

JANILSON ALVES DA SILVA

**Estimativa de crescimento em altura de
Leucena [*Leucaena leucocephala* (Lam.) de Wit.] por meio
do Modelo ARIMA**

RECIFE-PE - MAR/2008



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

Estimativa de crescimento em altura de Leucena
[*Leucaena leucocephala* (Lam.) de Wit.] por meio do
Modelo ARIMA

Dissertação apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Mestre.

Área de Concentração: Modelagem Estatística e Computacional

Orientador: Prof. Dr. José Antônio Aleixo da Silva de Oliveira

Co-orientador: Professora Dra Viviane Moraes de Oliveira

Professor Dr Rinaldo Luiz Caraciolo Ferreira

RECIFE-PE - MAR/2008.

Ficha catalográfica

S586e Silva, Janilson Alves da
Estimativa de crescimento em altura de Leucena [*Leucaena leucocephala* (Lam.) de Wit.] por meio do modelo ARIMA / Janilson Alves da Silva. – 2008.
60 f. : il.

Orientador: José Antônio Aleixo da Silva.
Dissertação (Mestrado em Biometria e Estatística Aplicada) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em Biometria e Estatística Aplicada.
Inclui referências

1. Leucena 2. Análise de crescimento 3. Séries temporais
I. Silva, José Antônio Aleixo da II. Título

CDD 574.018

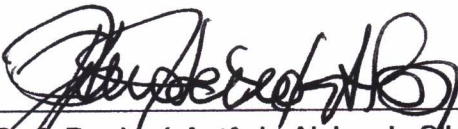
UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**Estimativa de crescimento em altura de Leucena [*Leucaena leucocephala* (Lam.) de
Wit.] por meio do Modelo ARIMA**

Janilson Alves da Silva

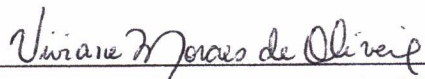
Dissertação julgada adequada para obtenção
do título de mestre em Biometria e Estatística
Aplicada, defendida e aprovada por unanimi-
dade em 31/03/2008 pela Comissão Exami-
nadora.

Orientador:




Prof. Dr. José Antônio Aleixo da Silva
Universidade Federal Rural de Pernambuco


Banca Examinadora:



Profa. Dra. Viviane Moraes de Oliveira
Universidade UFRPE



Prof. Dr. Gauss Moutinho Cordeiro
Universidade UFRPE



Prof. Dr. Tadeu Jankovski
Universidade UFRPE

Dedico esse trabalho

À minha esposa Dorai Queiroz

Agradecimentos

Antes de tudo preciso dizer que meus agradecimentos não são formais, assim como eu. Quero agradecer a todas as pessoas que se fizeram presentes, que se preocuparam, que foram solidárias, que torceram por mim. No entanto, sei que agradecer é sempre difícil e posso cometer injustiça esquecendo pessoas que me ajudaram do que fazer jus a todas que merecem. Também, caso o resultado daquilo pelo qual se agradece seja ruim, posso passar a impressão de distribuir a culpa entre todos aqueles mencionados neste item e minha intenção não jaz em nenhum desses dois pontos. Enfim todos que realizam um trabalho de pesquisa sabem que não o fazem sozinhos, embora seja solitário o ato de leitura e o do escrever, sinto que este trabalho não é só meu. Pelos autores que li, pelos professores com quem tive aulas na graduação e pós-graduação, pelos colegas de mestrado que me fizeram aprender com as discussões e conversas, bem como pelos comentários e sugestões feitos aos meus primeiros rabiscos da dissertação.

Ao meu bom Deus, e a São Francisco de Assis por estarem sempre presentes nas minhas orações e espiritualmente em todos os meus momentos de aflição e dificuldades em toda minha vida.

À Universidade Federal Rural de Pernambuco, em especial ao Departamento de Ciência Florestal, por ter dado todas as condições necessárias para a realização do curso. À Coordenação do Programa de Pós - Graduação em Biometria e Estatística Aplicada por ter dado todas as condições necessárias para o desenvolvimento do curso.

