODELOS DE BOX E JENKINS

Para a série cujos valores se encontram na tabela 4.1, aplicou-se a metodologia de Box e Jenkins com o objetivo de se prever os anos de 1971, 1968, 1970 e 1957.

Com respeito aos resultados propriamente ditos, pode-se destacar:

1) TRANSFORMAÇÃO DE PARÂMETROS

A transformação de Box e Cox, como descrita na seção III.2, foi aplicada à série original, tendo-se obtido para estimativas ótimas os valores: $\lambda_1^* = 0,0$, $\lambda_2^* = 20,0$ e $\chi^2 = 2,735$, os quais indicaram uma transformação do tipo: $\text{Ln}(Z_t + 20)$, onde Z_t é a série original.

Entretanto, obteve-se para $\lambda_1 = 0,0$ e $\lambda_2 = 0,0$ o valor $\chi^2 = 5,211$, sendo também aceita, neste caso, a hipótese nula a um nível de significância de 5% e 7 graus de liberdade. Em vista desse fato e também porque na prática, a transformação logarítmo natural é frequentemente razoável para dados sazonais de vazões e principalmente atendendo o critério da parcimônia, optou-se pela transformação:

$$Z_t^{(\lambda)} = \text{Ln}(Z_t)$$

A tabela 4.13 mostra o plano $(\lambda_1 \times \lambda_2)$, no entorno do ponto $\lambda_1 = 0,0$ e $\lambda_2 = 20,0$, enquanto que a figura 4.3, a transformação logarítmica adotada.

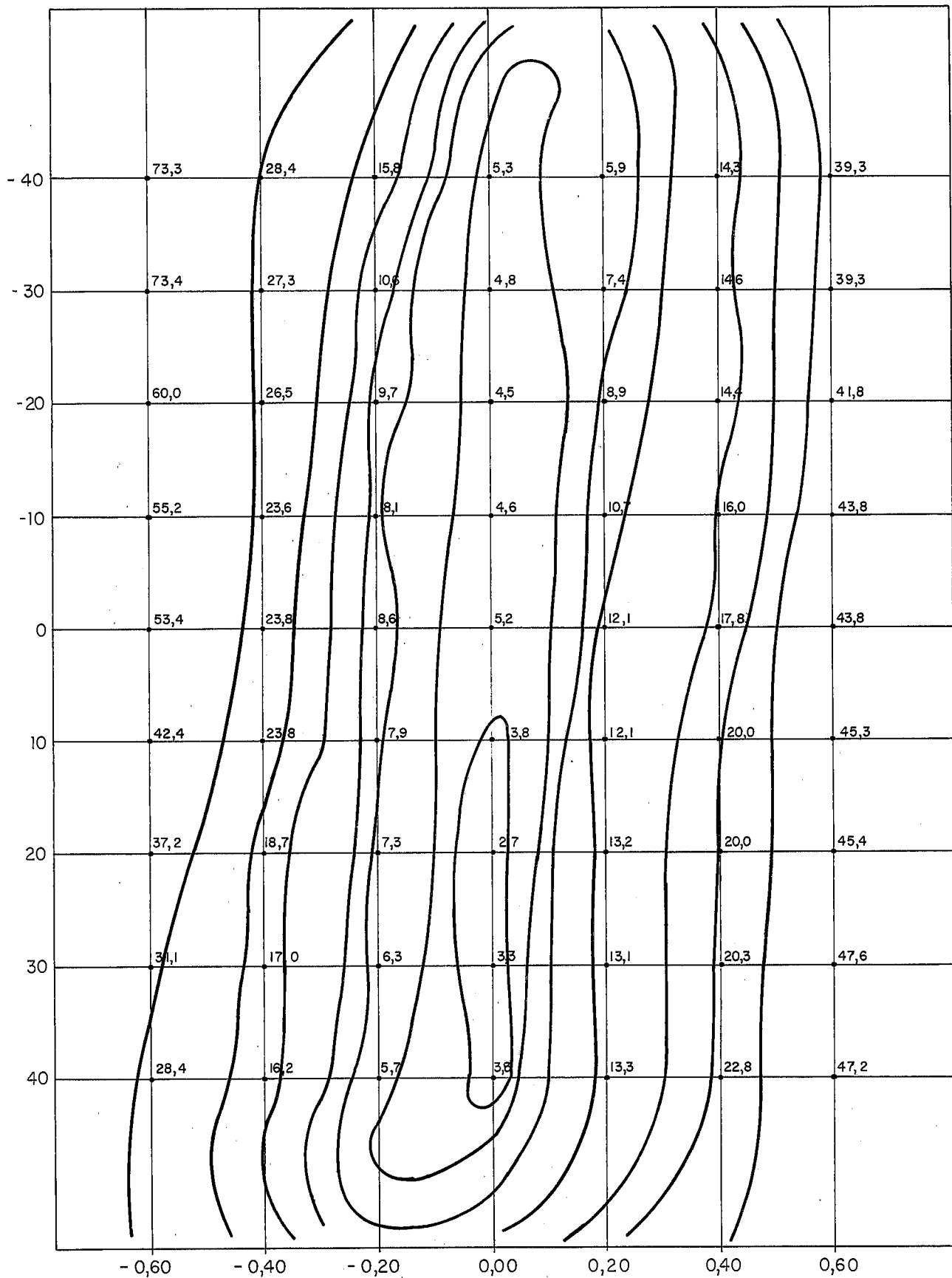


TABELA 4.13 - PLANO $(\lambda_1 \times \lambda_2)$ NO ENTORNO DO PONTO $\lambda_1 = 0,00$ e $\lambda_2 = 20,00$

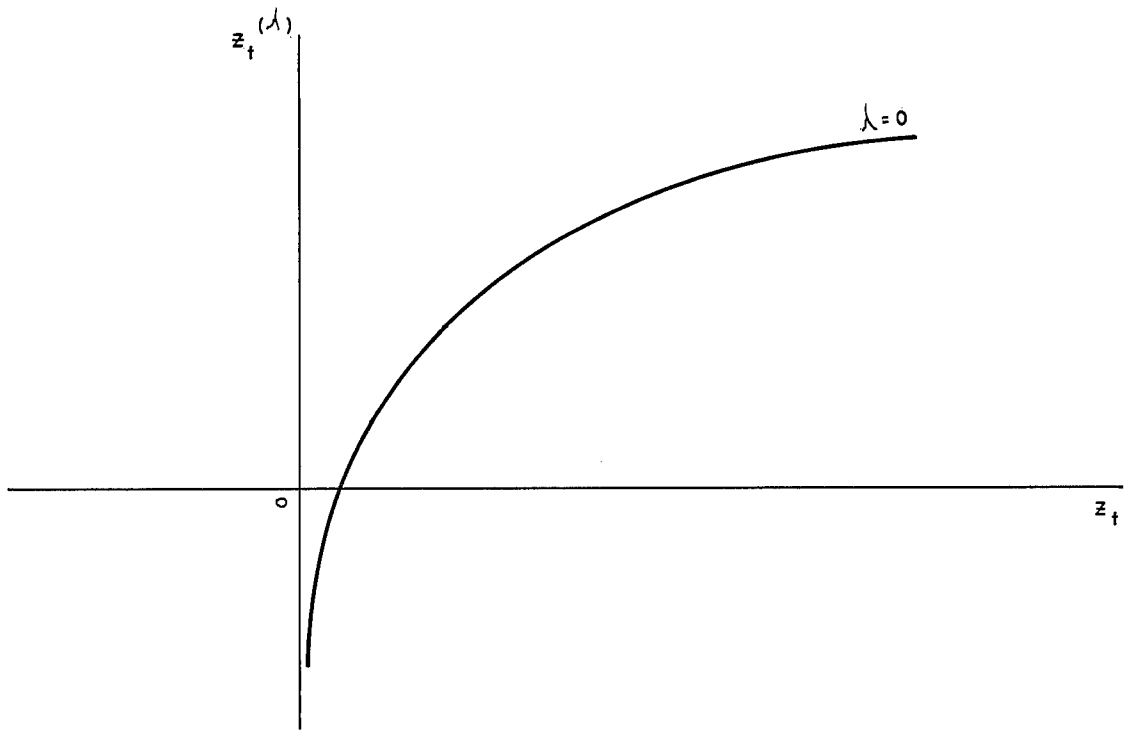


FIG. 4.3 — TRANSFORMAÇÃO ADOTADA

2) IDENTIFICAÇÃO DO (S) MODELO (S)

Uma inspeção nas funções de autocorrelação e autocorrelação parcial da série transformada $\{Z_t^{(\lambda)}\}$, estimadas a partir das equações (3.37) e (3.60), e nos seus respectivos gráficos indicou que uma diferença sazonal ($D=1$) era requerida. Então, a identificação do (s) modelo (s) foi feita para a série,

$$\omega_t^{(\lambda)} = \nabla_{12} Z_t^{(\lambda)}$$

Utilizando-se as equações (3.37) e (3.60), obteve-se as funções de autocorrelação e autocorrelação

parcial da sêrie $\omega_t^{(\lambda)}$. Os modelos identificados inicialmente, escritos de acordo com a equação (2.42), foram:

$$B-1: (1-\phi_1 B) \nabla_{12} Z_t^{(\lambda)} = (1 - \theta_1 B^{12}) a_t$$

$$B-2: (1 - \phi_1 B) (1 - \Phi_1 B^{12}) \nabla_{12} Z_t^{(\lambda)} = a_t$$

ou, de forma compacta:

$$B-1: (1,0,0) \times (0,1,1)_{12}$$

$$B-2: (1,0,0) \times (1,1,0)_{12}$$

3) ESTIMAÇÃO DOS PARÂMETROS DOS MODELOS B-1 e B-2

De acordo com a têcnica "back forecasting", descrita na seção III.10, estimativas aproximadas de máxima verossimilhança foram obtidas para os parâmetros dos dois modelos. A tabela 4.14 resume os resultados obtidos.

	COEFICIENTES	ERRO PADRÃO DA ESTIMATIVA	VARIÂNCIA DOS RESÍDUOS
B-1	$\hat{\phi}_1 = 0,6065$	0,0427	0,318846
	$\hat{\theta}_1 = 0,8375$	0,0305	
B-2	$\hat{\phi}_1 = 0,6033$	0,0428	0,405483
	$\hat{\Phi}_1 = -0,3586$	0,0507	

TABELA 4.14 - PARÂMETROS DOS MODELOS B-1 e B-2 PARA A SÊRIE DE VAZÕES MENSIS DE SALTO OSÓRIO.

4) TESTES DE VERIFICAÇÃO APLICADOS AOS MODELOS B-1 e B-2

Os testes aplicados aos modelos, explicados na seção III.11, foram:

(i) correlograma dos resíduos - os limites de 95% de confiança para as autocorrelações das estimativas dos resíduos, obtidos a partir da equação (3.78), mostraram que enquanto o modelo B-1 era adequado, o modelo B-2, por possuir os valores $R_{12}(\hat{a})$ e $R_{24}(\hat{a})$ excedendo os limites de confiança, poderia ser inadequado;

(ii) periodograma acumulado - os periodogramas acumulados para os dois modelos, obtidos a partir da equação (3.83), apresentaram gráficos que se assemelharam bastante ao gráfico de um ruído branco;

(iii) teste de portmanteau - as estatísticas Q , estimadas pela equação (3.82), indicaram que a hipótese de adequade do modelo B-1 era aceita a 95% de confiança, enquanto que para o modelo B-2, esta hipótese era rejeitada. Os resultados dos testes de portmanteau se encontram na tabela 4.15.

		MOD. B-1	MOD. B-2	QUI-QUADRADO TAB. (95%,40)
Q	PARA			
OBS		43,449	79,659	55,8
40	G L			

TABELA 4.15 - TESTE DE PORTMANTEAU PARA OS MODELOS B-1 E B-2 DA SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO.

(iv) sobreajustamento - em se levando em conta os resultados dos testes anteriores, os modelos selecionados para sobreajustar o modelo B-1, escritos de forma compacta, foram:

$$B-3: (2,0,0) \times (0,1,1)_{12}$$

$$B-4: (2,0,0) \times (1,1,0)_{12}$$

$$B-5: (1,0,0) \times (0,1,2)_{12}$$

$$B-6: (1,0,0) \times (1,1,1)_{12}$$

5) ESTIMAÇÃO DOS PARÂMETROS PARA OS MODELOS B-3, B-4, B-5 E B-6.

Estimativas aproximadas de máxima verossimilhança para os parâmetros dos modelos B-3, B-4, B-5 e B-6 e seus respectivos erros padrão encontram-se na tabela 4.16.

	COEFICIENTES	ERRO PADRÃO DA ESTIMATIVA	VARIÂNCIA DOS RESÍDUOS
B-3	$\hat{\phi}_1 = 0,6094$ $\hat{\phi}_2 = -0,0051$ $\hat{\theta}_1 = 0,8374$	0,0537 0,0538 0,0305	0,318837
B-4	$\hat{\phi}_1 = 0,5789$ $\hat{\phi}_2 = 0,0406$ $\hat{\phi}_1 = -0,3618$	0,0537 0,0537 0,0507	0,404818
B-5	$\hat{\phi}_1 = 0,5976$ $\hat{\theta}_1 = 0,6693$ $\hat{\theta}_2 = 0,1907$	0,0432 0,0534 0,0537	0,308887
B-6	$\hat{\phi}_1 = 0,5974$ $\hat{\phi}_1 = 0,1881$ $\hat{\theta}_1 = 0,8873$	0,0431 0,0642 0,0314	0,310600

TABELA 4.16 - PARÂMETROS DOS MODELOS B-3, B-4, B-5 e B-6 PARA A SÉRIE DE VAZÕES MENSAIS DE SALTO OSÓRIO.

6) TESTES DE VERIFICAÇÃO PARA OS MODELOS B-3, B-4, B-5 E B-6.

Inicialmente aplicou-se um teste para verificar a hipótese de que os parâmetros $\hat{\phi}_2$ dos modelos B-3 e B-4 fossem iguais a zero. Este teste foi realizado em vista dos estimadores de máxima verossimilhança serem assintoticamente normais e eficientes, e porque aqueles parâmetros apresentaram estimativas próximas de zero. Para ambos os modelos, a hipótese nula de que $\phi_2=0$ foi considerada consistente a um nível de significância de 5%. Então, prosseguiu-se com os demais testes apenas para os modelos B-5 e B-6, em virtude dos modelos B-3 e B-4 terem recaído nos modelos B-1 e B-2 anteriormente analisados.

O teste do correlograma dos resíduos apresentou, para os modelos B-5 e B-6, apenas um único valor não pertencente ao intervalo de confiança das autocorrelações estimadas dos resíduos ("lag 18"). Como o número de autocorrelações testadas foi de 43, aceita-se que este valor corresponda à faixa de rejeição de 5% do teste, ou seja, aceita-se, a um nível de significância de 5% que os resíduos destes modelos representem estimativas do ruído branco.

Os gráficos dos periodogramas acumulados dos modelos B-5 e B-6 mostraram um comportamento bastante linear entre os pontos (0;0) e (0,5;1), indicando a adequação dos modelos.

Quanto ao teste de portmanteau, a hipótese nula de que os modelos B-5 e B-6 eram adequados foi aceita para um grau de confiança de 95%. A tabela 4.17 sintetiza os resultados deste teste.

		MOD. B-5	MOD. B-6	QUI-QUADRADO TAB. (95%, 40)
Q OBS 40	PARA G L	36,400	37,623	55,8

TABELA 4.17 - TESTE DE PORTMANTEAU PARA OS MODELOS B-5 E B-6 DA SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO.

Tendo-se em vista os resultados dos testes aplicados aos resíduos dos modelos B-5 e B-6 e, ressaltando-se mais uma vez o critério da parcimônia, não se prosseguiu com o teste de sobreajustamento.

Com relação aos modelos B-5 e B-6 que serão levados à fase de previsão, note-se que eles podem ser escritos, de acordo com a equação (2.42), como:

MODELO B-5:

$$(1 - \hat{\phi}_1 B) \omega_t^{(\lambda)} = (1 - \hat{\theta}_1 B^{12} - \hat{\theta}_2 B^{24}) a_t$$

onde,

$$\omega_t^{(\lambda)} = \nabla_{12} Z_t^{(\lambda)} = (1 - B^{12}) \text{Ln} Z_t$$

ou,

$$(1-0,5976 B) \omega_t^{(\lambda)} = (1-0,6693 B^{12} - 0,1907 B^{24}) a_t$$

MODELO B-6:

$$(1-\hat{\phi}_1 B) (1-\hat{\phi}_1 B^{12}) \omega_t^{(\lambda)} = (1-\hat{\theta}_1 B^{12}) a_t$$

ou,

$$(1-0,5974 B) (1-0,1881 B^{12}) \omega_t^{(\lambda)} = (1-0,8873 B^{12}) a_t$$

7) PREVISÃO

As previsões e os intervalos de confiança de 50%, 90% e 95%, obtidas pelos modelos B-5 e B-6, para os anos de 1971, 1968, 1970 e 1957 constam, respectivamente, das tabelas 4.18, 4.19, 4.20 e 4.21.

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6						
JAN	3323	899	955	618	1308	656	1391	360	2243	382	2388	303	2672	320	2847
FEV	1212	1137	1246	734	1760	804	1931	392	3298	428	3625	320	4045	349	4447
MAR	1166	850	914	538	1342	578	1445	279	2589	299	2791	225	3205	241	3457
ABR	1238	499	532	314	794	334	846	161	1546	171	1652	130	1920	138	2052
MAI	2006	467	391	293	745	245	624	150	1457	125	1222	121	1811	101	1520
JUN	2537	813	814	509	1297	510	1301	260	2539	260	2551	209	3158	209	3175
JUL	1818	1074	990	672	1714	619	1582	343	3356	316	3104	276	4175	254	3863
AGO	822	456	417	285	727	261	667	146	1425	133	1309	117	1772	107	1630
SET	754	786	795	492	1255	497	1270	251	2458	253	2493	202	3058	204	3103
OUT	1003	1245	1211	780	1988	758	1936	398	3895	386	3799	320	4845	310	4728
NOV	437	909	979	570	1452	613	1566	291	2844	312	3072	234	3538	251	3824
DEZ	398	1232	1316	772	1967	823	2103	394	3854	419	4127	317	4794	337	5137

TABELA 4.18 - PREVISÕES DE VAZÕES MÉDIAS MENSIS PARA 1971 EM SALTO OSÓRIO
METODOLOGIA DE BOX E JENKINS

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%			INT. DE 90%			INT. DE 95%					
		MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6		
JAN	616	908	920	624	1322	632	1340	364	2266	368	2301	306	2700	309	2743
FEV	548	1455	1635	940	2252	1055	2533	508	420	562	4755	409	5175	458	5834
MAR	327	1296	1367	819	2043	864	2162	425	3941	447	4176	343	4878	361	5172
ABR	387	550	538	346	875	338	857	178	1705	173	1673	143	2117	139	2078
MAI	320	371	274	233	591	172	438	119	1156	88	858	96	1437	71	1067
JUN	244	660	617	413	1053	386	985	211	2061	197	1932	170	2563	158	2404
JUL	284	794	702	497	1267	439	1121	254	2481	224	2200	204	3087	180	2738
AGO	217	497	425	311	793	266	679	159	1553	135	1332	128	1932	109	1658
SET	224	986	1011	618	1574	632	1615	315	3084	322	3170	253	3836	259	3945
OUT	407	1150	1118	720	1835	699	1786	368	3595	356	3506	295	4473	286	4363
NOV	744	1075	1170	673	1717	732	1870	344	3363	373	3670	276	4183	300	4568
DEZ	542	1399	1560	876	2234	976	2494	447	4377	497	4894	360	5445	400	6092

TABELA 4.19 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS PARA 1968 (ANO SECO) EM SALTO OSÓRIO
METODOLOGIA DE BOX E JENKINS

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD.B-5	MOD.B-6	MOD.B-5	MOD.B-6	MOD.B-5	MOD.B-6	MOD.B-5	MOD.B-6						
JAN	736	906	888	623	1319	609	1293	363	2261	355	2220	305	2693	298	2646
FEV	625	1217	1276	786	1884	823	1978	419	3530	439	3713	342	4329	358	4555
MAR	517	1017	1007	644	1606	637	1592	334	3097	330	3076	270	3834	266	3810
ABR	456	735	690	462	1169	433	1098	237	2277	222	2143	191	2828	179	2663
MAI	651	524	396	328	835	248	632	168	1632	127	1238	135	2029	102	1539
JUN	1575	903	790	566	1440	494	1261	289	2820	252	2474	232	3507	202	3078
JUL	2121	989	831	619	1579	520	1329	316	3092	265	2607	254	3846	213	3245
AGO	541	481	409	301	768	256	653	154	1504	130	1281	124	1871	105	1595
SET	693	806	773	505	1287	484	1236	258	2521	247	2426	207	3136	198	3020
OUT	1303	1333	1212	835	2129	758	1938	426	4170	386	3803	343	5187	310	4733
NOV	561	1238	1271	775	1976	795	2032	396	3871	405	3988	318	4816	326	4964
DEZ	1147	1166	1239	730	1862	775	1981	373	3647	395	3888	300	4537	317	4839

TABELA 4.20 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS PARA 1970 (ANO MÉDIO) EM SALTO OSÓRIO.
METODOLOGIA DE BOX E JENKINS.

