CRITÉRIOS DE IDENTIFICAÇÃO DA ORDEM DO MODELO AUTORREGRESSIVO PERIÓDICO – PAR(P)

Fernando Luiz Cyrino Oliveira; Cristina Vidigal Cabral de Miranda
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900
fcyrino@ele.puc-rio.br; cristina@ele.puc-rio.br

Pedro Guilherme Ferreira; Rodrigo Flora Calili
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900
ferreira@ele.puc-rio.br; rcalili@ele.puc-rio.br

Reinaldo Castro Souza
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225, Gávea - Rio de Janeiro, RJ - Brasil - 22451-900
reinaldo@ele.puc-rio.br

RESUMO

No Brasil a geração de energia elétrica depende essencialmente dos aproveitamentos hidroelétricos, sendo as usinas termoelétricas a segunda maior fonte de geração de energia. No entanto, como essas têm custos mais elevados, é importante o estudo do comportamento das séries de vazões, objetivando otimizar as vazões fluviais. Estudos mostram ser bastante adequado para este tipo de dados o modelo Periódico Autorregressivo – PAR (p), cuja fase de identificação é importante e os métodos adotados nesta fase interferem no resultado final. Assim, este trabalho tem por finalidade discutir o melhor critério a ser adotado na fase de identificação do modelo autorregressivo periódico aplicado à série citada. Os dois critérios adotados se diferem por admitir ou não *lags* intermediários não significativos na fase de identificação. Ver-se-á que o critério que não utiliza os *lags* intermediários têm como resultado, cenários hidrológicos mais parcimoniosos e, por conseguinte, um aprimoramento do modelo de simulação das ENAs.

PALAVRAS CHAVE: Séries hidrológicas. Decomposição clássica de séries. Modelos Periódicos Autorregressivos.

ABSTRACT

The generation of electrical energy in Brazil depends, heavily, on the hydro plants, having the thermal plants as the second source of energy generation in the country. However, as these have higher costs, it becomes essential the study of the behavior of the inflow series at the reservoirs aiming the optimal use of the available inflows. Some existing studies show that the so called periodic autoregressive formulation (PAR(p)), is the recommended one to model such inflows, whose identification phase is important and the methods adopted at this stage affect the final result. This paper discuss the questions associated with this structural identification of the order p of the PAR(p) formulation. Two possible identification alternatives are presented and they differ on the assumption they adopted for the intermediate lags. It will be shown that the approach that does not consider the intermediate lags within the structure produce more parsimonious hydrological scenarios.

KEY WORDS: Hydrological plants. Classical time series decomposition. Periodical autoregressive models.

1. Introdução

A geração de energia elétrica no Brasil depende basicamente das vazões que naturalmente afluem aos 167 aproveitamentos hidroelétricos brasileiros distribuídos por quinze bacias hidrográficas do país, onde, deste total, 119 usinas são despachadas pelo ONS – Operador Nacional do Sistema. O processo natural de vazões fluviais tem como característica principal a sua inconstância, que é dependente da ocorrência de precipitações, fenômeno sabidamente intermitente, e de comportamento sazonal ditado pelo clima (ANEEL, 2010).

Devido à irregularidade das vazões fluviais e à necessidade de manter a continuidade do fornecimento de energia elétrica, o sistema de geração de energia elétrica brasileiro, além de contar com 1.330 usinas termoelétricas de complementação (o que implica em custos de combustíveis), conta também com um conjunto de reservatórios de acumulação, cuja função é essencialmente a regularização dos regimes fluviais, através do armazenamento de água nos períodos de maiores afluências naturais, para ser utilizada durante os períodos mais secos.

A coordenação da operação do sistema de reservatórios do setor elétrico em conjunto com a operação do sistema de usinas termoelétricas de complementação, permite o melhor aproveitamento das vazões naturais, evitando o desperdício de água e gastos excessivos com combustíveis. Esta coordenação é feita no âmbito do chamado Planejamento da Operação do Sistema Interligado, executado atualmente no ONS. Vale também ressaltar que, além da função de regularização para efeito de geração de energia elétrica, parte da capacidade dos reservatórios do setor elétrico é utilizada para controle de cheias. Neste caso, reserva-se uma parcela superior do reservatório para ser ocupada somente durante a ocorrência de grandes cheias.

Assim, no sistema brasileiro, a disponibilidade, ou a capacidade de geração das usinas hidráulicas num dado instante de tempo depende das afluências ocorridas até aquele instante e das decisões operativas já tomadas, ou seja, decisões associadas ao despacho hídrico, térmico, vertimentos, déficits de energia e intercâmbios energéticos entre submercados (Barros et al, 2009).