Ao Professor Dr. José Antônio Aleixo da Silva, pela dedicação concedida durante a realização da dissertação, pelas sugestões, ensinamentos, confiança e amizade, pois, orientador não é um chefe, é um amigo, em que você confia e que está sempre pronto para ajudar nos momentos difíceis. Ao Professor Dr. Rinaldo Luiz Caraciolo Ferreira, por toda ajuda e orientações possíveis e impossíveis, como co-orientador, pela grande amizade concedida durante todo o curso. À Professora Viviane Moraes de Oliveira pelas aulas, pelas sugestões, pelos conselhos e dicas informais, bem como o apoio dispensado durante todo o processo até a defesa. Ao professor Dr. Borko Stosic pelos conselhos, colaborações, análise, ensinamentos e incentivo a este trabalho, principalmente em sua fase exaustiva de conclusão. Ao professor Dr. Tadeu Jankovski , pelo apoio solidário, com quem compartil-

hei mais de perto, dúvidas, angústias e conhecimentos. Ao professor Dr. Gauss Moutinho Cordeiro pelas dicas de pesquisa e pelas horas de leituras gastas no meu trabalho. Muitas vezes até, indicando material que eu desconhecia e fazendo perguntas sobre meu trabalho. Suas sugestões nunca soaram arrogância de quem detém o título de doutor, mas foram sempre úteis, sempre bem-vindas e suas orientações acabaram por nortear e constituir-se na etapa final deste trabalho. Gauss, muito obrigado. A todos os Professores do Programa de Pós-graduação em Biometria que diretamente ou indiretamente contribuíram para o meu sucesso. Ao meu pai, Agamenon Alves, e minha mãe Neuza Eulália (In memorian), que me deram não somente a vida; agradeço por tudo e sempre.

Às minhas irmãs Simone Alves e Patrícia Alves por acreditarem na minha capacidade e compartilharem comigo, dificuldades e conquistas.

A todos os amigos que me ajudaram direta ou indiretamente e em especial aos meus amigos Adriano Regis, Frank Valdomiro e Ronaldo Alves.

À Mônica Pedrosa, pelo seu incentivo abraçando minha causa em momentos diversos e de forma solidária, contribuindo de maneira significativa para a conclusão deste trabalho.

Arquimedes
"Dê-me um ponto de apoio e moverei o mundo"

Resumo

O principal objetivo deste trabalho é utilizar modelos ARIMA para o ajuste das estimativas de crescimento em altura de leucena (*Leucaena leucocephala* (Lam.) de Wit.), no Agreste de Pernambuco. O experimento foi conduzido na Estação Experimental da Empresa Pernambucana de Pesquisa Agropecuária - IPA, no município de Caruaru - PE. Foram utilizadas 544 árvores de leucena, da variedade Hawaii (cv. K8), divididas em 24 tratamentos com 24 repetições. As fontes de variações estudadas foram: níveis de adubação fosfatada, composto orgânico de resíduo urbano e inoculação de rizóbio (NFB 466 e 473) aplicadas isoladamente. Foram consideradas para esse estudo 5 anos de medições e utilizado o Modelo de Chapman-Richards para remover a tendência da série em estudo, após remoção da tendência a nova série S_t foi modelada utilizando modelos ARIMA (1,1,0);(1,1,1)(1,1,2) e (1,1,3). Entretanto, os resultados não foram superiores aos dos modelos não-lineares tradicionais, frequentemente usados na modelagem do crescimento de florestas.

Palavras-chave: Análise de Crescimento, Leucena, Séries Temporais

Abstract

The main objective of this work is to use models ARIMA to adjust the estimates of growth in height of leucaena (*Leucaena leucocephala* (Lam.) de Wit.) Agreste of Pernambuco. The experiment was conducted at the Experimental Station Company Research Pernambuco Agropecuária - IPA, the municipality of Caruaru - PE. 544 trees were used to leucaena variety, of Hawaii (cv. K8), divided into 24 treatments with 24 repetitions. The sources of variations were: levels of phosphorus, organic compounds and urban waste inoculation with rhizobia (NFB 466 and 473) applied alone. Were considered for this study 5 years of measurements and used the Chapman-Richards model to remove the trend in the series study, after removal of the new trend set $S(t)$ was modeled using models ARIMA (1,1,0), (1,1,1) (1,1,2) and (1,1,3). However, the results were not superior to traditional non-linear models, often used in modeling the growth of forests.