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6						
										MOD. B-5	MOD. B-6	MOD. B-5	MOD. B-6		
JAN	432	259	228	178	377	157	332	104	646	91	570	87	770	77	680
FEV	1115	376	308	243	582	199	477	130	1091	106	896	106	1338	86	1099
MAR	534	300	240	190	473	152	380	98	913	79	734	79	1130	63	909
ABR	433	487	450	306	775	283	717	157	1510	145	1400	127	1875	117	1739
MAI	421	445	409	279	709	256	653	143	1386	131	1280	115	1723	105	1592
JUN	1090	861	1018	540	1373	637	1626	276	2689	325	3189	222	3344	261	3968
JUL	3262	625	699	391	998	438	1118	200	1954	223	2193	161	2438	180	2729
AGO	4796	525	572	329	838	358	914	168	1641	182	1794	135	2042	147	2233
SET	5206	1080	1385	677	1725	866	2213	346	3379	441	4343	278	4203	355	5406
OUT	1887	1361	1399	853	2173	675	2236	435	4257	446	4388	350	5296	358	5462
NOV	1498	956	1054	599	1527	660	1685	306	2990	336	3307	246	3720	270	4116
DEZ	823	580	640	363	926	400	1022	185	1814	204	2006	150	2257	164	2497

TABELA 4.21 - PREVISÕES DE VAZÕES MÉDIAS MENSIAIS PARA 1957 (ANO ÚMIDO) EM SALTO OSÓRIO
METODOLOGIA DE BOX E JENKINS

IV.4 - PREVISÃO COM MODELOS MISTOS YEVJEVICH-BOX e JENKINS

Inicialmente as previsões foram calculadas para a componente estocástica dependente padronizada . Esta componente foi obtida a partir do modelo de composição selecionado na análise estrutural procedida na série transformada $\{Z_t^{(\lambda)}\}$. Este modelo de composição, para a média e variância mensais, em combinação com uma transformação inversa, forneceu os resultados requeridos.

As fases e os resultados obtidos pela integração das metodologias propostas por YEVJEVICH (1972) e por BOX e JENKINS (1970) estão relacionadas a seguir:

1) TRANSFORMAÇÃO DE PARÂMETROS

A mesma transformação de parâmetros explicada na seção IV.3 foi realizada na série original. São válidos, portanto, os resultados apresentados na tabela 4.13, bem como a figura 4.3, representando a transformação adotada.

2) ANÁLISE ESTRUTURAL PARA A SÉRIE TRANSFORMADA

Desde que a transformação adotada foi a logaritmo natural, as estatísticas mensais da série transformada são as mesmas obtidas para o modelo A-2, da seção IV.2. Os resultados constam da tabela 4.2.

Para a separação das componentes periódica e estocástica do modelo foi adotado o método paramétrico ,

descrito em II.2, em virtude da hipótese básica d , da seção III.3.

O modelo de composição para a média e variância mensais foi obtido a partir dos testes de constância, descritos em III.7, realizados nas médias, variâncias e coeficientes de variação mensais. Os resultados dos testes estão, respectivamente, nas figuras 4.4, 4.5 e 4.6 e indicaram o modelo de composição semelhante ao da equação (2.18), ou seja, tanto a média quanto a variância e coeficiente de variação mensais são periódicos

Como o modelo de composição foi aquele dado pela equação (2.18) uma análise harmônica nas médias e variâncias foi realizada. Com relação aos coeficientes de Fourier estimados pelas equações (3.7) e (3.8) e das amplitudes dos harmônicos, são válidos os mesmos resultados encontrados para o modelo A-2, os quais se encontram relacionados, respectivamente, nas tabelas 4.3 e 4.5.

A significância dos harmônicos foi testada de acordo com o procedimento descrito em III.6. Para as médias mensais foram considerados significativos os harmônicos de ordem 3, 1 e 2; para as variâncias mensais, os de ordem 1, 4, 2 e 5. Note-se que quanto às médias houve uma economia de um harmônico em relação ao teste empírico proposto por Yevjevich. Às médias e variâncias mensais ajustou-se as seguintes funções periódicas:

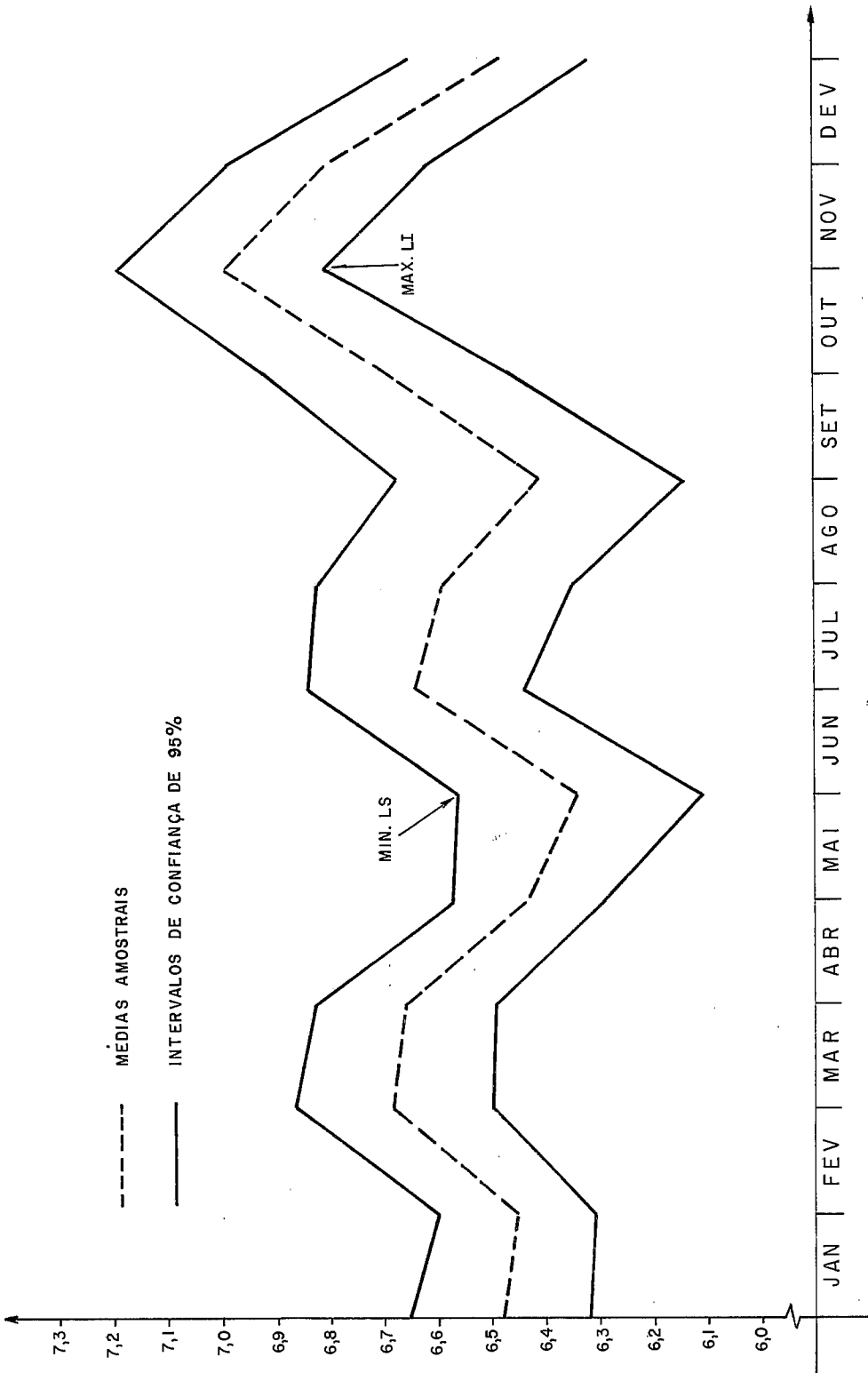


FIGURA 4.4 - TESTE DE CONSTÂNCIA PARA AS MÉDIAS MENSIS DA SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO (SÉRIE TRANSFORMADA)

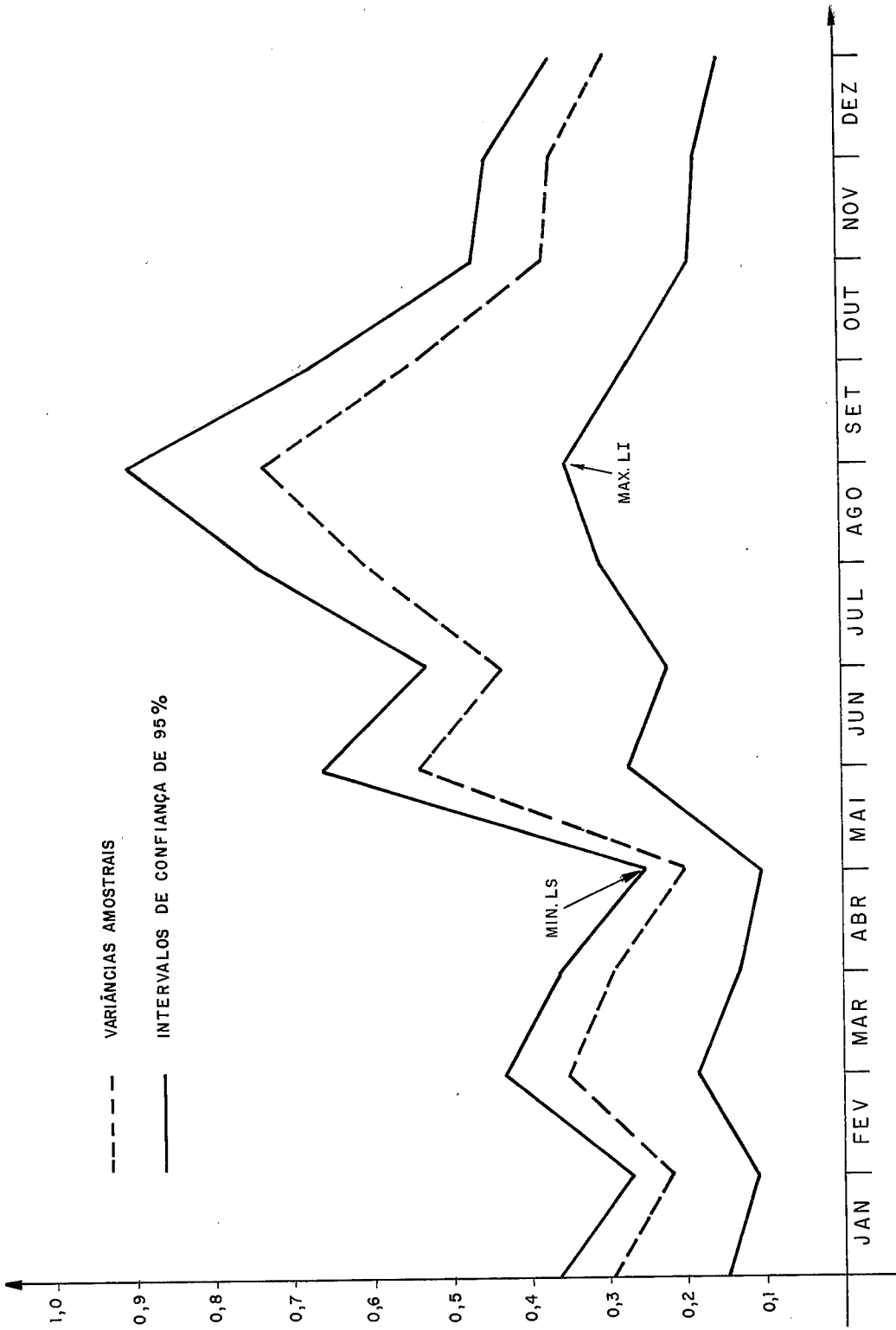


FIGURA 4.5 - TESTE DE CONSTÂNCIA PARA AS VARIÂNCIAS MENSAS DA SÉRIE DE VAZÕES MENSAS DE SALTO OSÓRIO (SÉRIE TRANSFORMADA)

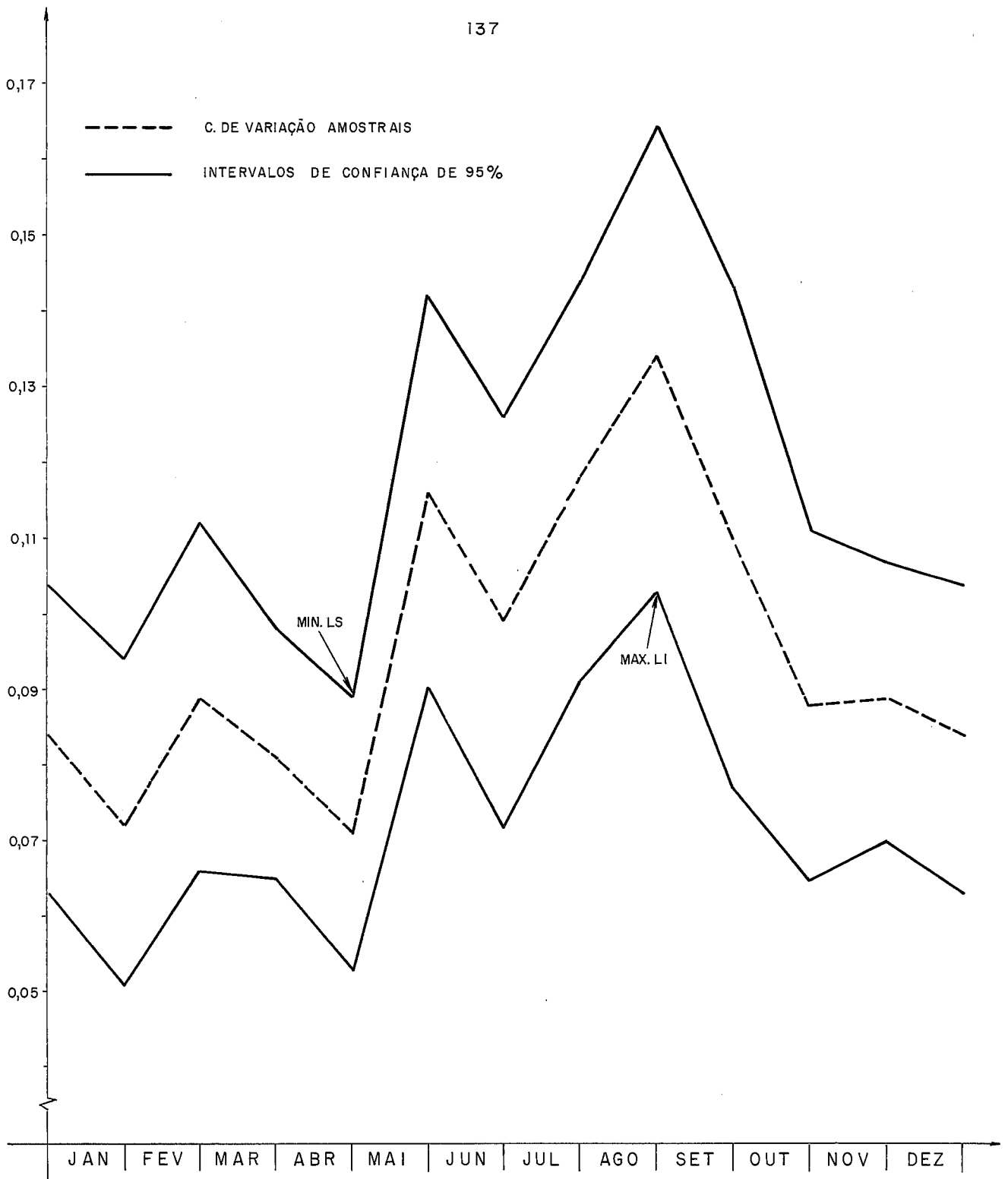


FIGURA 4.6 - TESTE DE CONSTÂNCIA PARA OS COEFICIENTES DE VARIAÇÃO MENSIS DA SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO (SÉRIE TRANSFORMADA)

$$\begin{aligned}\bar{X}_\tau = & 6,599 - 0,165 \cos\left(\frac{\pi}{2} \cdot \tau\right) - 0,095 \sin\left(\frac{\pi}{2} \cdot \tau\right) + \\ & + 0,092 \cos\left(\frac{\pi}{6} \cdot \tau\right) - 0,098 \sin\left(\frac{\pi}{6} \cdot \tau\right) - 0,066 \\ & \cos\left(\frac{\pi}{3} \cdot \tau\right) - 0,062 \sin\left(\frac{\pi}{3} \cdot \tau\right).\end{aligned}$$

$$\begin{aligned}S_\tau^2 = & 0,413 - 0,121 \cos\left(\frac{\pi}{6} \cdot \tau\right) - 0,139 \sin\left(\frac{\pi}{6} \cdot \tau\right) + \\ & - 0,022 \cos\left(\frac{2\pi}{3} \cdot \tau\right) - 0,086 \sin\left(\frac{2\pi}{3} \cdot \tau\right) + \\ & - 0,013 \cos\left(\frac{\pi}{3} \cdot \tau\right) + 0,061 \sin\left(\frac{\pi}{3} \cdot \tau\right) + 0,040 \\ & \cos\left(\frac{5\pi}{6} \cdot \tau\right) + 0,022 \sin\left(\frac{5\pi}{6} \cdot \tau\right)\end{aligned}$$

As médias e variâncias mensais observadas e ajustadas e os respectivos intervalos de confiança constam das figuras 4.7 e 4.8.

A função periódica ajustada aos desvios padrão mensais foi calculada como sendo a raiz quadrada da variância ajustada. Neste caso, porque a função periódica ajustada às variâncias mensais foi a mesma que a ajustada às variâncias mensais do modelo A-2, são válidos os resultados apresentados na tabela 4.6.

As estimativas da média e desvio padrão da componente estocástica dependente $\varepsilon_{p,\tau}$, estimada a partir da equação (2.18) foram, respectivamente, 0,004 e 1,002. Então, a componente $\zeta_{p,\tau}$ foi estimada como: $\zeta_{p,\tau} = \frac{\varepsilon_{p,\tau} - 0,004}{1,002}$.