O plano de operação de um sistema hidrotérmico de energia elétrica deve definir em todos os instantes a geração de cada unidade, de tal forma que a demanda seja atendida a um custo mínimo. A existência de reservatórios com capacidade de regularização significativa torna o problema do planejamento da operação uma questão não-separável no tempo, uma vez que qualquer decisão de deplecionamento destes reservatórios influencia a capacidade de geração do sistema no futuro, sendo essa também influenciada pelas afluências hidrológicas futuras, cuja natureza extrinsecamente aleatória deve ser considerada no planejamento da operação do sistema.

Em geral, o uso de previsões hidrológicas pode melhorar o desempenho da operação do sistema aumentando benefícios e confiabilidade e, reduzindo custos. Potencialmente, o sistema brasileiro de geração de energia elétrica, que contém uma grande capacidade de regularização, pode beneficiar-se largamente do desenvolvimento de melhores modelos de previsão hidrológica. Desta forma, este trabalho tem como objetivo principal discutir o melhor critério a ser adotado na fase de identificação da ordem do modelo autorregressivo periódico utilizando, como estudo de caso, as séries de Energia Natural Afluente (ENA) do Subsistema Sudeste/Centro-oeste. Sendo a ENA de um subsistema igual à soma dos produtos da vazão natural afluente de cada usina, considerando sua produtividade média.

Os objetivos específicos deste trabalho são: (i) fazer uma análise descritiva da série de ENA da região Sudeste/Centro-oeste; (ii) e fazer a identificação dos parâmetros do método periódico para a série em estudo.

A série de ENA utilizada para resolução deste problema é mensal e o horizonte de estudo inicia-se em janeiro de 1931 e termina em dezembro de 2008, totalizando 78 anos de dados e um total de 936 observações.

2. Análise exploratória da série de ENA

Primeiramente, é apresentada, na figura 1, a série da evolução temporal da ENA para o subsistema Sudeste/Centro-oeste. Nesse, é possível notar claramente que esta é uma série altamente sazonal, refletindo o comportamento das chuvas dentro do ano. Além disso, não se observa tendência de crescimento e a variação parece constante, isto é, os processos estocásticos geradores das séries parecem estacionários de segunda ordem (média e variância constantes ao longo do período de análise).

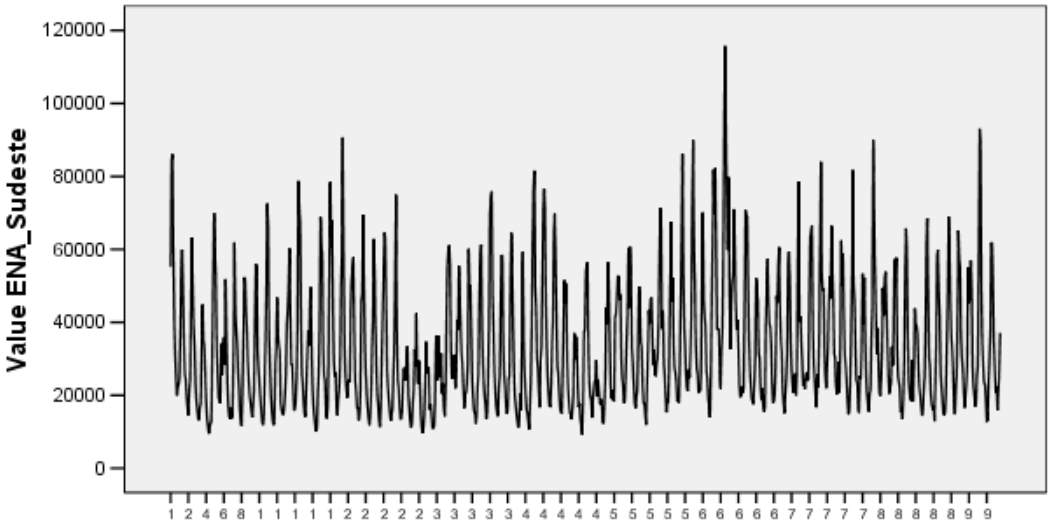

Figura 1 – ENA subsistema Sudeste/Centro-Oeste

Os métodos de decomposição assumem que uma série temporal é constituída por um conjunto de componentes não-observáveis. Dessa forma, pela identificação das componentes individuais presentes no padrão básico da série histórica de dados (tendência, sazonalidade e aleatoriedade), a extrapolação para o futuro pode ser realizada.