Os coeficientes de correlação mensais da com-

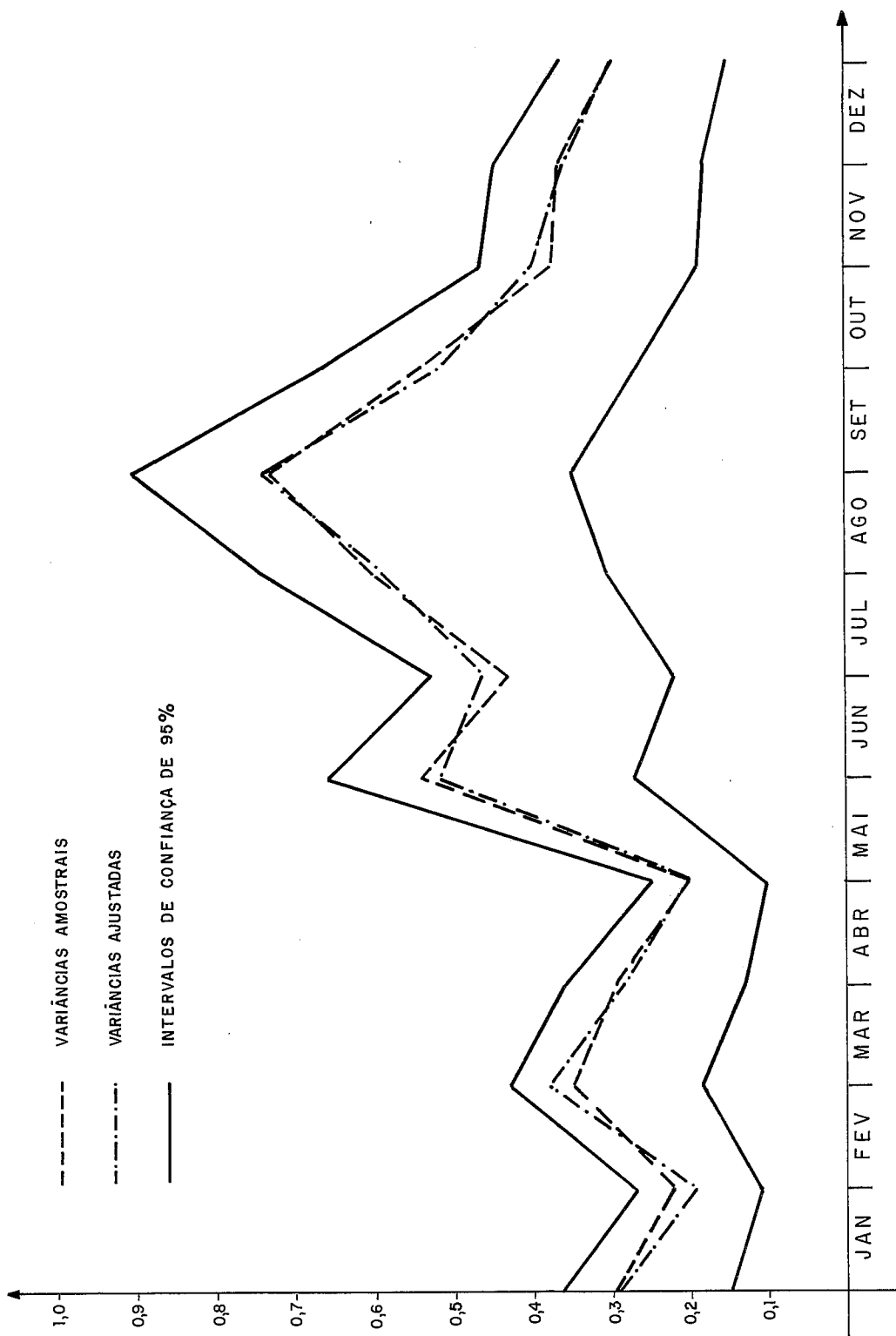


FIGURA 4.8 - TESTE PARA SIGNIFICÂNCIA DE HARMÔNICOS NA ANÁLISE HARMÔNICA NAS VARIÂNCIAS MENSAS DA SÉRIE DE VAZÕES MENSAS DE SALTO OSÓRIO (SÉRIE TRANSFORMADA)

ponente $\varepsilon_{p,\tau}$, foram estimados a partir da equação (3.36). O teste de constância, descrito em III.7, indicou que estes coeficientes não eram significativamente diferentes de uma constante sendo, então, estimados por 0,5943. A figura 4.9 apresenta o referido teste.

3) IDENTIFICAÇÃO DO (S) MODELO (S) PARA A COMPONENTE ESTOCÁSTICA DEPENDENTE PADRONIZADA

Com as funções de autocorrelação e autocorrelação parcial das séries $\{\zeta_t\}$ e $\{\omega_t = \nabla_{12} \zeta_t\}$, estimadas pelas equações (3.37) e (3.60) identificou-se, inicialmente, os modelos:

$$C-1 : (1,0,0) \times (0,0,0)_{12}$$

$$C-2 : (1,0,0) \times (1,0,0)_{12}$$

$$C-3 : (1,0,0) \times (0,0,1)_{12}$$

$$C-4 : (1,0,0) \times (0,1,1)_{12}$$

De acordo com a equação (3.34) estes modelos podem, também, ser escritos por,

$$C-1 : (1 - \phi_1 B) \zeta_t = a_t$$

$$C-2 : (1 - \phi_1 B) (1 - \phi_1 B^{12}) \zeta_t = a_t$$

$$C-3 : (1 - \phi_1 B) \zeta_t = (1 - \theta_1 B^{12}) a_t$$

$$C-4 : (1 - \phi_1 B) \nabla_{12} \zeta_t = (1 - \theta_1 B^{12}) a_t$$

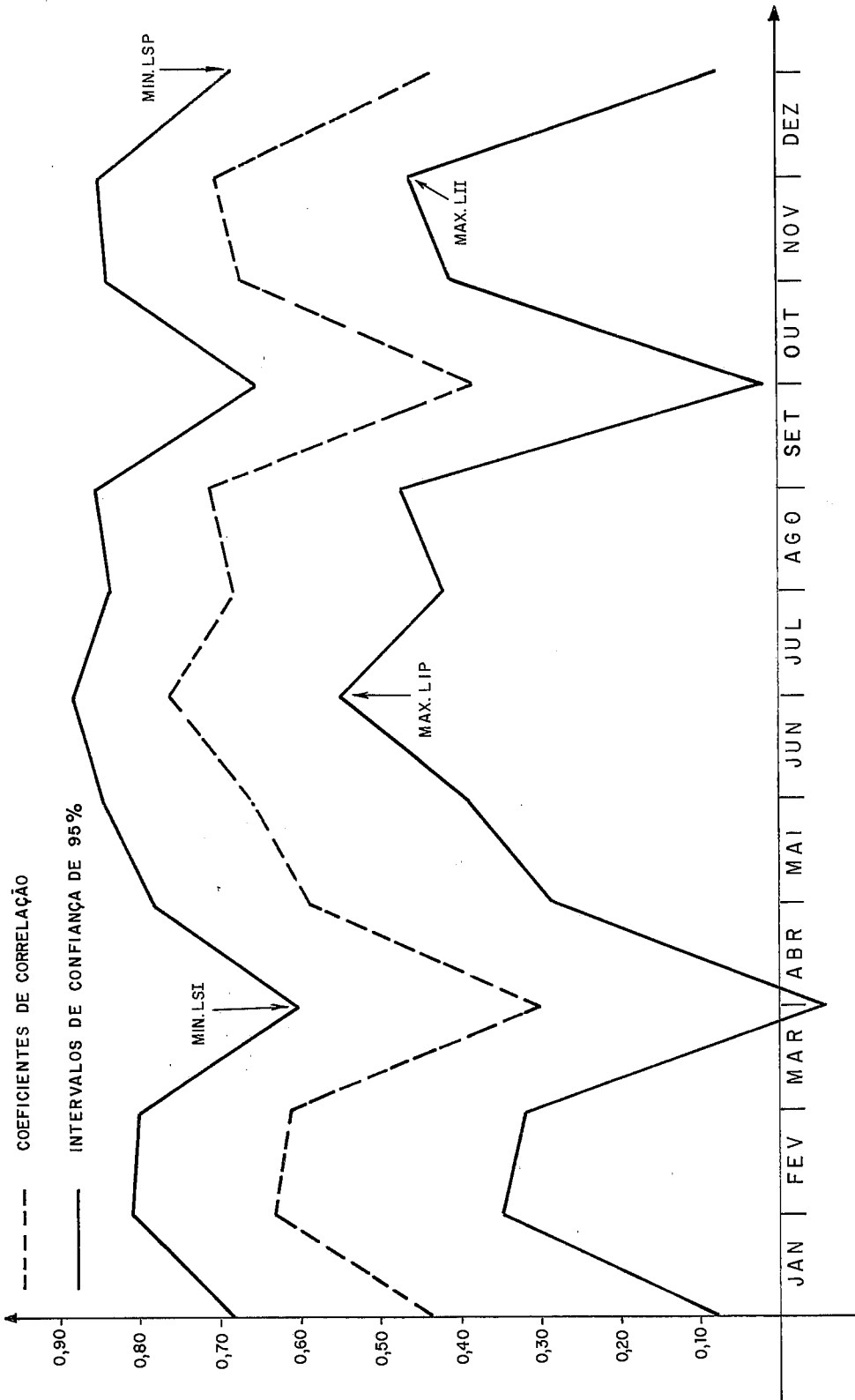


FIGURA 4.9 - TESTE DE CONSTÂNCIA PARA OS COEFICIENTES DE CORRELAÇÃO MENSIAIS DA SÉRIE DE VAZÕES MENSIAIS DE SALTO OSÓRIO (SÉRIE TRANSFORMADA)

4) ESTIMAÇÃO DOS PARÂMETROS DOS MODELOS C-1, C-2, C-3 e C-4

Estimativas aproximadas de máxima verossimilhança, obtidas a partir da técnica "back forecasting", descrita na seção III.10, para os quatro modelos constam da tabela 4.22.

	COEFICIENTES	ERRO PADRÃO DA ESTIMATIVA	VARIÂNCIA DOS RESÍDUOS
C-1	$\hat{\phi}_1 = 0,5922$	0,0425	0,651451
C-2	$\hat{\phi}_1 = 0,5897$	0,0426	0,648921
	$\hat{\phi}_1 = 0,0645$	0,0534	
C-3	$\hat{\phi}_1 = 0,5889$	0,0426	0,647987
	$\hat{\theta}_1 = -0,0886$	0,0533	
C-4	$\hat{\phi}_1 = 0,5991$	0,0430	0,780761
	$\hat{\theta}_1 = 0,8675$	0,0279	

TABELA 4.22 - PARÂMETROS DOS MODELOS C-1, C-2, C-3 e C-4 PARA A SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO.

5) TESTES DE VERIFICAÇÃO APLICADOS AOS MODELOS C-1, C-2, C-3 e C-4

Segundo o descrito na seção III.11, os testes aplicados aos modelos C-1, C-2, C-3 e C-4 foram:

(i) correlograma dos resíduos - as autocorrelações de ordem 24, 26 e 31 das estimativas dos resíduos, para os modelos C-1, C-2 e C-3, excederam os limites do intervalo de 95% de confiança. Com relação ao modelo C-4, a única autocorrelação que não pertenceu ao intervalo de 95% de confiança foi a de "lag" 26. Como o número de autocorrelações, testadas para o primeiro modelo foi de 41 enquanto que para os demais foi de 42, conclui-se que apenas o modelo C-4 é significante a 95%. Os modelos C-1, C-2 e C-3 são, entretanto, significantes a 99%;

(ii) periodograma acumulado - os periodogramas acumulados para os quatro modelos, estimados a partir da equação (3.83), mostraram que os resíduos destes modelos poderiam ser considerados como estimativa de um ruído branco;

(iii) teste de portmanteau - estimativas da estatística Q, obtidas a partir da equação (3.82), indicaram que todos os quatro modelos eram significantes a 95% de confiança. Os resultados deste constam da tabela 4.23.

		MOD.C-1	MOD.C-2	MOD.C-3	MOD.C-4	QUI-QUADRADO TAB (95%,40)
Q _{OBS}	PARA					
40	G L	48,683	46,845	45,947	40,470	55,8

TABELA 4.23 - TESTE DE PORTMANTEAU PARA OS MODELOS C-1, C-2, C-3 e C-4 DA SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO

(iv) sobreajustamento - de acordo com os testes anteriores e considerando a hipótese de parcimônia selecionou-se, para sobreajustar os modelos C-1 e C-4, os seguintes modelos:

$$C-5 : (2,0,0) \times (0,0,0)_{12}$$

$$C-6 : (1,0,0) \times (0,1,2)_{12}$$

$$C-7 : (1,0,0) \times (1,1,1)_{12}$$

6) ESTIMAÇÃO DOS PARÂMETROS DOS MODELOS C-5, C-6 e C-7

As estimativas aproximadas de máxima verossimilhança e seus respectivos erros padrão para os parâmetros dos modelos C-5, C-6 e C-7, encontram-se na tabela 4.24.

	COEFICIENTES	ERRO PADRÃO DA ESTIMATIVA	VARIÂNCIA DOS RESÍDUOS
C-5	$\hat{\phi}_1 = 0,5930$ $\hat{\phi}_2 = -0,0018$	0,0529 0,0530	0,651447
C-6	$\hat{\phi}_1 = 0,5918$ $\hat{\theta}_1 = 0,7501$ $\hat{\theta}_2 = 0,1316$	0,0433 0,0538 0,0544	0,769959
C-7	$\hat{\phi}_1 = 0,5918$ $\hat{\phi}_1 = 0,1220$ $\hat{\theta}_1 = 0,8956$	0,0433 0,0628 0,0292	0,772059

TABELA 4.24 - PARÂMETROS DOS MODELOS C-5, C-6 e C-7 PARA A SÉRIE DE VAZÕES MENSIS DE SALTO OSÓRIO.

7) TESTES DE VERIFICAÇÃO PARA OS MODELOS C-5, C-6 e C-7

O primeiro teste a ser aplicado, visou verificar a hipótese de que os parâmetros $\hat{\phi}_2$ e $\hat{\phi}_1$ dos modelos C-5 e C-7, fossem iguais a zero. Para ambos os modelos, as hipóteses nulas de que $\phi_2=0$ e $\phi_1=0$ foram consideradas consistentes a um nível de significância de 5%. Prosseguiu-se, então, com os demais testes, apenas para o modelo C-6, visto que os modelos C-5 e C-7 recaíram nos C-1 e C-4, analisados anteriormente.

O teste do correlograma dos resíduos apresentou as autocorrelações de "lag 18" e "lag 26" não pertencentes ao intervalo de 95% de confiança. Como o número de autocorrelações testadas foi de 43, aceita-se que este resultado corresponda ao nível de significância do teste.

O gráfico do periodograma acumulado mostrou um comportamento linear entre os pontos (0;0) e (0,5;1), indicando que os resíduos do modelo C-6 representavam estimativas do ruído branco.

Com relação ao teste de portmanteau, a hipótese nula da adequacidade do modelo C-6 foi aceita a um nível de significância de 5%. A estimativa da estatística Q, para este caso, foi de 38,661 enquanto que o valor crítico tabelado a 95% e 40 graus de liberdade é de 55,8.

Em vista do critério de parcimônia adotado neste trabalho, além de não se prosseguir com o teste de sobreajustamento para o modelo C-6, considerou-se que os

resultados dos testes aplicados a este modelo não justificavam sua escolha em detrimento do modelo C-4. Por estes motivos levou-se à fase de previsão os modelos C-1 e C-4, os quais, de acordo com a equação (3.34), podem ser escritos como:

MODELO C-1:

$$(1-0,5922 B) \zeta_t = a_t$$

MODELO C-4:

$$(1-0,5889 B) \nabla_{12} \zeta_t = (1-0,8675 B^{12}) a_t$$

8) PREVISÃO

As previsões e intervalos de confiança de 50%, 90%, 95%, como explicadas na seção III.12, foram obtidas para os anos de 1971, 1968, 1970 e 1957 através da integração dos modelos C-1 e C-4 com a componente determinística resultante da análise estrutural. Os resultados constam, respectivamente, das tabelas 4.25, 4.26, 4.27 e 4.28.

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4						
JAN	3323	839	974	658	1069	746	1271	464	1516	509	1862	414	1698	450	2109
FEV	1212	1030	1256	697	1520	818	1926	398	2660	443	3561	332	3192	363	4347
MAR	1166	883	959	619	1259	648	1417	371	2078	370	2484	315	2475	308	2980
ABR	1238	633	569	468	858	408	795	303	1326	252	1284	263	1527	216	1501
MAI	2006	636	449	391	1036	263	769	194	2086	122	1662	155	2620	95	2134
JUN	2537	765	734	483	1210	443	1218	250	2341	214	517	202	2899	169	3186
JUL	1818	736	906	437	1240	510	1610	206	2622	224	3676	162	3346	171	4804
AGO	822	645	437	361	1153	230	830	157	2660	92	2082	119	3487	68	2808
SET	754	795	777	487	1296	453	1331	241	2617	209	2889	192	3289	162	3714
OUT	1003	1084	1213	705	1665	756	1947	380	3086	383	3845	311	3772	307	4793
NOV	437	941	1056	627	1412	676	1652	350	2531	355	3141	290	3059	288	3869
DEZ	398	641	1208	443	928	804	1816	260	1580	448	3261	219	1877	370	3945

TABELA 4.25 - PREVISÕES DE VAZÕES MENSIAIS PARA 1971 EM SALTO OSÓRIO
METODOLOGIA YEJJEVICH/BOX E JENKINS

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4						
JAN	616	377	269	296	481	206	351	209	682	141	514	186	764	124	582
FEV	548	536	298	363	791	195	458	207	1385	105	846	173	1661	86	1033
MAR	327	630	293	441	898	199	434	265	1496	113	761	224	1766	94	912
ABR	387	535	410	395	725	294	572	256	1121	182	925	222	1291	155	1080
MAI	320	543	336	334	883	197	575	166	1779	91	1242	132	2233	71	1596
JUN	244	700	904	442	1108	545	1498	229	2143	264	3097	185	2654	208	3923
JUL	284	693	648	412	1168	365	1151	195	2471	160	2630	152	3150	122	3437
AGO	217	620	435	347	1109	229	825	150	2557	91	2069	115	3352	68	2791
SET	224	779	1096	478	1270	639	1879	237	2566	295	4077	188	3225	229	5240
OUT	407	1073	1491	698	1649	929	2394	377	3055	471	4723	308	3734	377	5892
NOV	744	936	1129	624	1405	722	1766	348	2517	380	3336	288	3041	308	4135
DEZ	542	639	705	441	925	469	1060	260	1574	261	1903	218	1871	216	2303

TABELA 4.26 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS PARA 1968 (ANO SECO) EM SALTO OSÓRIO

METODOLOGIA YEVEVICH/BOX E JENKINS

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%		INT. DE 95%							
		MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4						
JAN	736	693	816	543	883	625	1064	383	1252	427	1560	342	1402	377	1766
FEV	625	881	1169	597	1300	763	1793	341	2276	412	3316	284	2730	388	4050
MAR	517	814	963	571	1161	652	1423	343	1934	372	2496	290	2283	310	2995
ABR	456	608	565	449	824	404	789	291	1274	250	1274	252	1467	214	1489
MAI	651	613	410	376	997	240	701	187	2008	111	1515	149	2521	86	1947
JUN	1575	749	643	473	1185	388	1067	245	2292	188	2206	198	2839	148	2793
JUL	2121	726	792	431	1222	446	1407	203	2586	195	3212	160	3297	149	4198
AGO	541	639	424	357	1142	224	805	155	2635	89	2020	118	3454	66	2722
SET	693	791	795	485	1290	464	1364	240	2605	214	2957	191	3272	166	3803
OUT	1303	1081	1209	704	1661	753	1941	380	3078	381	3830	311	3762	306	4778
NOV	561	940	1173	627	1411	751	1835	350	2528	395	3489	289	3054	302	4297
DEZ	1147	641	1228	443	928	817	1846	260	1578	455	3315	219	1876	376	4010

TABELA 4.27 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS PARA 1970 (ANO MÉDIO) EM SALTO OSÓRIO

METODOLOGIA YEVJEVICH/BOX E JENKINS

MÊS	OBSERVADO	PREVISÃO		INT. DE 50%		INT. DE 90%			INT. DE 95%						
		MOD.C-1	MOD.C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4	MOD. C-1	MOD. C-4				
JAN	432	777	882	610	991	676	1151	430	1405	461	1686	384	1574	407	1909
FEV	1115	968	1537	655	1428	1002	2357	374	2500	542	4357	312	2999	444	5322
MAR	534	855	1241	599	1219	840	1834	360	2031	479	3217	305	2398	399	3859
ABR	433	623	510	460	844	365	712	298	1305	226	1151	258	1503	193	1345
MAI	421	627	378	385	1020	221	647	191	2055	102	1399	152	2580	80	1797
JUN	1090	758	627	479	1200	378	1039	248	2322	183	2149	200	2876	144	2722
JUL	3262	732	858	434	1233	483	1524	205	2608	212	3480	161	3325	162	4548
AGO	4796	643	463	359	1149	244	878	156	2648	97	2203	119	3475	72	2972
SET	5206	795	1059	486	1293	618	1816	241	2613	285	3941	192	3282	222	5065
OUT	1887	1083	1437	705	1663	895	2307	380	3084	454	4555	311	3767	364	5679
NOV	1498	941	1258	627	1412	804	1967	350	2529	423	3739	290	3058	343	4607
DEZ	823	641	1541	443	928	1205	2317	260	1579	571	4160	219	1877	472	5032

TABELA 4.28 - PREVISÕES DE VAZÕES MÉDIAS MENSIS PARA 1957 (ANO ÚMIDO) EM SALTO OSÓRIO

METODOLOGIA YEVJEVICH/BOX E JENKINS

IV.5 - AVALIAÇÃO DAS PREVISÕES PELOS DIFERENTES MODELOS

Visando avaliar as previsões feitas pelos modelos apresentados, pode-se utilizar diversos critérios. Neste trabalho foram utilizados os seguintes: (1) soma dos quadrados dos afastamentos das previsões em relação aos valores observados; (2) desvio máximo, em valor absoluto, das previsões em relação às observações; e (3) soma dos quadrados dos afastamentos, dos limites superiores dos intervalos de 95% das previsões, em relação aos dados observados, doravante chamados de critérios 1, 2 e 3. Estes critérios foram utilizados de duas formas distintas. Em primeiro lugar, visou-se uma avaliação das previsões doze meses à frente de uma origem pré-fixada. Estas previsões corresponderam aos anos de 1971 (não pertencente ao histórico), 1968 (ano seco), 1970 (ano médio) e 1957 (ano úmido) sendo, então, as origens para as previsões tomadas em dezembro dos anos imediatamente anteriores. Em segundo lugar, objetivou-se avaliar os desempenhos dos modelos a medida que se passava a conhecer os primeiros 3, 6 e 9 meses de 1971. Neste caso, as origens corresponderam, respectivamente, aos meses de março, junho e setembro de 1971.

1) AVALIAÇÃO DAS PREVISÕES DOZE MESES À FRENTE

Os resultados da aplicação dos critérios 1, 2 e 3 aos diferentes modelos constam da tabela 4.29. Nela, nota-se que em relação ao ano de 1971, o critério

1 indicou como melhores os modelos A-2 e C-1, com o critério 2, nenhuma conclusão pode ser tirada e, com o critério 3, sobressairam os modelos A-1, A-2 e C-1. Para o ano de 1968, enquanto o critério 1 indicava os modelos A-1 e C-1, os critérios 2 e 3 indicavam os modelos A-1, A-2 e C-1. Com respeito aos anos de 1970 e 1957, o critério 1 apontou, respectivamente, os modelos B-5 e B-6. Houve uma pequena predominância do modelo B-5, para 1970, em relação ao critério 2, enquanto que para 1957, nenhum modelo predominou; com o critério 3, os modelos A-1, A-2 e C-1 sobressairam sobre os demais.

2) AVALIAÇÃO DAS PREVISÕES PARA 1971 COM 3, 6 e 9 MESES CONHECIDOS

As previsões e intervalos de confiança de 95% para 1971, tendo como origem respectivamente os meses de março, junho e setembro do mesmo ano, obtidas pelos seis modelos, estão representadas nas figuras 4.10 a 4.27.

A aplicação dos critérios 1, 2 e 3, neste caso, apresentou os resultados constantes da tabela 4.30. Estes resultados mostram uma nítida predominância dos modelos A-1, A-2 e C-1 sobre os demais.

3) DISCUSSÃO DOS RESULTADOS

Como se pode observar pelas análises anteriores os modelos A-1, A-2 e C-1 foram os que apresentaram melhores desempenhos. No entanto, ao se verificar com

	CRITÉRIO 1				CRITÉRIO 2				CRITÉRIO 3			
	Σ (obs. - prev.) ²				MAX obs. - prev.				Σ (obs. - LS (95%)) ²			
	1971	1968	1970	1957	1971	1968	1970	1957	1971	1968	1970	1957
MOD.A-1	13 583 427	1 681 812	3 380 274	46 298 464	2 474	634	1 440	4 477	31 238 592	54 643 728	34 710 336	20 787 136
MOD.A-2	13 166 442	2 087 429	3 178 856	45 303 168	2 478	691	1 387	4 414	41 458 976	69 513 040	43 422 976	22 546 112
MOD.B-5	13 533 495	4 364 303	2 932 301	43 535 280	2 424	969	1 132	4 271	68 964 320	132 325 152	95 205 152	37 258 672
MOD.B-6	13 778 516	5 080 309	3 361 780	40 264 848	2 368	1 087	1 290	4 224	75 404 464	147 042 848	90 426 608	40 892 016
MOD.C-1	13 189 738	1 556 187	3 276 612	44 527 424	2 484	666	1 395	4 411	39 285 536	58 775 824	40 960 336	29 853 776
MOD.C-4	13 749 064	2 910 704	3 620 113	43 629 888	2 349	1 084	1 329	4 333	75 445 264	102 634 416	76 135 072	83 085 984

TABELA 4.29 - AVALIAÇÃO DAS PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS PELOS MODELOS A-1, A-2, B-5, B-6, C-1 E C-4, PARA OS ANOS DE 1971, 1968, 1970 E 1957 EM SALTO OSÓRIO.

	CRITÉRIO 1				CRITÉRIO 2			CRITÉRIO 3		
	Σ (obs. - prev.) ²				MAX obs. - prev.			Σ (obs. - LS (95%)) ²		
	ABR-DEZ	JUL-DEZ	OUT-DEZ	ABR-DEZ	JUL-DEZ	OUT-DEZ	ABR-DEZ	JUL-DEZ	OUT-DEZ	
MOD.A-1	6 526 125	632 866	235 317	1 729	502	440	25 184 624	34 443 296	9 795 110	
MOD.A-2	6 139 284	595 370	262 778	1 654	535	445	33 723 952	78 495 456	13 291 158	
MOD.B-5	6 996 394	1 480 746	932 954	1 668	900	823	57 758 464	85 605 312	32 063 008	
MOD.B-6	7 691 901	1 653 392	1 119 226	1 679	987	903	60 185 632	85 694 896	35 669 664	
MOD.C-1	6 282 779	651 981	299 448	1 714	580	490	34 884 800	67 464 224	11 111 006	
MOD.C-4	7 774 819	1 585 802	1 055 380	1 763	867	804	62 410 048	115 142 784	27 901 376	

TABELA 4.30 - AVALIAÇÃO DAS PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS PELOS MODELOS A-1,

A-2, B-5, B-6, C-1 e C-4, PARA OS NOVE, SEIS E TRÊS ÚLTIMOS MESES DE 1971 EM

SALTO OSÓRIO.

mais cuidado os resultados pertinentes ao modelo A-1, nota-se que este modelo apresenta algum dos limites inferiores negativos. Sugere-se, então, os modelos A-2 e C-1, como os que apresentam uma melhor performance.

Deve-se notar, entretanto, que se o objetivo é a obtenção de previsões a curto prazo, uma maior precisão poderia ser alcançada com o relacionamento dos processos de afluência e precipitação. Por outro lado, é inegável a dificuldade na coleta e análise de dados para este propósito. Por este motivo, tanto o modelo A-2 quanto o C-1, devem ser encarados apenas como uma aproximação limitada do sistema físico real referente à séries hidrológicas temporais.

É importante que se observe, também, que além de não ser de todo correto uma avaliação das previsões para anos pertencentes à série histórica, visto que eles foram utilizados na determinação dos parâmetros dos modelos, esta avaliação ficou um tanto prejudicada por causa de mudanças bruscas de comportamento da própria série. Assim, o ano anterior ao ano seco selecionado (1968), foi úmido e, o anterior ao úmido selecionado (1957), foi seco. Ora, diante de situações adversas como estas, era de se esperar que nenhum dos modelos testados apresentassem resultados altamente satisfatórios. Um outro ponto importante é que o próprio ano de 1971 apresenta alguns valores divergentes em relação à série histórica, especialmente nos meses de janeiro, abril, maio e junho.

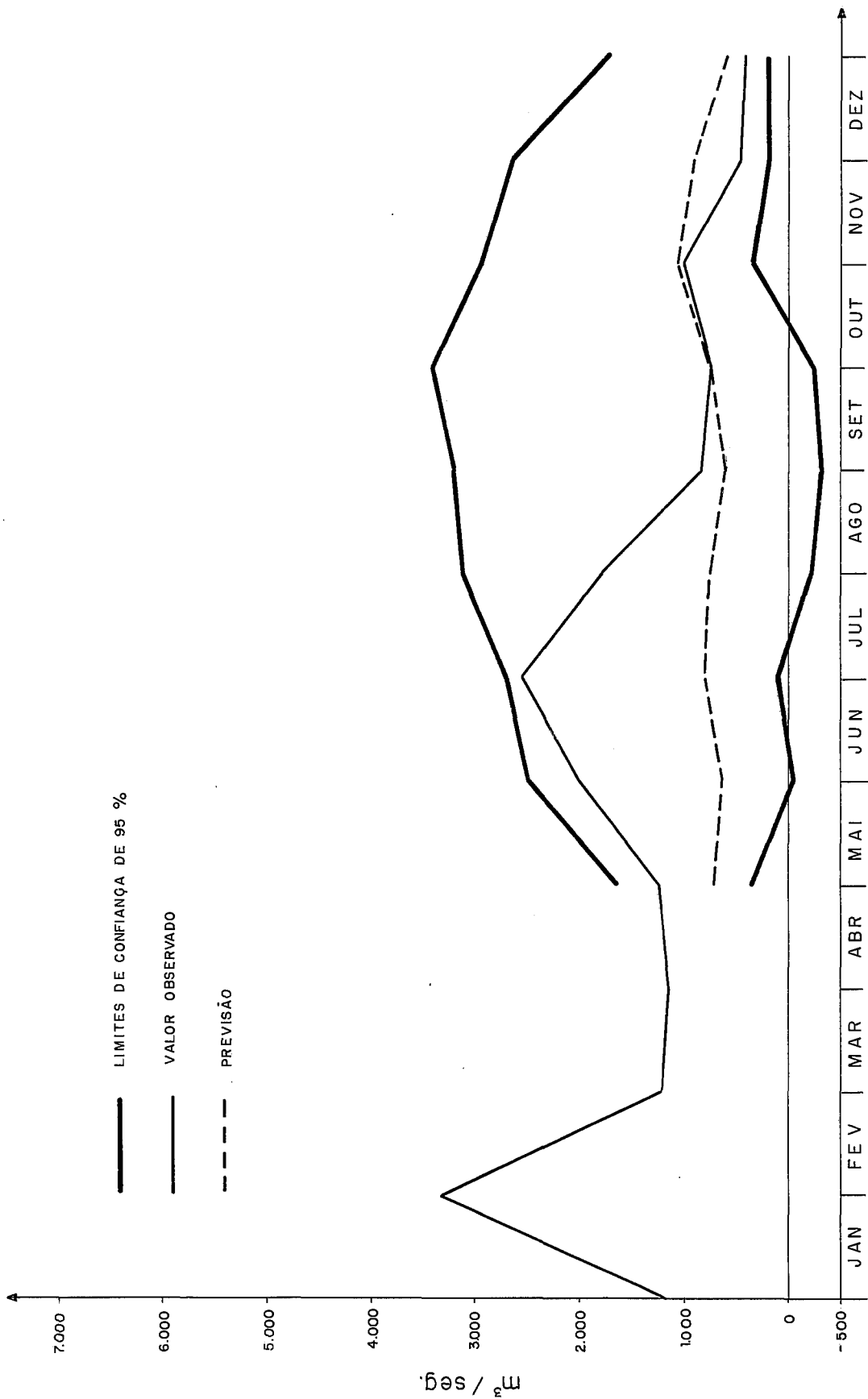


FIG. 4.10 - PREVISÕES DE VAZÕES MÉDIAS MENSIS, OBTIDAS A PARTIR DO MODELO A-1 PARA OS NOVE ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE YEVEVICH).

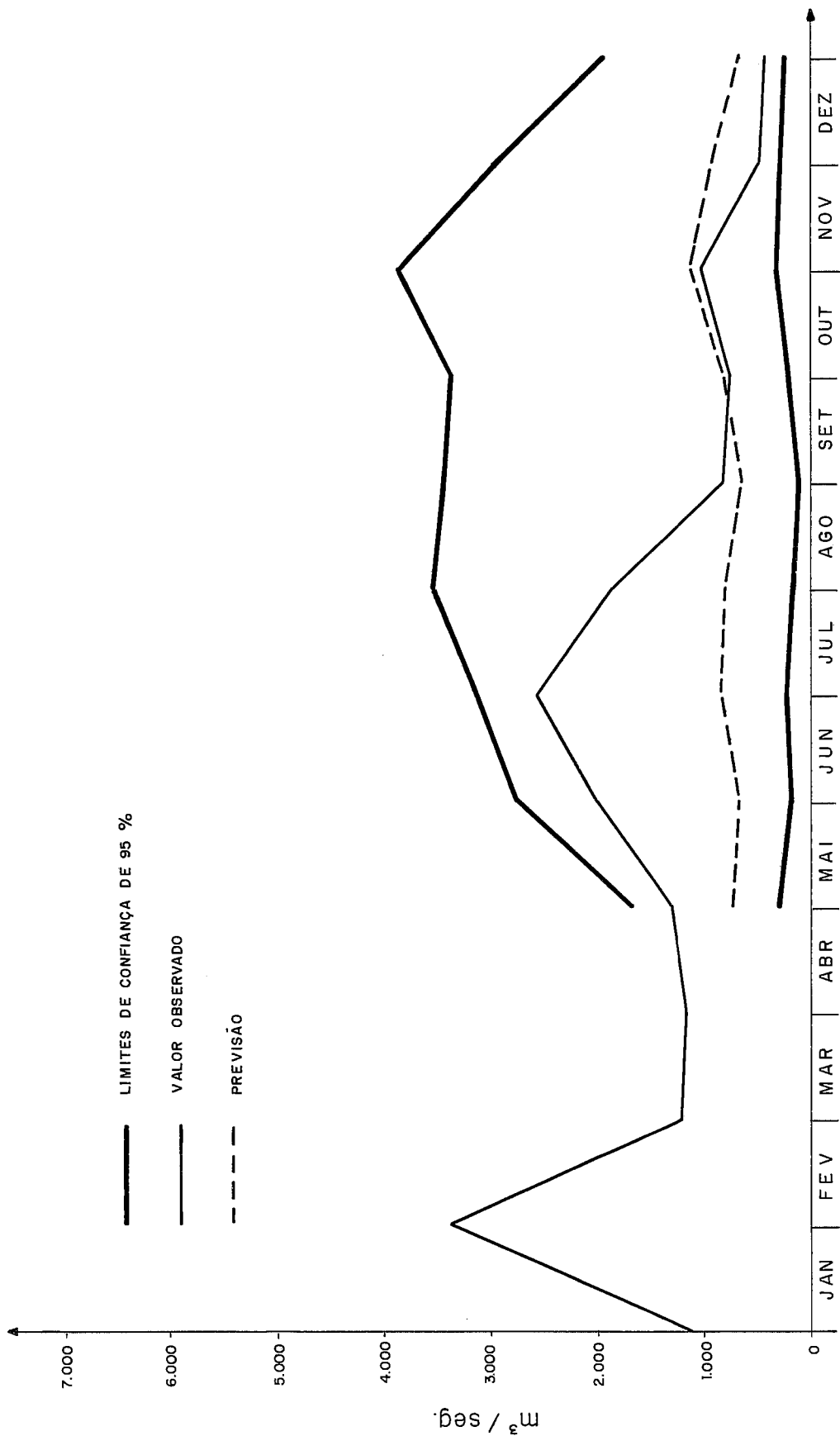


FIG. 4.11 - PREVISÕES DE VAZÕES MÉDIAS MENSIS, OBTIDAS A PARTIR DO MODELO A-2 PARA OS NOVE ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE YEVEVICH)

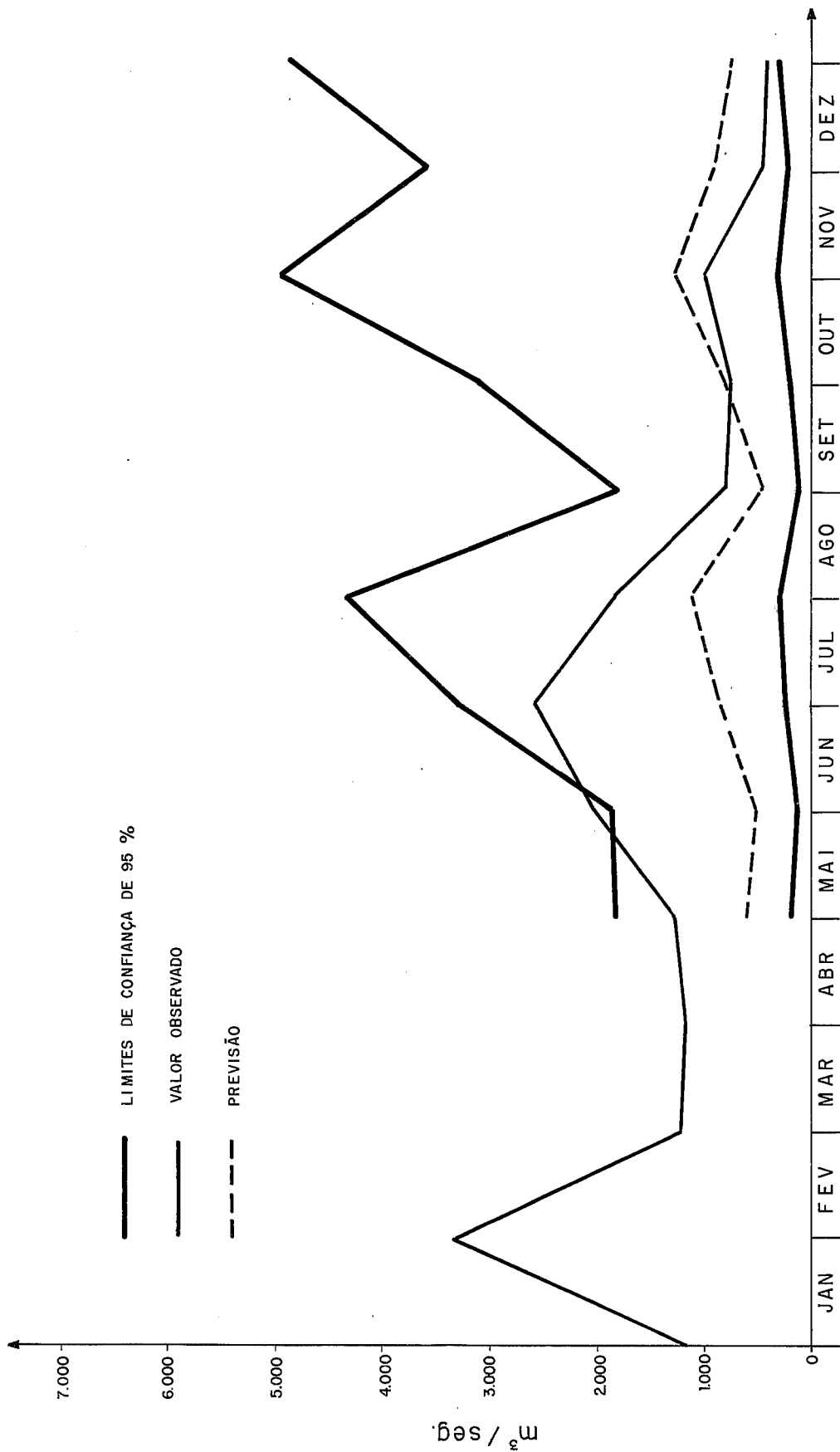


FIG. 4.12 - PREVISÕES DE VAZÕES MÉDIAS MENSIS, OBTIDAS A PARTIR DO MODELO B-5, PARA OS NOVE ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE BOX E JENKINS).