Segundo Morettin & Toloi (1987), a componente sazonal representa as flutuações da série de acordo com algum fator de sazonalidade. O ciclo apresenta um comportamento similar à componente sazonal, embora tenha normalmente comprimento maior que aquela. Justamente pelo fato de não apresentar duração uniforme, a identificação da componente ciclo é mais problemática. A tendência representa o aumento ou declínio gradual nos valores das observações de uma série temporal. Com a remoção das componentes de sazonalidade e tendência, a componente aleatória fica determinada.

Wheelwright & Makridakis (1989) mostram que vários procedimentos para a decomposição de séries temporais foram desenvolvidos, cada qual tentando isolar as componentes não-observáveis da série o mais acuradamente possível. O objetivo desses procedimentos consiste em remover cada uma das componentes, permitindo que o comportamento da série temporal seja mais bem compreendido e, conseqüentemente, prognosticar valores futuros mais apropriados.

A seguir, na tabela 1, têm-se os resultados referentes à decomposição nas componentes de tendência, sazonalidade e irregular para a série do subsistema Sudeste/Centro-oeste, sendo predominante a componente sazonal e sendo a componente aleatória (irregular) a que tem menor participação na série, indicando, a princípio, o comportamento padrão da mesma.

Tabela 1 – Decomposição da série – ENA (MWmed)

Componentes (%)	Subsistema Sudeste / Centro-oeste
Tendência	14,3%
Sazonalidade	72,9%
Irregular	12,8%

Este resultado pode ser confirmado na figura 2 a seguir, onde é apresentada a função de autocorrelação (FAC). Como esta decai lentamente ao observar os *lags* sazonais, pode-se afirmar que se trata de uma série com sazonalidade, como já esperada na série do subsistema Sudeste/Centro-Oeste.

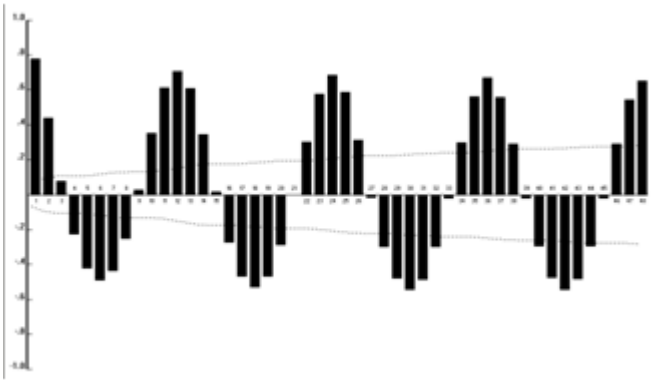

Figura 2 – FAC subsistema Sudeste/Centro-Oeste

Dando continuidade às análises exploratórias das componentes da série histórica e dada a importância do efeito sazonal, a seguir retrata-se de forma mais detalhada o comportamento desta componente na série. Os processos de ajuste sazonal são largamente empregados na análise de dados econômicos, sendo uma das principais razões para esta análise o fato deles poderem representar o efeito de fatores exógenos e não controláveis.

Existem dois interesses principais no ajuste sazonal: o estudo da sazonalidade propriamente dita e a remoção da sazonalidade da série para posterior estudo de seus demais aspectos. Este último é o que está presente na maioria dos trabalhos, nos quais está implícita a idéia de que a existência de movimentos sazonais afetam o reconhecimento e a interpretação de importantes movimentos não-sazonais numa série, tais como: mudanças de tendência e outros eventos cíclicos, padrões emergentes ou ocorrências inesperadas para os quais possíveis causas são procuradas.

Em outras palavras, é possível determinar um fator sazonal para cada um dos períodos analisados (e.g. mês). A análise de tais fatores permite reconhecer padrões que irão influenciar o nível da série hidrológica, refletindo, no caso dos fenômenos naturais, o comportamento esperado ao longo do ano, isto é, espera-se que em meses de chuva um fator sazonal positivo e o contrário em períodos de secas.

Com o objetivo de facilitar a interpretação dos fatores sazonais obtidos, a seguir é mostrado o gráfico que reflete o comportamento de tais fatores ao longo dos meses analisados para o subsistema estudado.

Figura 3 – Fatores Sazonais – Sudeste/Centro-Oeste

Como pode ser observado na figura 3, os fatores sazonais no subsistema Sudeste/Centro-Oeste são bem definidos de acordo com as estações do ano, isto é, conforme esperado, valores sazonais elevados – refletindo a intensidade das chuvas – nos meses de dezembro a março. Em contrapartida, nos períodos secos (inverno – julho a setembro), os fatores sazonais são negativos, indicando a queda da intensidade dos regimes hidrológicos.

É possível que efeitos de longo prazo, como El Niño, La Niña e Manchas Solares, estejam presentes na série estudada. Para que estes fenômenos sejam investigados modelos fracionários – ARFIMA (p,d,q) – onde o parâmetro "d" de diferenciação é fracionário podem ser adotados (Trevisan et al, 2000).

A próxima seção contempla a análise da série através dos modelos autorregressivos periódicos – PAR (p) – com vistas à simulação de cenários sintéticos que reproduzam as características das séries originais e sirvam de base para o planejamento da operação energética de médio prazo.

3. Modelo Autorregressivo Periódico – PAR(p)

De acordo com Oliveira (2010) e conforme mostrado em Hipel & McLeod (1994), algumas séries históricas, dentre essas as hidrológicas sazonais, exibem uma estrutura de autocorrelação que depende não somente do intervalo de tempo entre as observações, mas também do período observado. Ainda, segundo Salas, Boes & Smith (1982), os processos estocásticos naturais são, em geral, estacionários em sentido lato, isto é, os momentos de primeira e segunda ordem da distribuição de probabilidades não são afetados por variações devido à escolha da origem dos tempos, um dos pressupostos para a aplicação da metodologia Box & Jenkins (Harvey, 1981).

Na classe de modelos periódicos, dois modelos se destacam: PAR (*periodic autoregressive*) e PARMA (*periodic* ARMA). O modelo PAR (p) ajusta para cada período da série um modelo AR (p). Em hidrologia, a modelagem PAR (p) surgiu a partir das pesquisas de Thomas & Fiering (1962), de acordo com Hipel & McLeod (1994).

De acordo com Maceira (1989), séries hidrológicas com periodicidade menor que o ano, tais como séries mensais, têm como característica o comportamento periódico das suas propriedades probabilísticas, como, por exemplo, a média, a variância, a assimetria e a estrutura de autocorrelação.

Para se obter os modelos periódicos, tem-se como primeiro passo fazer a identificação deste modelo. No entanto, essa fase tem sido alvo de críticas por alguns autores, como Stedinger (2001) e Oliveira (2010), dado que ordens mal escolhidas podem gerar resultados não confiáveis e previsões/simulações errôneas.

Segundo Kirchner et al (2008), as funções de autocorrelação (FAC) e de autocorrelação parcial (FACP) na análise clássica de séries temporais têm obtido grande aceitação na sua aplicação para identificação de estruturas lineares de séries temporais.

Desta forma, a identificação clássica do modelo consiste em determinar as ordens mais apropriadas dos operadores autorregressivos de cada período p_m, m = 1, ..., s. Isto pode ser feito obtendo-se estimativas dessas autocorrelações, isto é, φ_{kk}^{m}, k = 1, ...,N/4 e substituindo na equação de Yule-Walker as autocorrelações pelos respectivos valores amostrais. Se a ordem do operador autorregressivo de um período qualquer m é p_m então φ_{kk}^{m} para k > p_m tem distribuição aproximadamente Gaussiana com média zero e variância 1/N (aproximação de Bartlett). Para cada período m procura-se a maior ordem i tal que todas as estimativas φ_{kk}^{m} para k > i não sejam mais significativas, isto é, procura-se qual é a maior ordem da autocorrelação cujo coeficiente é significativo. Esta forma de identificação é a utilizada no Newave (Cepel, 2010), sendo este o programa oficial utilizado pelo ONS, empregado para resolver problemas de planejamento da operação de sistemas hidrotérmicos interligados empregando a técnica de programação dinâmica dual estocástica. Nesse, a ordem máxima considerada é seis. Assim, tem-se que seis é o número máximo de defasagens que o programa permite atualmente (sendo que este número máximo já chegou a ser 11, no entanto esta ordem tão elevada e a opção de encontrar a ordem através do último *lag* significante podem ter sido a causa do problema ocorrido em 2007).

Na prática, modelos autorregressivos com ordem elevada apresentam maiores chances de conter coeficientes autorregressivos negativos que eventualmente, poderão produzir coeficientes positivos indesejáveis nos cortes de Benders (fase de otimização). Neste caso, a princípio, seria razoável reduzir a ordem dos modelos. Contudo, a incorporação de um número de afluências passadas no modelo de geração de séries sintéticas tem o objetivo bem sucedido de melhorar a representação de secas longas. O modelo PAR (p) é capaz de gerar séries sintéticas que reproduzem com sucesso os valores médios e dispersões mensais e anuais, assim como as estatísticas dos períodos secos. No caso deste problema, uma solução possível é analisar novamente a função de autocorrelação parcial do mês em questão de k = p_{original} até k = 1, procurando o próximo coeficiente φ_{kk}^{m} significativo que não gere problemas futuros. A ordem deste novo coeficiente será o novo valor de p.