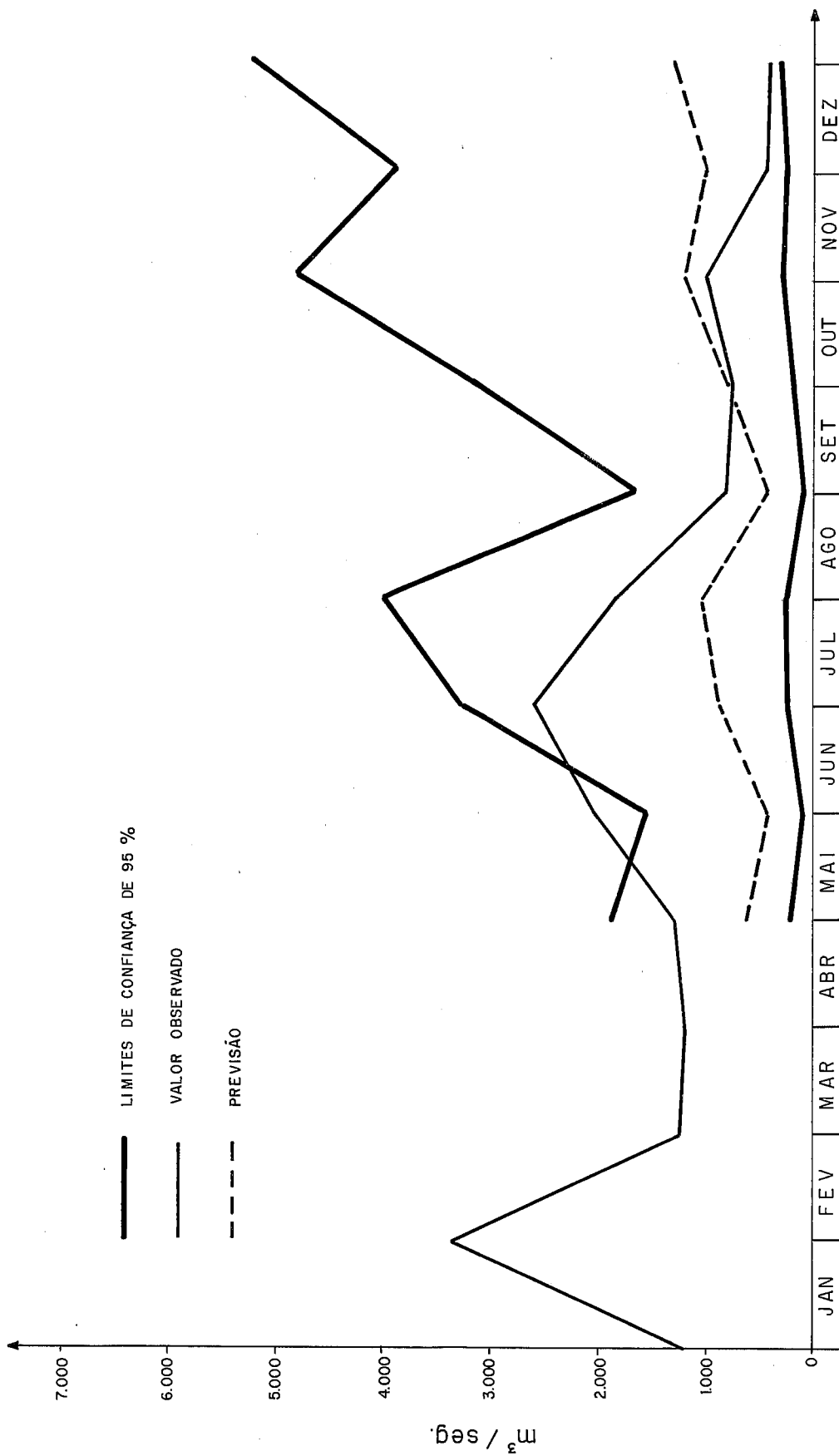


FIG. 4.13 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS A PARTIR DO MODELO B-6, PARA OS NOVE ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE BOX E JENKINS).

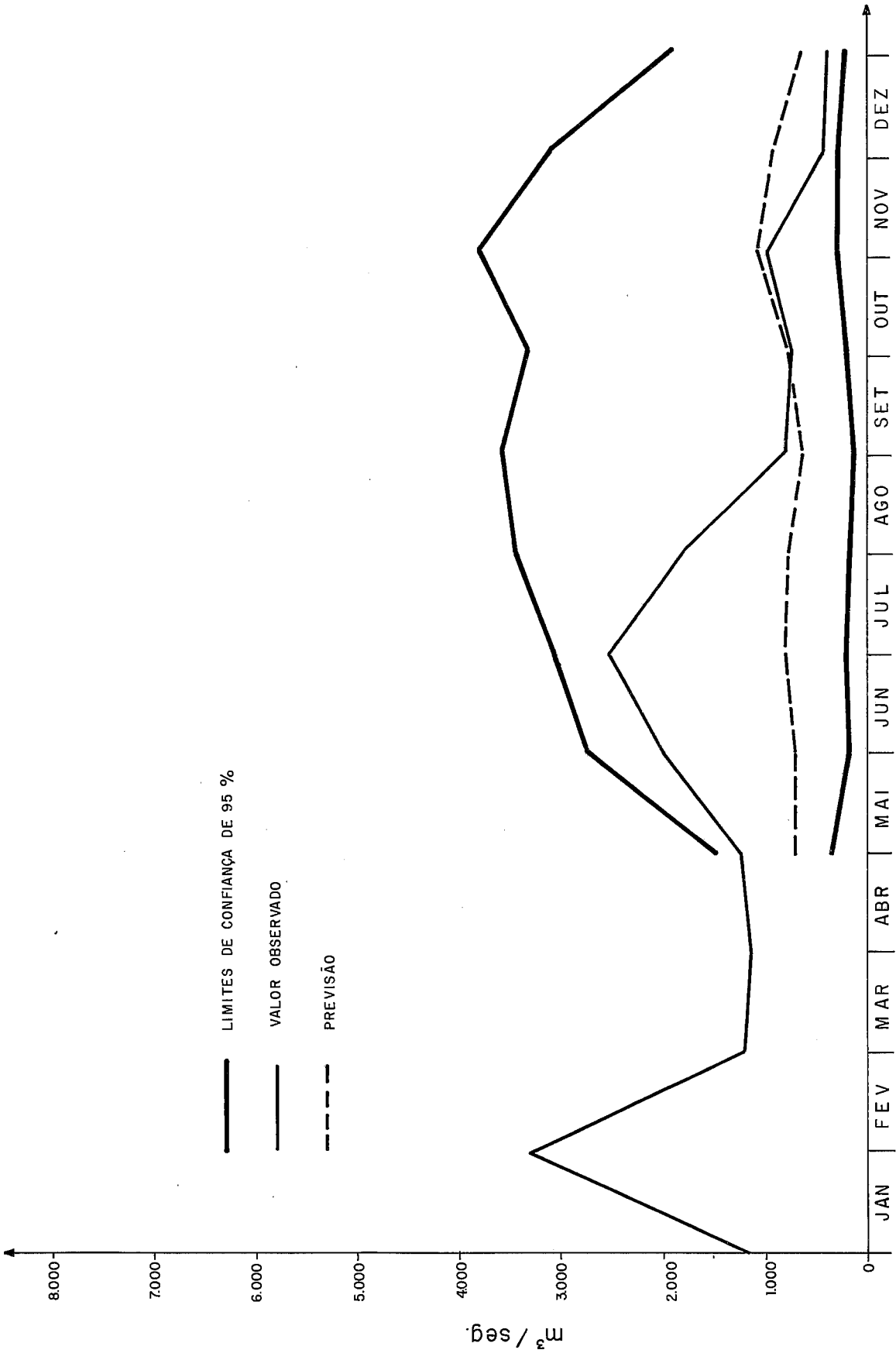


FIG. 4.14 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS A PARTIR DO MODELO C-1, PARA OS NOVE ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA YEVJEVICH / BOX E JENKINS).

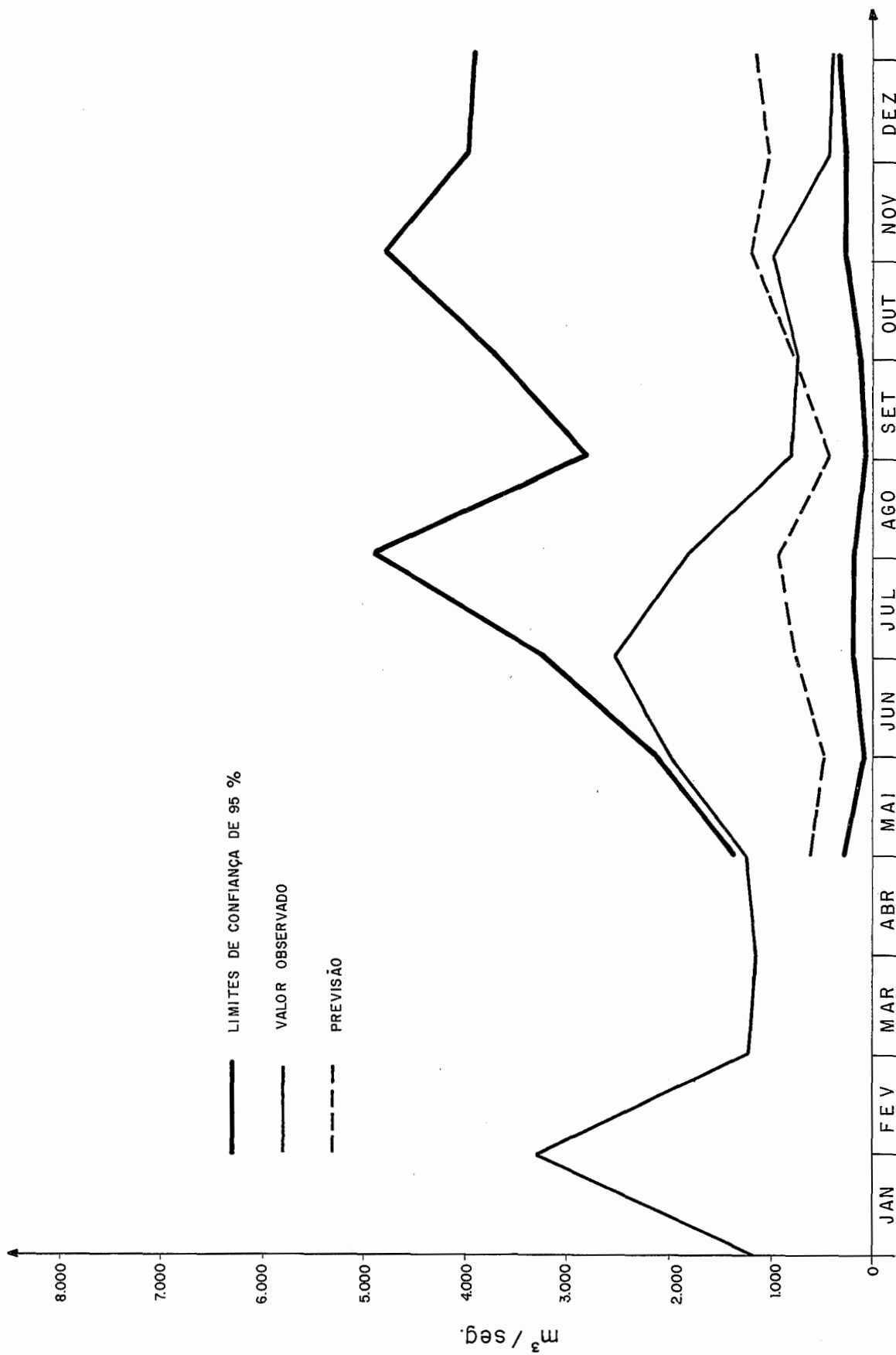


FIG. 4-15 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS A PARTIR DO MODELO C-4, PARA OS NOVE ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA YEVJEVICH / BOX E JENKINS).

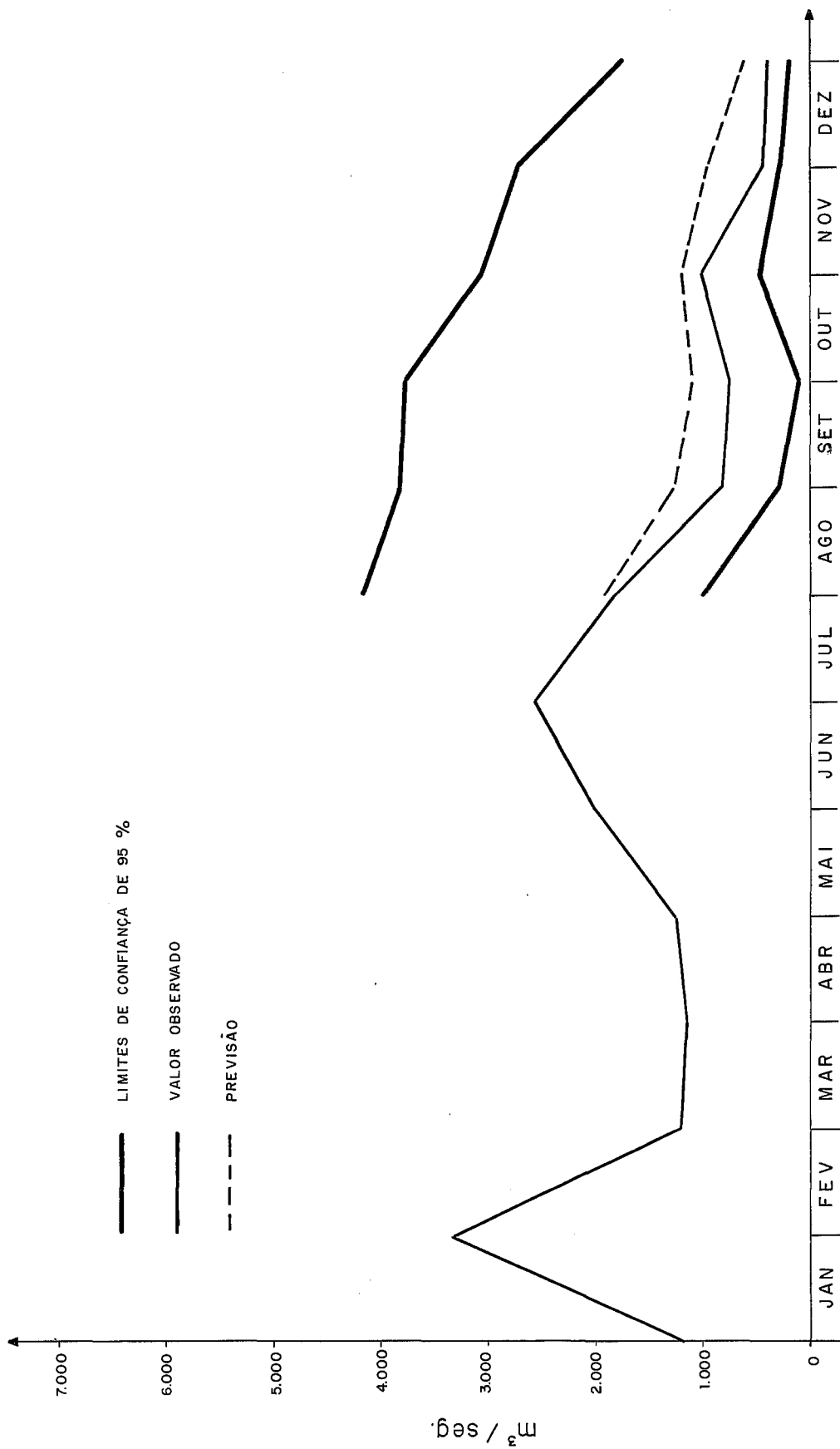


FIG. 4-16 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS A PARTIR DO MODELO A-1 PARA OS SEIS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE YEVEVICH).

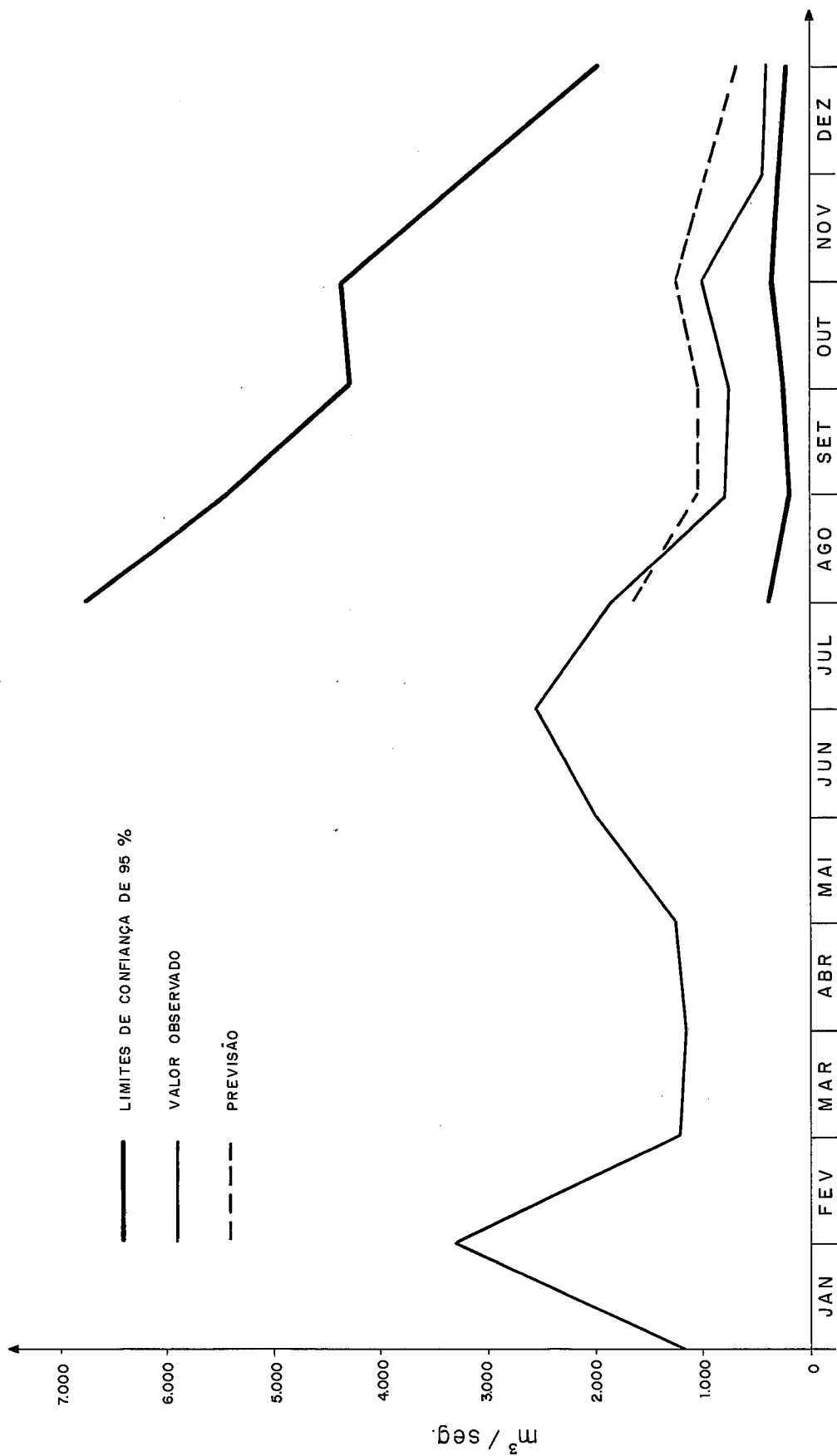


FIG. 4-17 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS A PARTIR DO MODELO A-2 PARA OS SEIS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE YEVJEVICH)

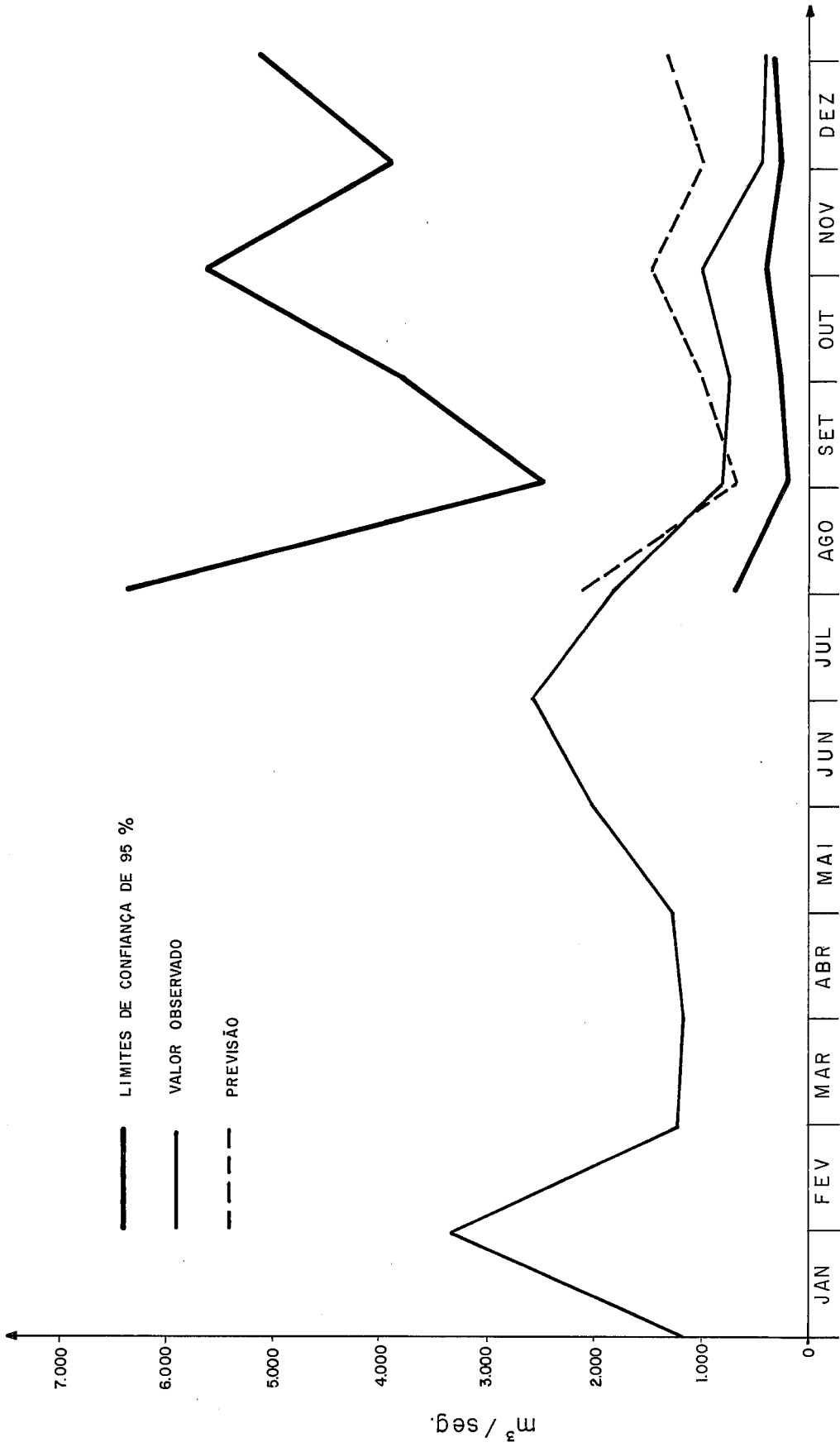


FIG. 4.18 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS A PARTIR DO MODELO B-5, PARA OS SEIS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE BOX E JENKINS).