Desta forma, a fim de conciliar os objetivos originais do modelo PAR(p) com a sua incorporação no modelo de otimização Newave de forma explícita, deve-se buscar uma ordem suficientemente grande, mas que não produza coeficientes positivos nos cortes de Benders associados à afluência prevista para o mês em curso. O procedimento de análise dos coeficientes é descrito em Maceira & Damazio (2004) e deve ser repetido até que o conjunto de parâmetros de todos os meses não produza contribuição negativa em qualquer mês futuro. A possibilidade de um coeficiente autorregressivo negativo gerar coeficiente positivo associado à afluência prevista para o mês em curso no corte de Benders pode ser avaliada a partir do exame da equação de regressão da afluência de um mês futuro qualquer, em função da afluência prevista para o mês em curso e das afluências passadas.

No entanto, algumas considerações ainda merecem destaque. Conforme já mencionado, consideremos o exemplo em que um PAR (6) é utilizado para modelar o mês de fevereiro de uma série sintética. A FACP deste período seguiria o seguinte comportamento hipotético:

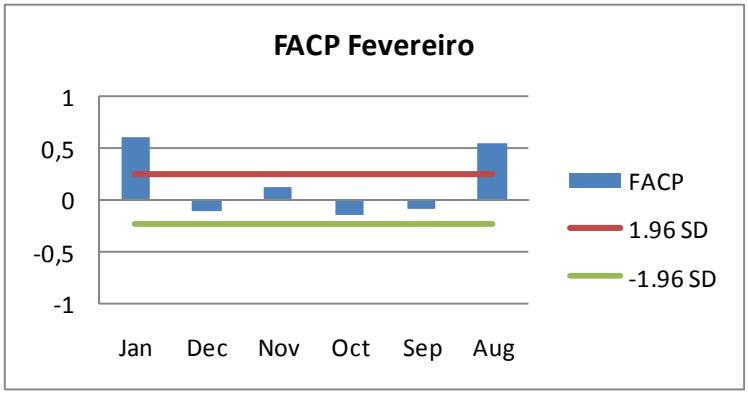

Figura 4: FACP hipotética de um PAR(6) para o mês de fevereiro

Conforme mostrado em Stedinger (2001), para o mês estudado, a escolha de seis parâmetros autorregressivos se dá em função da teoria proposta: "para cada período m procura-se a maior ordem i tal que todas as estimativas φ_{kk}^{m} para k > i não sejam mais significativas". Apesar de φ_2^{Fev} a φ_5^{Fev} serem não significativos (comportando-se como uma variável aleatória de média nula e variância conhecida, ruído branco), φ_6^{τ} mostra-se relevante do ponto de vista estatístico.

Entretanto, Stedinger (2001) propõe algumas reflexões acerca da identificação e da estimação dos parâmetros. Conforme já dito, os resultados obtidos através das modelagens estocásticas servirão de base para os sistemas de otimização via programação dinâmica dual estocástica, pelo Newave. Os coeficientes negativos obtidos em alguns períodos podem gerar valores positivos associados ao aumento de afluências nas funções que representam o custo esperado de operação do sistema. Isto significa que elevados fluxos durante o mês de setembro irão aumentar o custo da operação de fevereiro. Evidentemente, o resultado esperado é exatamente o oposto. Altos níveis de vazão normalmente reduzem o custo de operação um sistema hidroelétrico.

Stedinger (2001) propõe considerar φ_6^{Fev} como um caso espúrio e justifica a possibilidade de ocorrência em função dos níveis de significância considerados. O pesquisador sugere que a identificação dos modelos siga outra regra e afirma que uma modelagem que considera individualmente cada mês pode não produzir o melhor ou um conjunto adequado de modelos. Sem maiores detalhes matemáticos e dada a natureza dos fenômenos naturais envolvidos, o autor questiona que não faria sentido considerar que os fluxos de fevereiro, por exemplo, dependam daqueles de dezembro, novembro, outubro, setembro e agosto e, para as previsões de janeiro seja necessário somente o histórico de dezembro.

Ele ainda finaliza afirmando que a solução para o desafio da modelagem enfrentado seria a adoção de uma visão mais ampla do modelo de decisões que deve ser guiado pela razoabilidade hidrológica dos modelos selecionados com o reconhecimento da intrínseca ligação entre os modelos escolhidos para diferentes períodos.

Haja vista o exposto, Oliveira (2010) propõe as seguintes alternativas para as regras de seleção das ordens dos polinômios autorregressivos periódicos:

Critério 1: para cada período m procura-se a maior ordem i tal que todas as estimativas φ_{kk} para k > i não sejam mais significativas, com ordem máxima igual a seis (seguindo a orientação usada pelo Newave).

Critério 2: para cada período m procura-se a maior ordem i tal que todas as estimativas φ_{kk} para k < i sejam significativas. Este é o critério sugerido por Stedinger (2001) e não admite *lags* intermediários não-significantes. Nesse caso, a ordem máxima considerada também é seis.

A seguir, têm-se os resultados numéricos das funções de autocorrelação parcial para todos os períodos do subsistema Sudeste/Centro-Oeste. Com isso, pode-se ver que há diversos meses em que existem *lags* não significantes no meio de *lags* significantes, o que estaria sobre-estimando a ordem dos modelos se esses forem calculados pelo critério 1. Ao final são mostradas as ordens p identificadas utilizando os Critérios 1 e 2.

Conforme descrito, o teste para verificação da significância dos parâmetros baseia-se nas aproximações de Bartllet (FAC) e Quenouille (FACP) e o valor absoluto do limites de significância é aproximadamente $1{,}96*N^{0,5}$, em que N corresponde ao número de anos do histórico.

Nas próximas figuras, os valores dos coeficientes são expostos e os significativos estão marcados por fundo cinza. Para a FAC, em geral (com exceção de alguns períodos e lags espúrios), é possível notar o decrescimento dos valores para cada mês à medida que o número de *lags* cresce, o que caracteriza a função de autocorrelação teórica para um modelo autorregressivo, de acordo com Souza & Camargo (2004). Segundo o autor, para um modelo AR (p), a função de autocorrelação é infinita (exponenciais e/ou senóides amortecidas) e os coeficientes não se anulam bruscamente (significância estatística). Entretanto, este comportamento deve ser analisando conjuntamente com a função de autocorrelação parcial, a fim de identificar possíveis estruturas mistas (ARMA).

Diferentemente da FAC, a FACP do modelo PAR (p) apresenta-se finita, com poucos *lags* significativos e cortes bruscos nos primeiros (em geral). Comumente, a identificação da ordem dos polinômios é feita somente com base na observação visual dos gráficos da função. Entretanto, como neste trabalho a intenção é analisar a fundo e propor alternativas à fase de inicial da metodologia Box & Jenkins para os modelos periódicos utilizados no planejamento energético brasileiro, é importante conhecer cada um dos coeficientes das funções (FAC e FACP) a fim de que sua real significância seja investigada.

Mês	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11
Jan	0,5913	0,3835	0,2984	0,2734	0,3658	0,3455	0,2541	0,2759	0,2777	0,1711	0,1228
Fev	0,5808	0,2473	0,2027	0,0873	0,0605	0,2286	0,2461	0,1095	0,1731	0,1551	0,1423
Mar	0,6383	0,3808	0,1078	0,1387	0,0642	-0,0603	0,1476	0,2074	0,0696	0,1516	0,1744
Abr	0,7730	0,6251	0,3576	0,1957	0,2539	0,2144	0,1567	0,3419	0,3426	0,1923	0,2230
Mai	0,7911	0,6525	0,6890	0,4207	0,2708	0,2335	0,2422	0,1881	0,4066	0,4107	0,2591
Jun	0,8132	0,6413	0,5314	0,5613	0,4480	0,3242	0,3031	0,2505	0,1834	0,3507	0,3947
Jul	0,8815	0,8009	0,7268	0,6060	0,6032	0,5412	0,3959	0,3720	0,2573	0,1733	0,3665
Ago	0,8221	0,7361	0,7284	0,6098	0,5018	0,4749	0,4361	0,3607	0,3328	0,2407	0,2150
Set	0,8125	0,7106	0,6643	0,6084	0,5158	0,4157	0,4308	0,3706	0,3344	0,3458	0,3048
Out	0,6683	0,6528	0,6546	0,6203	0,5702	0,5373	0,3614	0,3962	0,2411	0,3023	0,2844
Nov	0,7196	0,4905	0,4634	0,5236	0,5233	0,4138	0,3268	0,2445	0,3544	0,2961	0,2303
Dez	0,6957	0,5382	0,4730	0,5008	0,4770	0,4733	0,3451	0,3054	0,1645	0,2134	0,2033