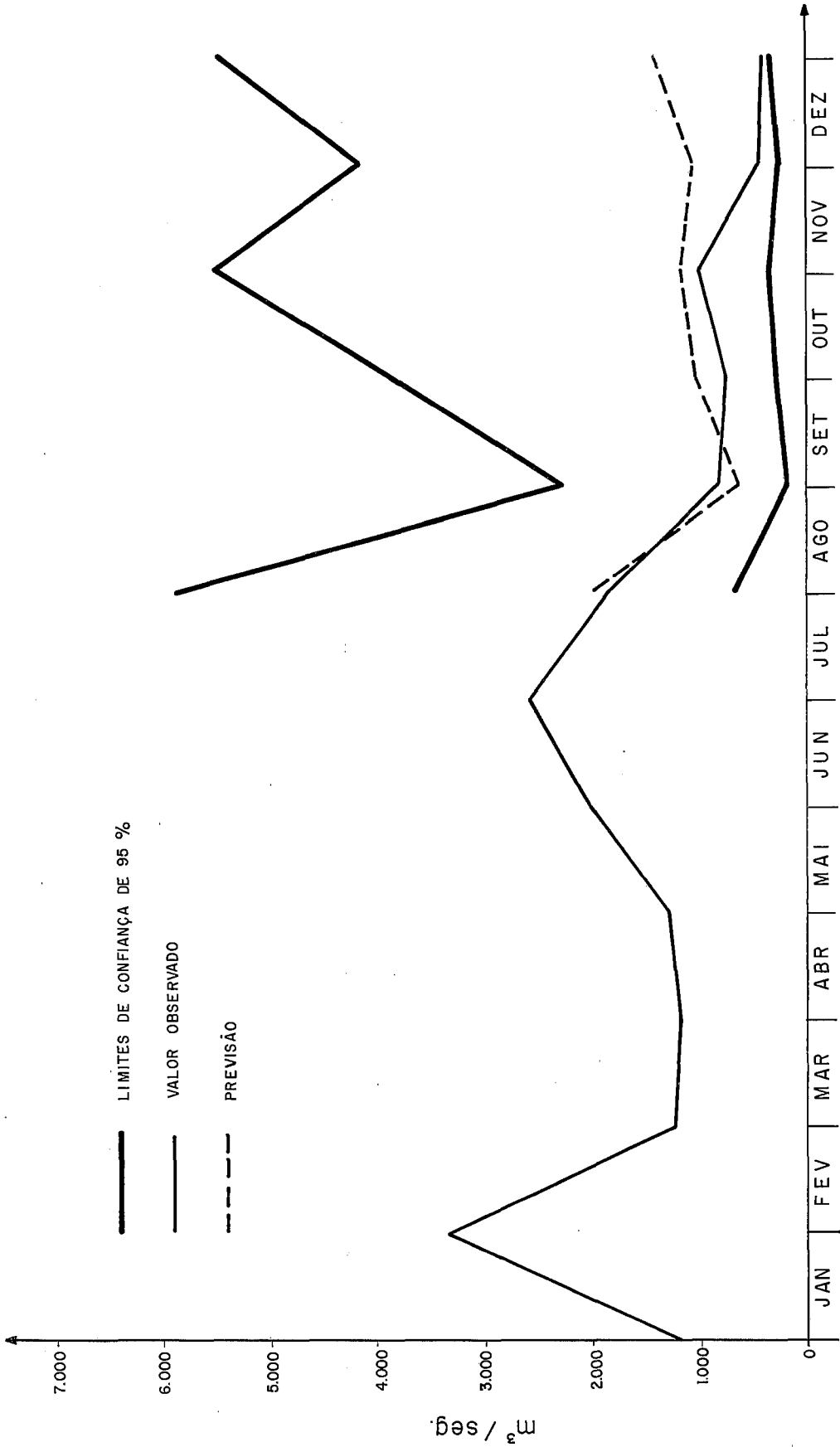


FIG. 4.19 - PREVISÕES DE VAZÕES MÉDIAS MENSIS, OBTIDAS A PARTIR DO MODELO B-6, PARA OS SEIS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE BOX E JENKINS).

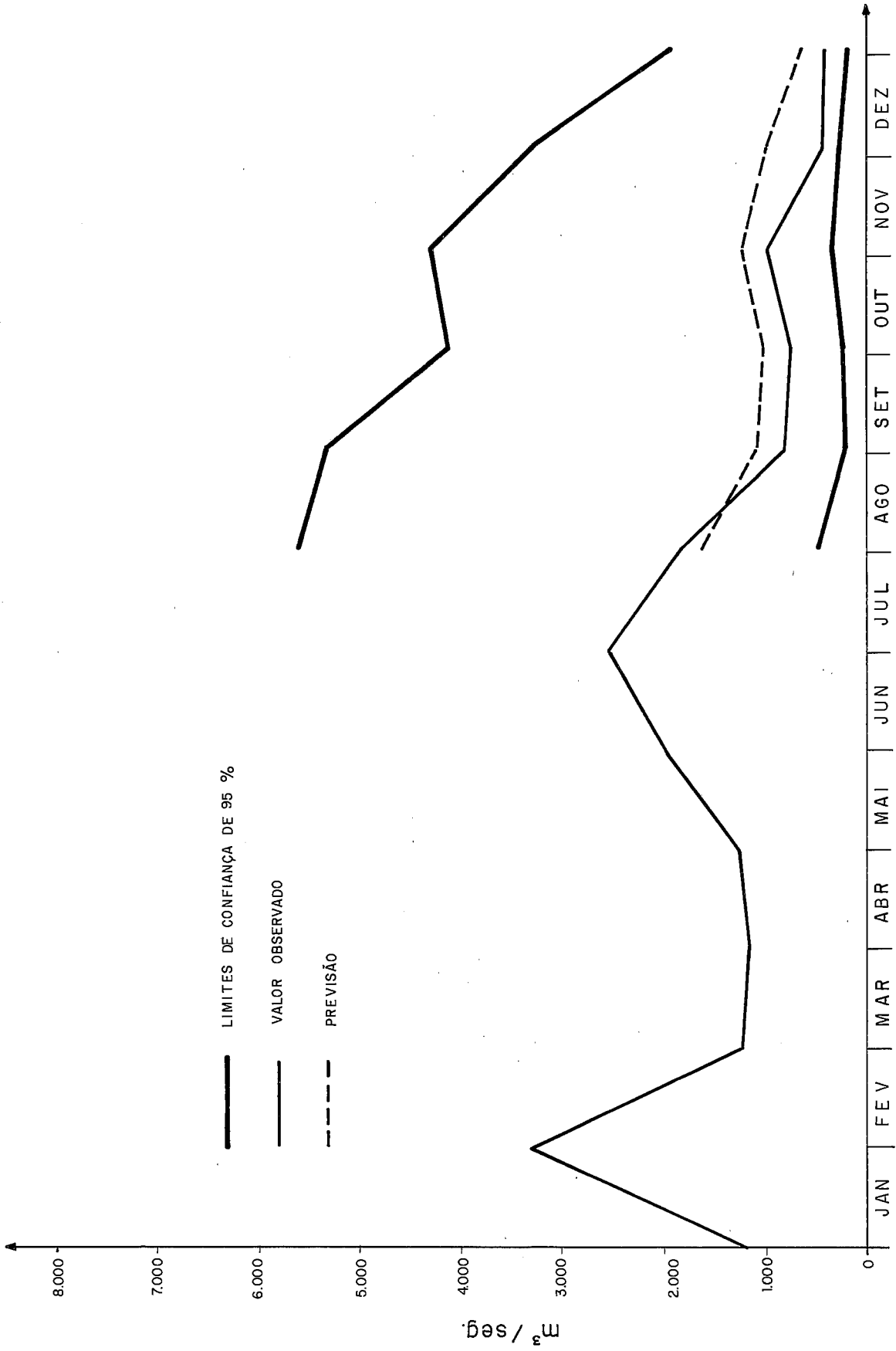


FIG. 4.20 - PREVISÕES DE VAZÕES MÉDIAS MENSIS, OBTIDAS A PARTIR DO MODELO C-1, PARA OS SEIS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA YEVJEVICH / BOX E JENKINS).

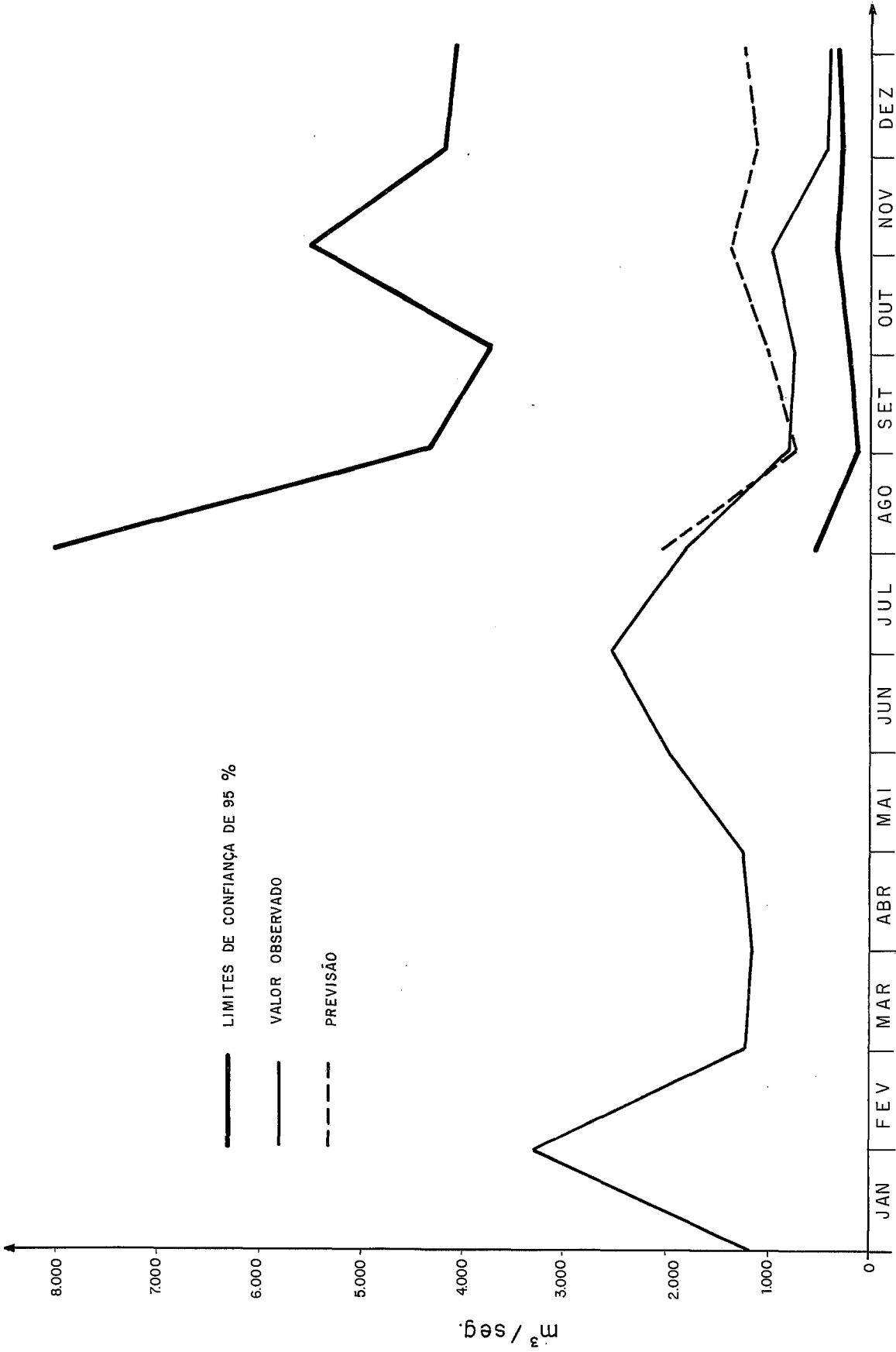


FIG. 4-21 - PREVISÕES DE VAZÕES MÉDIAS MENSIS, OBTIDAS A PARTIR DO MODELO C-4, PARA OS SEIS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA YEVJEVICH / BOX E JENKINS).

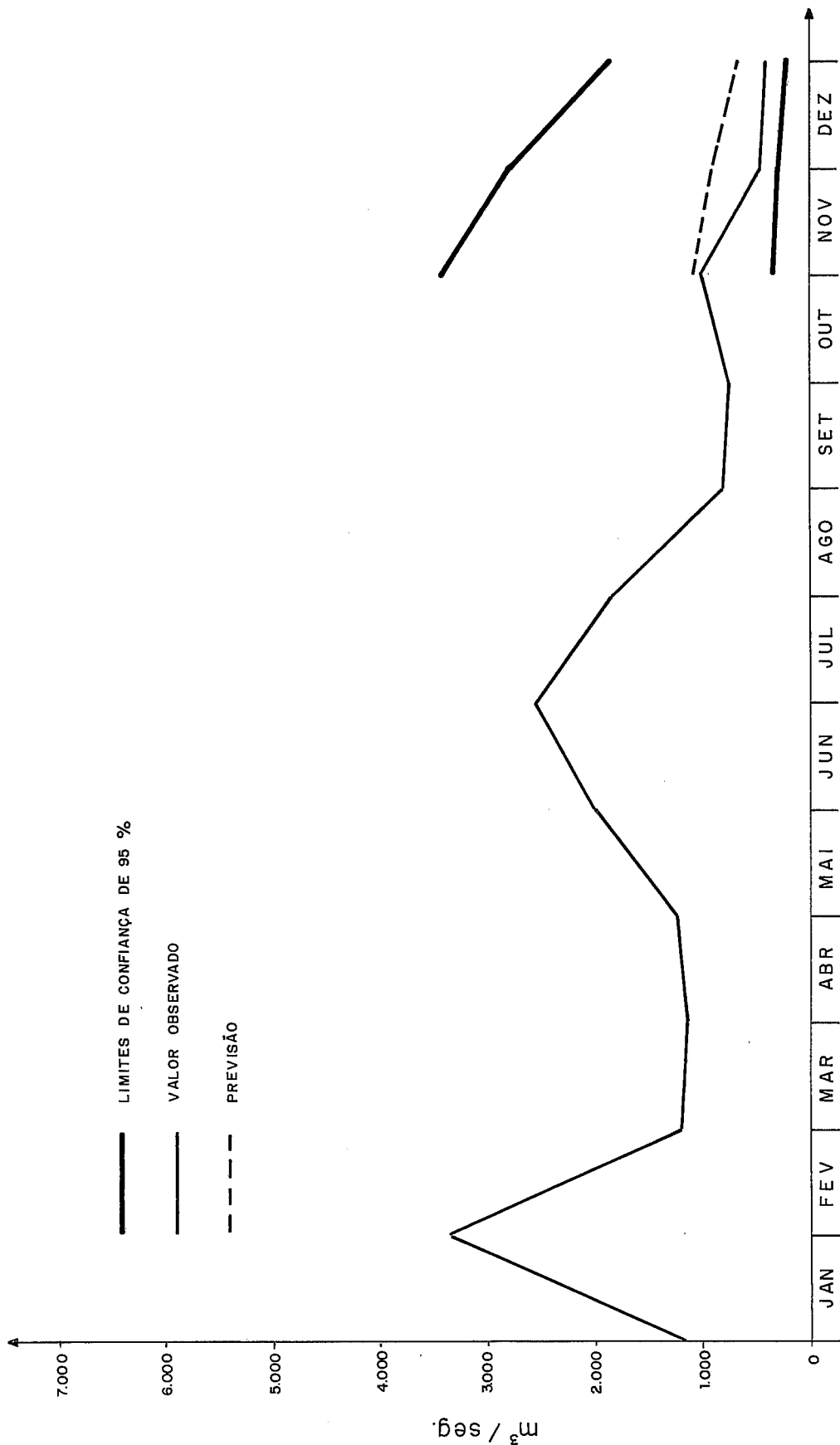


FIG. 4-23 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS A PARTIR DO MODELO A-2 PARA OS TRÊS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE YEVEVICH)

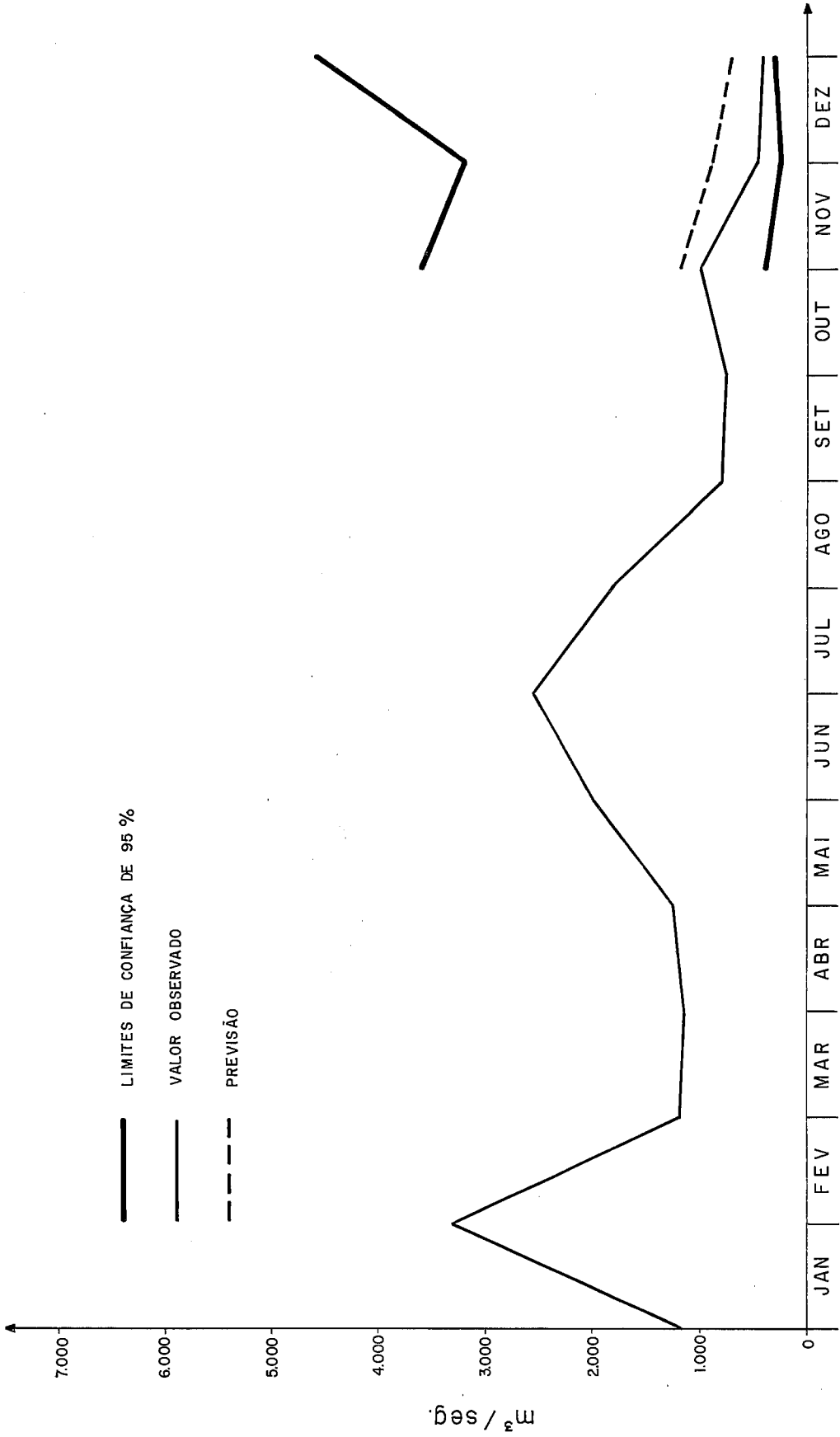


FIG. 4-24 - PREVISÕES DE VAZÕES MÉDIAS MENSAIS, OBTIDAS A PARTIR DO MODELO B-5 PARA OS TRÊS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE BOX E JENKINS).

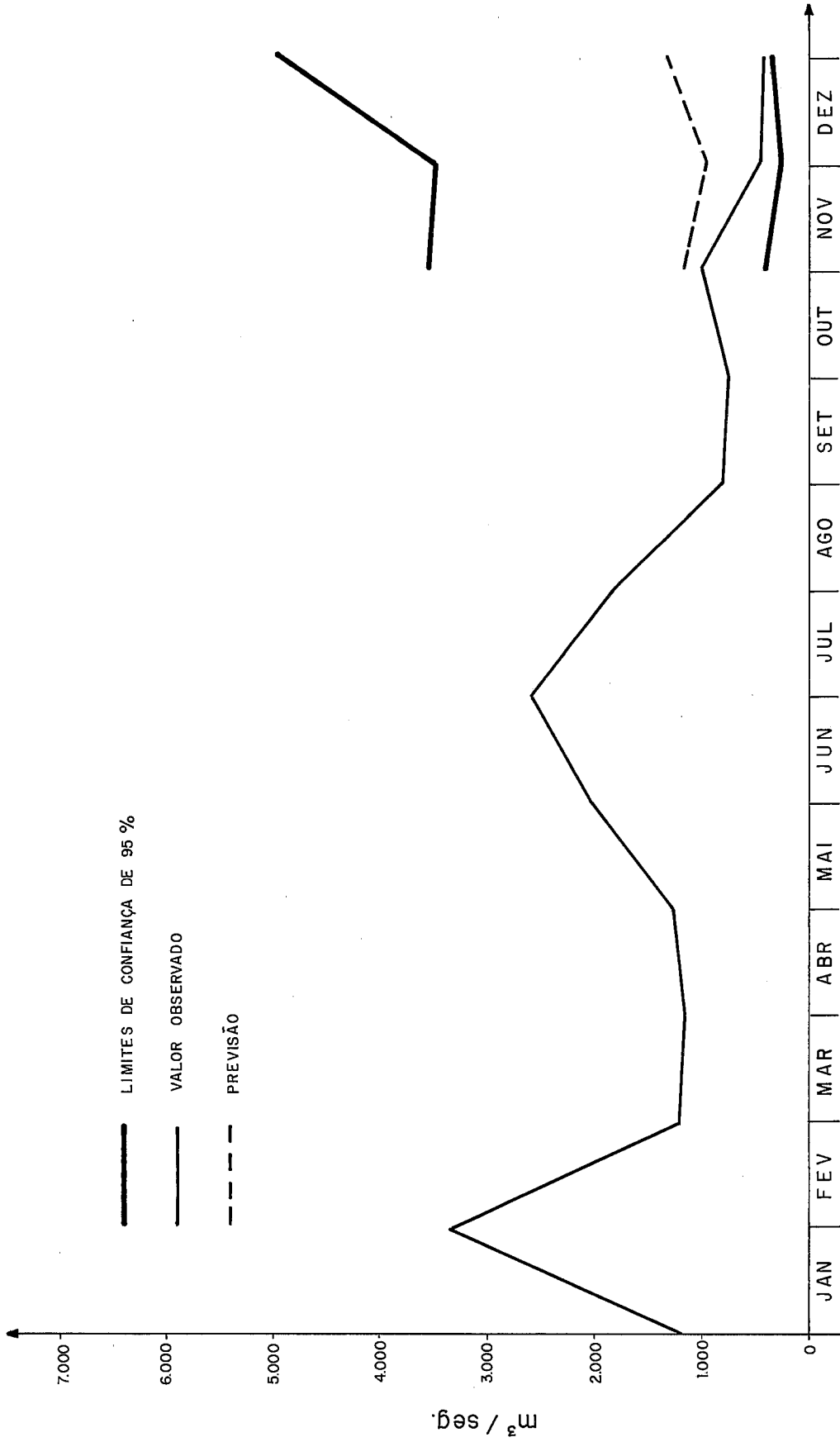


FIG. 4-25 - PREVISÕES DE VAZÕES MÉDIAS MENSIAIS, OBTIDAS A PARTIR DO MODELO B-6 PARA OS TRÊS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA DE BOX E JENKINS).

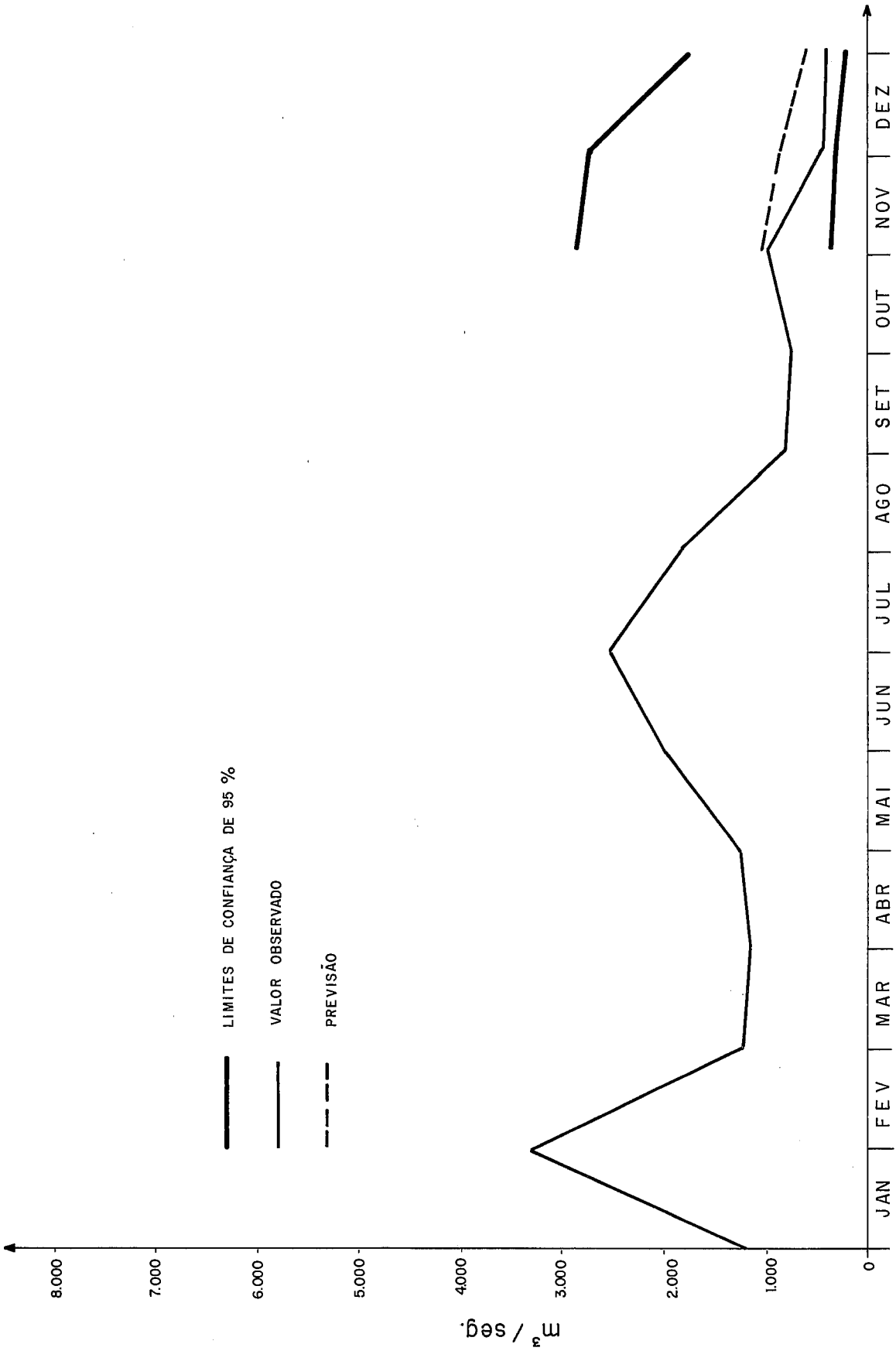


FIG. 4-26 - PREVISÕES DE VAZÕES MÉDIAS MENSIS, OBTIDAS A PARTIR DO MODELO C-1, PARA OS TRÊS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA YEVJEVICH / BOX E JENKINS).

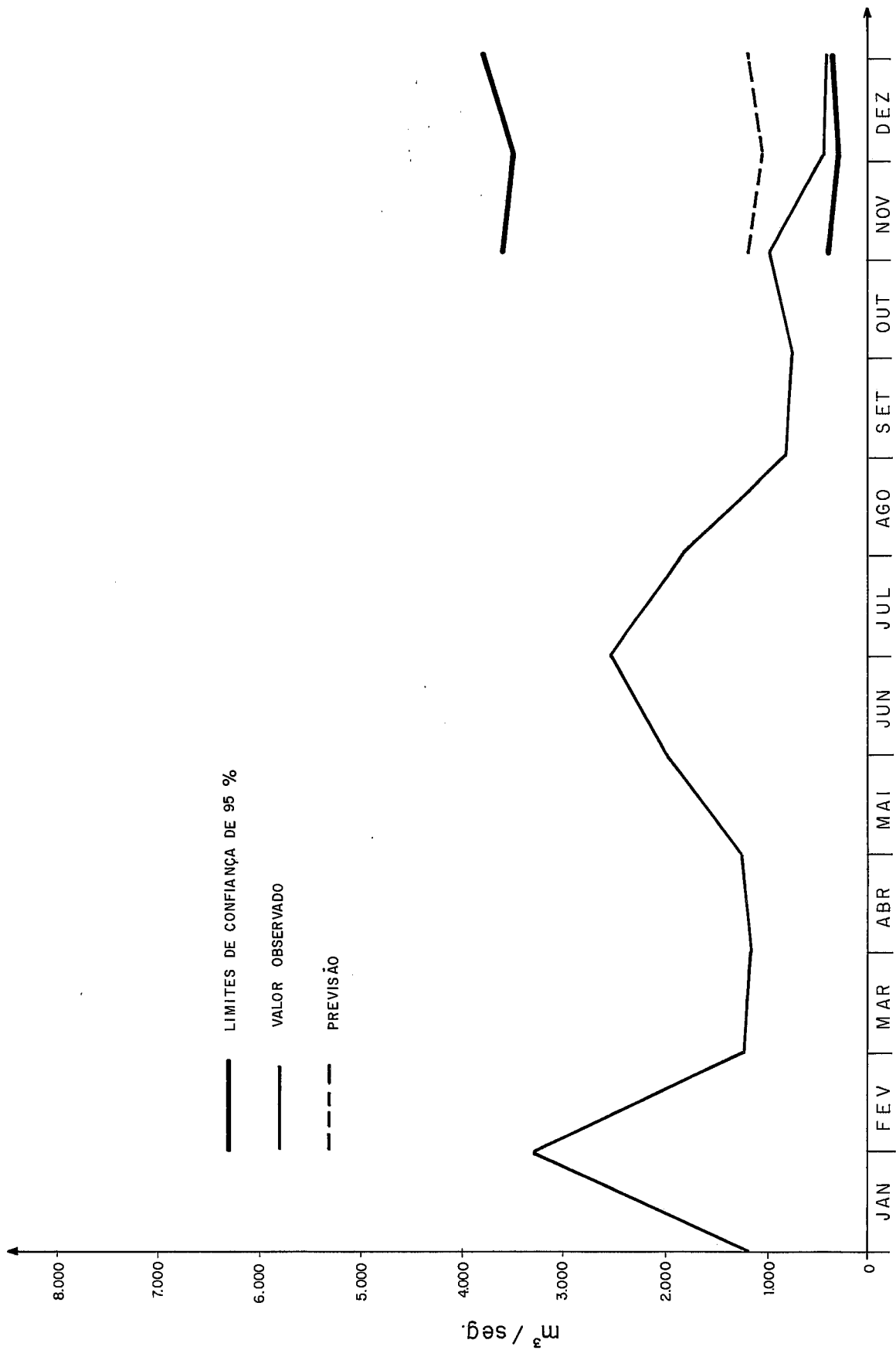


FIG. 4.27- PREVISÕES DE VAZÕES MÉDIAS MENSIS, OBTIDAS A PARTIR DO MODELO C-4, PARA OS TRÊS ÚLTIMOS MESES DE 1971 EM SALTO OSÓRIO (METODOLOGIA YEVJEVICH / BOX E JENKINS).

V - CONCLUSÕES E RECOMENDAÇÕES

Em vista dos resultados anteriores, as seguintes conclusões podem ser tiradas:

- (i) o modelamento de séries hidrológicas temporais objetivando a obtenção de previsões, com vistas ao planejamento da operação de sistemas hidrotérmicos, através da metodologia mista, mostrou-se amplamente viável, do ponto de vista computacional;
- (ii) a integração das metodologias de Yevjevich e Box e Jenkins além de possibilitar a obtenção de um modelo capaz de representar com boa precisão o processo periódico-estocástico da série hidrológica analisada possibilitou, também, a obtenção de um método que satisfaz amplamente como modelo de previsão, no contexto dos que são atualmente utilizados com este propósito;
- (iii) a transformação de Box e Cox aplicada à série coincidiu com a tradicionalmente utilizada em Hidrologia, ou seja, a transformação logarítmica;
- (iv) a seleção do modelo de dependência para a componente estocástica dependente padronizada, de acordo com a metodologia de Box e Jenkins, conduziu a um modelo da classe sugerida por Yevjevich para esta componente (autoregressivo linear de 1ª ordem). Entretanto, não se pode afirmar, *à priori*, que para todas as séries hidrológicas, o modelo

autoregressivo seja o mais apropriado.

(v) em vista de (iv) e (v) era de se esperar, como de fato ocorreu, que os modelos A-2 e C-1 tivessem desempenhos semelhantes;

(vi) os novos testes propostos para seleção do número de harmônicos em parâmetros periódicos e constância de parâmetros apresentaram resultados que não diferiram de maneira significativa daqueles sugeridos por Yevjevich (o que, de certa forma, é muito importante). Entretanto, enquanto no primeiro caso, o teste para seleção do número de harmônicos baseado nos intervalos de confiança dos parâmetros deixava de ser um teste puramente empírico, no segundo caso houve um ganho no tempo de processamento dos programas.

Visando possíveis desenvolvimentos do presente trabalho, as seguintes recomendações podem ser feitas:

(i) o desenvolvimento de um estudo semelhante ao aqui apresentado para os demais postos da Região Sul, bem como para outras regiões e, a partir daí, a seleção de um modelo de previsão que possa ser utilizado para toda uma região;

(ii) desenvolvimento de outros critérios probabilísticos para seleção do melhor (es) modelo (s) para uma dada região;

- (iii) maior importância ao tempo de resposta dos modelos, isto é, uma verificação da adaptação dos modelos a medida que novos valores da série vão sendo conhecidos:
- (iv) seleção de um modelo a partir de previsões obtidas um passo a frente (modelos adaptivos);
- (v) obtenção de previsões multivariadas e desenvolvimento de critérios para seleção de modelos com esta finalidade;
- (vi) adaptação dos modelos para geração de séries univariadas e multivariadas;
- (vii) desenvolvimento de critérios para seleção do melhor modelo de geração para uma dada região.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ANDERSON, O.D. - Time Series Analysis and Forecasting, The Box - Jenkins Approach, Butterworth, 1975.
2. ARAUJO, R.P. - Transformação e Estimação de Parâmetros Para Modelos Adaptados à Previsão de Séries Temporais, Tese de Mestrado, PUC - RJ, setembro, 1974.
3. BARRETTO, L.A.L. - Otimização Implícita Estocástica para Operação de Reservatório, Tese de Mestrado, COPPE-UFRJ, junho, 1974.
4. BLOOMFIELD, P. - Fourier Analysis of Times Series: An Introduction, Wiley, 1976.
5. BOX, G.E.P. e COX, D.R. - An Analysis of Transformations, Journal Royal Statistical Society, B26, 211, 1964.
6. BOX, G.E.P. e JENKINS, G.M. - Times Series Analysis - Forecasting and Control, Holden - Day, 1970.
7. BOX, G.E.P. e PIERCE, D.A. - Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models, Jour. Amer. Stat. Assoc., 64, 1970.

8. CARLSON, R.F., MAC CORMICK, A.J.A. e WATTS, D.G. - Appli
cation of Linear Random Models to Four Annu-
al Flow Series, Water Resources Research ,
Vol. 6, nº 4, setembro, 1970.
9. CEPEL - ELETROBRÁS - Modelo de Programação Dinâmica Es-
tocástica Para Operação de Sistemas Hidro-
Térmicos, 1976.
10. CHATFIELD, C. - The Analysis of Time Series: Theory and
Practice, Chapman and Hall, 1975.
11. CRAMER, H. - Elementos de la Teoria de Probabilidades
y Algunas de sus Aplicaciones, Aguilar, 1970.
12. CRAMER, H. - Metodos Matematicos de Estatística, Agui-
lar, 1970.
13. ELETROBRÁS - Modelo Estatístico Para Previsão de Va -
zões, Departamento de Operação Energética das
Centrais Elétricas Brasileiras S.A., julho ,
1975.
14. ELETROBRÁS - Modelo a Sistema Equivalente para Opera -
ção de Sistemas Hidro-Térmicos, Departamento
de Operação Energética das Centrais Elétri -
cas Brasileiras S.A., 1975.
15. FIERING, M.B. e JACKSON, B.B. - Synthetic Streamflows,
Water Resources Monograph, American Geophys*ic*
al Union, Washington, D.C., 1971.

16. FISHER, R.A. - On the Probable Error of a Coefficient of Correlation Deduced from a Small Sample ,
Metron, 1, nº 4, 3.
17. FISHER, R.A. e YATES, F. - Tabelas Estatísticas Para
Pesquisa em Biologia, Medicina e Agricultura,
Editora Polígono, 1971.
18. FURNAS - Centrais Elétricas S.A. - Previsão de Afluên -
cias a Furnas: Método Box e Jenkins, Coordena -
ção de Planejamento, Assessoria de Métodos
e Modelos, Nota CP. P - 122, outubro, 1975.
19. FURNAS - Centrais Elétricas S.A. - Previsão de Afluên -
cias, Método Box e Jenkins: Testes e Avalia -
ções, Coordenação de Planejamento, Assesso -
ria de Métodos e Modelos, Nota CP. P - 126 ,
janeiro, 1976.
20. HOEL, P.G. - Introduction to Mathematical Statistics ,
Wiley, 1971.
21. HOEL, P.G., PORT, S.C. e STONE, C.J. - Introduction to
Probability Theory, Houghton Mifflin, 1971.
22. HOEL, P.G., PORT, S.C. e STONE, C.J. - Introduction to
Statistical Theory, Houghton Mifflin, 1971.
23. HYPEL, K.W. - Contemporary Box - Jenkins Modelling in
Hydrology, Tese de Doutorado, Waterloo, Ontaa
rio, Canadá, setembro, 1975.
24. JACKSON, B.B. - The Use of Streamflow Models in Planning,

Water Resources Research, Vol. II, nº 1, fevereiro, 1975.