Figura 5: Função de autocorrelação – subsistema Sudeste/Centro-oeste

Mês	Lag 1	Lag 2	Lag 3	Lag 4	Lag 5	Lag 6	Lag 7	Lag 8	Lag 9	Lag 10	Lag 11
Jan	**0,5913**	-0,0541	-0,0023	0,0059	**0,2449**	0,0786	**-0,3386**	0,1883	0,0510	-0,0282	-0,1420
Fev	**0,5808**	-0,1477	0,0956	-0,1366	-0,0488	**0,3567**	0,1597	**-0,2899**	0,0151	-0,0944	0,0354
Mar	**0,6383**	0,0152	-0,0894	0,0920	0,0084	-0,1769	**0,2592**	0,1817	-0,1840	0,0569	0,0619
Abr	**0,7730**	**0,2223**	-0,0179	0,1280	0,1307	0,1463	0,1605	0,1830	-0,0442	-0,1296	-0,1456
Mai	**0,7911**	0,1019	**0,3252**	0,0363	0,0944	-0,0692	0,1500	0,0174	0,2086	0,1113	-0,0432
Jun	**0,8132**	-0,0055	0,0055	0,0015	0,1593	0,0550	0,0958	-0,0742	-0,0290	-0,0323	0,1092
Jul	**0,8815**	**0,2483**	**0,2526**	0,0452	0,0526	0,1728	0,0662	0,0453	-0,1038	-0,0536	0,0583
Ago	**0,8221**	0,0515	0,2099	-0,0955	-0,0273	-0,1170	0,0338	0,0536	0,0310	-0,0445	0,0889
Set	**0,8125**	0,1315	0,1343	-0,0817	0,0232	-0,0209	0,0677	-0,0310	0,0581	0,0785	0,1489
Out	**0,6683**	**0,3233**	**0,3156**	0,1287	0,0382	0,1611	-0,1794	0,0148	-0,1569	0,1318	-0,0580
Nov	**0,7196**	0,0173	-0,0405	0,1880	0,1506	-0,1447	-0,1982	0,0515	0,1380	0,0951	-0,0996
Dez	**0,6957**	0,0778	0,1938	**0,2535**	-0,0512	0,0594	-0,1947	0,1089	-0,1914	-0,0876	-0,0653

Figura 6: Função de autocorrelação parcial – subsistema Sudeste/Centro-oeste

Para este subsistema, verifica-se um decaimento exponencial para a FAC da maioria dos meses, o que indica a ocorrência de um processo AR. Entretanto, no mês de março nota-se um corte brusco a partir do *lag* 2.

Quando se analisa a FACP percebe-se que para a maioria dos meses existem cortes bruscos até os primeiros 3 *lags*, reafirmando a ocorrência do processo AR e justificando o uso do PAR. No entanto, existem para os meses de janeiro, fevereiro e março *lags* significativos de ordem maior que seis. As ordens para estes três meses seriam 7, 8 e 7, respectivamente. Vale ressaltar que pelo Critério 1, especificado anteriormente e adotado pelo software Newave, as ordens para os meses supracitados seriam, 5, 6 e 1, respectivamente.

A seguir as ordens p identificadas para cada um dos períodos de todos os subsistemas com base nos Critérios 1 e 2, definidos anteriormente e a diferença entre as ordens desses critérios.

Tabela 2 – Ordens identificadas nos Critérios 1 e 2

Mês	Criério 1	Criério 2	Diferença
Janeiro	5	1	4
Fevereiro	6	1	5
Março	1	1	0
Abril	2	2	0
Maio	3	1	2
Junho	1	1	0
Julho	3	3	0
Agosto	1	1	0
Setembro	1	1	0
Outubro	3	3	0
Novembro	1	1	0
Dezembro	4	1	3

Pela figura 6, é claro notar diferenças nas ordens identificadas em alguns períodos e é nítido que o Critério 2 é consideravelmente mais parcimonioso que o Critério 1, sendo a parcimônia um adjetivo importante de um modelo de previsão.

Há períodos em que houve divergências entre os critérios empregados para identificação das ordens. Assim, é possível verificar que essas diferenças são bastante elevadas em alguns casos, o que certamente ocasionará impactos no processo de geração de cenários visto que a ordem de defasagem é fundamental para a definição do modelo PAR (p) empregado.

Corroborando o exposto em Stedinger (2001) e Oliveira (2010), o Critério 1 apresenta uma tendência a superestimar as ordens dos polinômios autorregressivos, provavelmente em função da consideração de *lags* elevados espúrios.

4. Conclusão

Neste trabalho foram apresentadas análises da série de ENA para o subsistema Sudeste/Centro-oeste. A decomposição esta mostrou a predominância da componente sazonal e que a componente aleatória (irregular) foi a que apresentou menor participação nesta série, indicando, a princípio, o comportamento padrão da mesma.

Com relação ao fator sazonal desse subsistema, verificou-se que esse é bem definido de acordo com as estações do ano, refletindo a maior intensidade dos regimes hidrológicos nos meses de dezembro a março e o período de seca para os demais meses.

Na segunda etapa do trabalho, partiu-se para a análise da série de forma desagregada mensal através dos modelos autorregressivos periódicos - PAR (p) - com vistas à simulação de cenários sintéticos que reproduzam as características das séries originais e que sirvam de base o planejamento da operação energética de médio prazo.