25. JENKINS, G.M. e WATTS, D.G. - Spectral Analysis and its Applications, Holden - Day, 1968.
26. KENDALL, M.G. - Time Series, Griffin, 1973.
27. KENDALL, M.G. e STUART, A. - The Advanced Theory of Statistics Design and Analysis, and Time - Series Vol. 3, C.Griffin, 1967.
28. KMENTA, J. - Elements of Econometrics, Macmillan, 1971.
29. LENTON, R.L., RODRIGUEZ - ITURBE, I.R. e SHAAKE, J.C. - The Estimation of ρ in the First-Order Autoregressive Model: A Bayesian Approach, Water Resources Research, Vol. 10, nº 2, abril, 1974.
30. MATALAS, N.C. - Mathematical Assessment of Synthetic Hydrology, Water Resources Research, Vol. 3, nº 4, 1967.
31. MC KERCHAR, A.I. e Delleur, J.W. - Application of Seasonal Parametric Linear Stochastic Models to Monthly Flow Data, Water Resources Research, Vol. 10, nº 2, abril, 1974.
32. MC MICHEL, F.C. e HUNTER, J.S. - Stochastic Modeling of Temperature and Flow in Rivers, Water Resources Research, Vol. 8, nº 1, fevereiro, 1972.

33. MOOD, A.M., GRAYBILL, F.A. e BOES, D.C. - Introduction to the Theory of Statistics, Mc Graw - Hill, 1974.
34. NELSON, C.R. - Applied Time Series Analysis for Managerial Forecasting, Holden-Day, 1973.
35. NIE, N.H., HULL, C.H. JENKINS, J.G., STEINBRENNER, K. e BENT, D.H. - Statistical Package for the Social Sciences, Second Edition, Mc Graw-Hill, 1975.
36. O'CONNEL, P.E. - A Simple Stochastic Modelling of Hurst's Law, Paper Presented at IASH International Symposium on Mathematical Models in Hydrology, Int. Ass. Sci. Hydrol., Warsaw, Poland, julho, 1971.
37. RALSTON, A. e WILF, H.S. - Mathematical Methods for Digital Computers, Vol. I, Wiley, 1960.
38. ROSS, S.M. - Introduction to Probability Models, Academic Press, 1972.
39. SILVA, L.F.C.A. e SALES, P.R.H. - Aplicação de Modelos Paramétricos Estocásticos Sazonais Lineares e Dados de Vazões Mensais, Projeto Final da Disciplina COP - 760 - Análise e Previsão de Séries Temporais, COPPE, dezembro, 1975 e ELETROBRÁS, Departamento de Operação Energética, Divisão de Estudos de Operação, dezembro, 1975.

40. SINGH, K.P. e LONNQUIST, C.G. - Two - Distribution Method for Modeling and Sequential Generation of Monthly Streamflows, Water Resources, Vol. 10, nº 4, agosto, 1974.
41. SOUZA, R.C. - Identificação e Testes Estatísticos e Modelos Adaptados à Previsão de Séries Temporais, Tese de Mestrado, PUC - RJ, setembro, 1974.
42. VAN DER WAERDEN, B.L. - Statistique Mathématique, Dunod, 1967.
43. WILKS, S.S. - Mathematical Statistics, Wiley, 1962.
44. YEVJEVICH, V.M. - Fluctuations of Wet and Dry Years - Part II - Analysis by Serial Correlation, Hydrology Papers, Colorado State University, Fort Collins, junho, 1964.
45. YEVJEVICH, V.M. - Probability and Statistics in Hydrology, Water Resources Publications, Fort Collins, Colorado, 1972.
46. YEVJEVICH, V.M. - Stochastic Processes in Hydrology, Water Resources Publications, Fort Collins, Colorado, 1972.
47. YEVJEVICH, V.M. - Structural Analysis of Hydrologic Time Series, Hydrology Papers, Colorado State University, Fort Collins, novembro, 1972.

APÊNDICE I

PREVISÃO A PARTIR DA METODOLOGIA DE YEVJEVICH

(i) INTRODUÇÃO

De acordo com a seção IV.2, a análise estrutural segundo YEVJEVICH (1972) foi realizada tanto na série de vazões mensais de Salto Osório quanto no logaritmo neperiano da série, resultando daí, dois modelos: A-1 e A-2. Para o modelo A-1, a distribuição de probabilidade ajustada à componente estocástica independente do modelo foi uma log-normal de três parâmetros, enquanto que, para a mesma componente do modelo A-2, foi ajustada uma distribuição normal.

A apresentação neste apêndice, restringe o desenvolvimento das previsões, L passos à frente de uma origem t , pré-fixada, bem como de seus respectivos intervalos de confiança, à modelos do tipo A-2. Um procedimento bastante análogo ao aqui descrito pode ser utilizado na derivação das previsões e intervalos de confiança para modelos do tipo A-1.

O problema maior, quando se utiliza a metodologia de Yevjevich para previsão, reside na obtenção da variância do erro desta previsão. Como o modelo é composto de uma parte determinística e outra estocástica, as previsões e os intervalos de confiança são obtidas, inicialmente, para a componente estocástica dependente padronizada

do modelo. Uma integração com a componente determinística em conjunto com uma transformação no sentido inverso fornece os resultados requeridos.

(ii) PREVISÃO APÓS TRANSFORMAÇÃO LOGARÍTMICA

Sejam, respectivamente, Z_t e $Z_t^{(\lambda)} = L_n(Z_t)$, a série e sua transformada logarítmica.

Supondo a previsão feita pelo valor esperado, em geral tem-se que:

$$(ii.1) \quad e^{E(Z_{t+1})} \neq E(e^{Z_{t+1}})$$

onde as expressões entre parênteses denotam um valor esperado condicionado à série histórica $Z_t, Z_{t-1}, Z_{t-2}, \dots$

Além disso, por causa da convexidade da função exponencial, neste caso tem-se:

$$(ii.2) \quad e^{E(Z_{t+1})} \leq E(e^{Z_{t+1}})$$

Se o logaritmo da série é normalmente distribuído, então a relação exata vale:

$$(iii.3) \quad E(e^{Z_{t+1}}) = e^{E(Z_{t+1}) + \frac{1}{2} \text{VAR}(Z_{t+1})} \\ > e^{E(Z_{t+1})}$$

A expressão anterior resulta do fato de que se $Z_t^{(\lambda)}$ é normal então Z_t é log-normal.

Recorde-se que se y é uma variável aleatória com distribuição $N(\mu, \sigma^2)$ e x uma variável aleatória tal que $x = \text{antilog}(y)$ ou $x = e^y$, então, diz-se que x possui uma função de densidade log-normal.

A moda da distribuição ocorre no ponto $x = e^{\mu - \sigma^2}$ e a mediana no ponto $x = e^\mu$. Os momentos em relação à origem são dados por:

$$(ii.4) \quad a_j = e^{j\mu + \frac{1}{2} j^2 \sigma^2}, \quad j = 1, 2, 3, \dots$$

Em consequência, tem-se que:

$$\text{m\u00e9dia:} \quad E(x) = a_1 = e^{\mu + \frac{1}{2} \sigma^2}$$

$$\begin{aligned} \text{vari\u00e2ncia:} \quad \text{VAR}(x) &= a_2 - a_1^2 = e^{2\mu + \frac{1}{2} 4\sigma^2} - e^{2\mu + \sigma^2} \\ &= e^{2\mu + \sigma^2} (e^{\sigma^2} - 1) \end{aligned}$$

Desde que y é normal, a probabilidade de que y seja menor ou igual a $(\mu - k\sigma)$, depende somente de k , por exemplo, $P(y \leq \mu - 1,96\sigma) = 0,025$. Em vista da relação entre y e x , tem-se que a probabilidade de que $x < e^{\mu - k\sigma}$ é igual a probabilidade de que $y < \mu - k\sigma$ e, portanto, depende somente de k . A situação é a mesma para $x > e^{\mu - k\sigma}$.

Consequentemente, se os limites de confiança para uma realização de y são $\mu \pm k\sigma$, então, tem-se o

mesmo grau de confiança para que x pertença ao intervalo $e^{\mu} \pm k\sigma$.

Sejam, $Z_{t+L}^{(\lambda)}$ o valor futuro da variável $Z_t^{(\lambda)}$, L passos a frente da origem t e $\hat{Z}_t^{(\lambda)}(L)$ a previsão deste valor.

A distribuição condicional de Z_{t+L} , dado a série $Z_t, Z_{t-1}, Z_{t-2}, \dots$, vale,

$$Z_{t+L} \sim N \{ \hat{Z}_t^{(\lambda)}(L) ; \text{VAR} [e_t^{(\lambda)}(L)] \}$$

onde $e_t^{(\lambda)}(L)$ representa o erro de previsão: $e_t^{(\lambda)}(L) = Z_{t+L}^{(\lambda)} - \hat{Z}_t^{(\lambda)}(L)$.

Logo, se $\hat{Z}_t^{(\lambda)}(L) \pm k_{\lambda/2} \text{DP} [e_t^{(\lambda)}(L)]$ constitui um intervalo de confiança de $(1-\lambda)$ para $Z_{t+L}^{(\lambda)}$ então, $e^{\hat{Z}_t^{(\lambda)}(L) \pm k_{\lambda/2} \text{DP} [e_t^{(\lambda)}(L)]}$, constitui um intervalo de confiança de $(1-\lambda)$ para Z_{t+L} .

(iii) PREVISÃO COM O MODELO A-2

O modelo A-2 se escreve:

$$(iii.1) \quad X_t^{(\lambda)} = e^{X_t^{(\lambda)}}$$

onde,

$$(iii.2) \quad X_t^{(\lambda)} = \bar{X}_T + s_T \varepsilon_t$$

e,

$$(iii.3) \quad \zeta_t = R_1 \zeta_{t-1} + \sqrt{1 - R_1^2} \xi_t$$

com,

$$(iii.4) \quad \zeta_t = \frac{\varepsilon_t - \bar{\varepsilon}}{S_\varepsilon}$$

e,

$$(iii.5) \quad \xi_t : N(0,1) \text{ e independentes entre si.}$$

As previsões são obtidas inicialmente para a componente estocástica dependente padronizada. Seja, então, $\hat{\zeta}_t(L)$ a previsão do valor futuro ζ_{t+L} . Esta previsão é dada por,

$$(iii.6) \quad \hat{\zeta}_t(L) = R_1 \hat{\zeta}_t(L-1)$$

Com relação à previsão para a variável transformada, $X_t^{(\lambda)}(L)$, tem-se:

$$(iii.7) \quad \frac{\varepsilon_t(L) - \bar{\varepsilon}}{S_\varepsilon} = R_1 \frac{\varepsilon_t(L-1) - \bar{\varepsilon}}{S_\varepsilon}$$

ou,

$$(iii.8) \quad \frac{\hat{X}_t^{(\lambda)}(L) - \bar{X}_\tau - \bar{\varepsilon}}{S_\tau} = R_1 \left(\frac{\hat{X}_t^{(\lambda)}(L-1) - \bar{X}_{\tau-1} - \bar{\varepsilon}}{S_{\tau-1}} \right)$$

onde, $L = \tau, 12 + \tau, 24 + \tau, \dots, \tau = \overline{1, 12}$ e, $\hat{X}_t^{(\lambda)}(0) = X_t^{(\lambda)}$

Explicitando $\hat{X}_t^{(\lambda)}(L)$ na equação (iii.8) chega-se a:

$$(iii.9) \quad \hat{X}_t^{(\lambda)}(L) = \bar{X}_T + S_T \bar{\varepsilon} (1-R_1) + S_T \cdot R_1 \left(\frac{\hat{X}_t^{(\lambda)}(L-1) - \bar{X}_{T-1}}{S_{T-1}} \right)$$

A previsão para a variável original \bar{e} dada por,

$$(iii.10) \quad \hat{X}_t^{(\lambda)}(L) = e^{\hat{X}_t^{(\lambda)}(L)}$$

(iv) INTERVALOS DE CONFIANÇA PARA AS PREVISÕES OBTIDAS COM O MODELO A-2

Visando o bom entendimento, a derivação dos intervalos de confiança será feita objetivando a componente estocástica dependente padronizada do modelo. Uma integração com a componente determinística fornece os resultados desejados. Assim, para:

$L = 1$:

$$(iii.11) \quad \zeta_{t+1} = R_1 \zeta_t + \sqrt{1 - R_1^2} \xi_{t+1}$$

$$(iii.12) \quad \hat{\zeta}_t(1) = R_1 \zeta_t$$

erro da previsão:

$$(iii.13) \quad e_t(1) = \zeta_{t+1} - \hat{\zeta}_t(1) = \sqrt{1 - R_1^2} \xi_{t+1}$$

$$(iii.14) \quad DP(e_t(1)) = \sqrt{1 - R_1^2}$$

intervalo de confiança de $(1 - \lambda) \%$:

$$(iii.15) \quad \hat{\zeta}_t(1) \pm k_{\lambda/2} DP(e_t(1))$$

ou

$$(iii.15.a) \quad R_1 \zeta_t \pm k_{\lambda/2} \sqrt{1 - R_1^2}$$

$L = 2$:

$$(iii.16) \quad \begin{aligned} \zeta_{t+2} &= R_1 \zeta_{t+1} + \sqrt{1 - R_1^2} \xi_{t+2} \\ &= R_1^2 \zeta_t + R_1 \sqrt{1 - R_1^2} \xi_{t+1} + \sqrt{1 - R_1^2} \xi_{t+2} \end{aligned}$$

$$(iii.17) \quad \hat{\zeta}_t(2) = R_1 \hat{\zeta}_t(1) = R_1^2 \zeta_t$$

erro da previsão:

$$(iii.18) \quad e_t(2) = \zeta_{t+2} - \hat{\zeta}_t(2) = R_1 \sqrt{1 - R_1^2} \xi_{t+1} + \sqrt{1 - R_1^2} \xi_{t+2}$$

$$(iii.19) \quad DP(e_t(2)) = \sqrt{1 - R_1^2} \sqrt{1 + R_1^2}$$

intervalo de confiança de $(1 - \lambda) \%$:

$$(iii.20) \quad \hat{\zeta}_t(2) \pm k_{\lambda/2} DP(e_t(2))$$

ou,

$$(iii.20.a) \quad R_1 \hat{\zeta}_t(1) \pm k_{\lambda/2} \sqrt{1-R_1^2} \sqrt{1+R_1^2}$$

$L = L$

$$(iii.21) \quad \zeta_{t+L} = R_1^L \zeta_t + R_1^{L-1} \sqrt{1-R_1^2} \xi_{t+1} + \dots + \sqrt{1-R_1^2} \xi_{t+L}$$

$$(iii.22) \quad \begin{aligned} \hat{\zeta}_t(L) &= R_1 \hat{\zeta}_t(L-1) \\ &= R_1^L \zeta_t \end{aligned}$$

erro da previsão:

$$(iii.23) \quad \begin{aligned} e_t(L) &= \zeta_{t+L} - \hat{\zeta}_t(L) = \\ &= R_1^{L-1} \sqrt{1-R_1^2} \xi_{t+1} + \dots + \sqrt{1-R_1^2} \xi_{t+L} \end{aligned}$$

$$(iii.24) \quad DP(e_t(L)) = \sqrt{1-R_1^2} \sqrt{1+R_1^2+R_1^4+\dots+R_1^{2L-2}}$$

intervalo de confiança de $(1-\lambda)\%$:

$$(iii.25) \quad \hat{\zeta}_t(L) \pm k_{\lambda/2} DP(e_t(L))$$

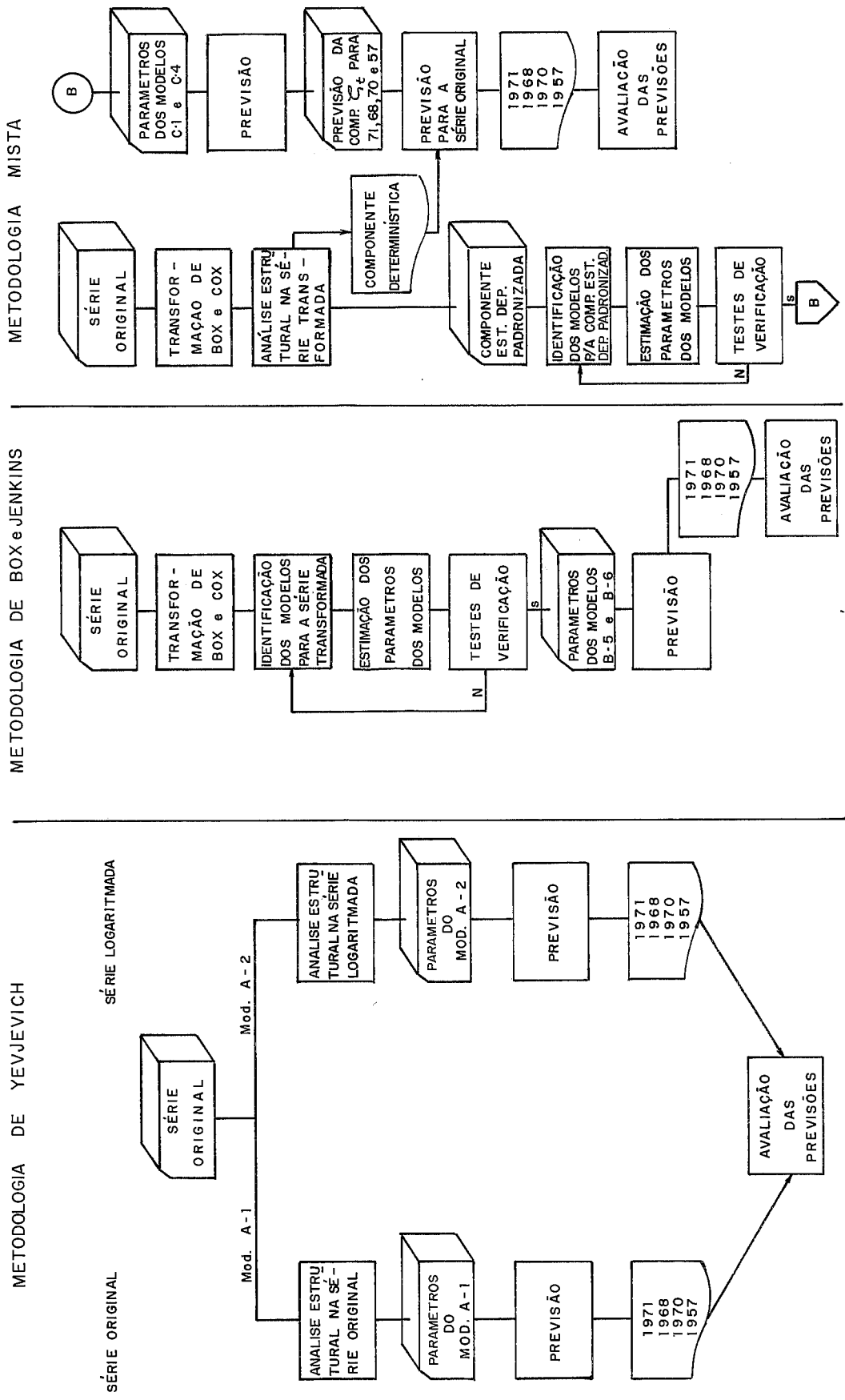
ou,

$$(iii.25.a) \quad R_1 \hat{\zeta}_t(L-1) \pm k_{\lambda/2} \sqrt{1-R_1^2} \sqrt{1+R_1^2+\dots+R_1^{2L-2}}$$

APÊNDICE II

FLUXOGRAMA DOS PROCEDIMENTOS ADOTADOS NAS METODOLOGIAS DE

YEVJEVICH , BOX e JENKINS e MISTA



OBS. OS PROGRAMAS CORRESPONDENTES AO FLUXOGRAMA ACIMA ESTAO DISPONIVEIS PARA OS LEITORES INTERESSADOS