Constatou-se que este tipo de modelagem pode apresentar o problema de superestimação dos parâmetros, uma vez que considera como ordem para o parâmetro "p" o último *lag* significativo, possuindo este a ordem máxima de seis. No entanto, pode ocorrer (e conforme demonstrado neste artigo, acontece com certa freqüência) que os *lags* intermediários ao último significante sejam não significativos, o que estaria superestimando a ordem do modelo e, conseqüentemente, deteriorando o ajuste das séries. Além disso, vale lembrar que de acordo com a teoria de Box & Jenkins, deve-se buscar a parcimônia na modelagem, priorizando modelos de ordens menores.

Dessa forma, sugere-se o uso do critério 2, o qual considera como ordem do modelo o último *lag* significante desde que todos os anteriores sejam também todos significativos. Com isso, espera-se obter um melhor ajuste para as séries em questão, o que irá resultar na criação de cenários hidrológicos mais consistentes e, conseqüentemente, em um aprimoramento do modelo de previsão das ENAs.

Finalmente, conforme é mostrado em Oliveira (2010), haja vista as divergências encontradas e o impacto negativo da escolha errônea das ordens dos modelos PAR (p), faz-se necessário uma investigação mais profunda desta fase da modelagem. Isto posto, sugere-se como trabalhos futuros: a finalização das análises do modelo PAR (p) e técnicas de computação intensiva que podem ser empregadas com o objetivo de uma detalhada verificação da correta ordem de defasagem dos modelos. Ademais, sugere-se que fenômenos de longo prazo, como El Niño, La Niña e Manchas Solares, também sejam investigados na série em estudo.

Referências Bibliográficas

ANEEL – *Agência Nacional de Energia Elétrica*. Disponível em: www.aneel.gov.br. Acesso em 29 de abril de 2010.

Barros, M., Mello, M. F., Souza, R. C. (2009), *Aquisição de energia no mercado cativo brasileiro: simulações dos efeitos da regulação sobre o risco das distribuidoras*. Pesquisa Operacional [versão impressa ISSN 0101-7438 / versão online ISSN 1678-5142]. vol.29, n.2, pp. 303-322.

CEPEL – *Centro de Pesquisa da Eletrobrás*. Disponível em: www.cepel.br. Acesso em 29 de abril de 2010.

Harvey, A. C. *Time Series Models*. Philip Allan. London, 1981.

Hipel, K. W., McLeod, A. I. *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier. Amsterdam, The Netherlands, 1994.

Kirchner, R. M., Souza, R. C. e Ziegelmann, Flávio A. (2008), *Identificação de estruturas não-lineares de séries temporais através de regressão linear local e modelos aditivos*. Pesquisa Operacional [versão impressa ISSN 0101-7438 / versão online ISSN 1678-5142]. vol.28, n.1, pp. 45-57.

Maceira, M. E. P. *Operação Ótima de Reservatórios com Previsão de Afluências*. Tese de M.Sc., COPPE/UFRJ, Rio de Janeiro, Brasil. 1989.

Maceira, M. E. P., Damázio, J. M. (2004**),** *The use of PAR(p) model in the stochastic dual dynamic programming optimization scheme used in the operation planning of the Brazilian hydropower system. 8th International Conference on Probabilistic Methods Applied to Power Systems*, Iowa State University, Ames, Iowa, September 12-16.

Morettin, P. A., Toloi, C. M. C. *Previsão de Séries Temporais*. Atual Editora, São Paulo, 1987.

Oliveira, F. L. C. *Nova abordagem para geração de cenários de afluências no planejamento da operação energética de médio prazo*. Dissertação de Mestrado Departamento de Engenharia Elétrica, PUC-Rio, 2010.

Salas, J. D., Boes, D. C., Smith, R. A. (1982), *Estimation of ARMA models with seasonal parameters*. Water Resour. Res. 18(4), 1006-10.

Souza, R. C., Camargo, M. E. (2004), Análise de Séries Temporais: os modelos ARIMA. Ijuí: SEDIGRAF.

Stedinger, J. R. (2001), Report on the evaluation of CEPEL's PAR models. *Cornell University. School of Civil and Environmental Engineering*. Ithaca, New York.

Trevisan, E. S.; Souza, L. R.; Souza, R. C. *Estimação do parâmetro "d" em modelos ARFIMA.* Pesquisa Operacional, v.20, n. 1 (jun), pp. 73-82, 2000.

Wheelwright, S. C., Makridakis, S. *Forecasting methods for management*. 4th edition, New York, John Wiley & Sons Inc, 1989.