Keywords: Analysis of Growth, Leucaena, Series Time

Lista de Figuras

1	Processo estocástico interpretado como uma família de variáveis aleatórias	p. 13
2	Altura da leucena	p. 35
3	FAC da série original	p. 35
4	Erros provenientes do Modelo de Chapman-Richards	p. 36
5	Função de autocorrelação parcial da Série S_t	p. 36
6	Função de autocorrelação da série S_t	p. 37
7	FAC - resíduos ARIMA(1,1,0)	p. 38
8	FACP - resíduos ARIMA(1,1,0)	p. 38
9	FAC - resíduos ARIMA(1,1,1)	p. 39
10	FACP - resíduos ARIMA(1,1,1)	p. 39
11	FAC - resíduos ARIMA(1,1,2)	p. 40
12	FACP - resíduos ARIMA(1,1,2)	p. 40
13	FAC - resíduos ARIMA(1,1,3)	p. 41
14	FACP - resíduos ARIMA(1,1,3)	p. 41

Lista de Tabelas

1	Altura da leucena nos 60 meses de medição	p. 31
2	Modelos não-lineares utilizados para remover a tendência	p. 33
3	Modelos não-lineares e as medidas de precisões	p. 35
4	Modelo ARIMA(1,1,0) ajustado à série S_t	p. 37
5	Teste de χ^2 de Box-Pierce para resíduos do modelo ARIMA(1,1,0)	p. 38
6	Modelo ARIMA(1,1,1) ajustado à série S_t	p. 39
7	Teste de χ^2 de Box-Pierce para resíduos do modelo ARIMA(1,1,1)	p. 39
8	Modelo ARIMA(1,1,2) ajustado à série S_t	p. 40
9	Teste de χ^2 de Box-Pierce para resíduos do modelo ARIMA(1,1,2)	p. 40
10	Modelo ARIMA(1,1,3) ajustado à série S_t	p. 41
11	Teste de χ^2 de Box-Pierce para resíduos do modelo ARIMA(1,1,3)	p. 42
12	Erros Quadráticos Médios dos Modelos ARIMA	p. 42
13	Valores do AIC e BIC dos modelos ARIMA	p. 43
14	Previsões da série S_t à 12 passos á frente	p. 43

Sumário

1	Introdução	p. 1
2	Revisão de Literatura	p. 3
2.1	Leucena	p. 3
2.2	Conceitos de crescimento florestal	p. 5
2.3	Modelagem matemática	p. 6
2.4	Curvas de Crescimento	p. 9
2.5	Modelos de crescimento em florestas	p. 10
2.5.1	Modelo de Von Bertalanffy	p. 10
2.5.2	Modelo de Richard's	p. 11
2.5.3	Modelo de Silva-Bayley	p. 11
2.6	Estacionariedade	p. 12
2.7	Processos estocásticos	p. 12
2.8	Processos estacionários	p. 13
2.9	Séries temporais	p. 13
2.10	Funções de Autocovariância e Autocorrelação	p. 14
2.11	Modelos Auto-Regressivos - AR(p)	p. 17
2.12	Modelos de Médias Móveis - MA(q)	p. 19
2.13	Etapas da Metodologia de Box-Jenkins	p. 19
2.13.1	Identificação	p. 20
2.13.2	Estimação	p. 21
2.13.3	Verificação	p. 26

2.13.4	Teste de autocorrelação residual	p. 26
2.13.5	Teste de Box-Pierce	p. 27
2.13.6	Previsão	p. 28
3	Material e Métodos	p. 31
3.0.1	Caracterização do experimento	p. 31
4	Resultados e Discussão	p. 34
4.0.1	Modelo Arima(1,1,0)	p. 37
4.0.2	Modelo Arima(1,1,1)	p. 38
4.0.3	Modelo Arima(1,1,2)	p. 39
4.0.4	Modelo Arima(1,1,3)	p. 41
5	Conclusões	p. 44
	Referências	p. 45
	Apêndice	p. 49
	Anexo	p. 50

1 Introdução

No Nordeste brasileiro, como na maioria das regiões tropicais em desenvolvimento, a utilização principal das florestas nativas tem se dado como lenha, carvão e madeira para serrarias e construções

A leucena [*Leucaena leucocephala* (Lam.) de Wit.] é uma leguminosa, perene, arbórea, originária da América Central e, atualmente, plantada em toda região tropical devido as suas múltiplas formas de utilização (forragem, produção de madeira, carvão vegetal, melhoramento do solo, sombreamento, quebra-vento e cerca - viva)(CAMPOS, 1980).

A leucena se desenvolve numa faixa de precipitação entre 600 e 3000 mm, prefere insolação direta perdendo as folhas quando sombreadas. Não se desenvolve bem em solos ácidos com altos teores de alumínio e deficientes em cálcio, magnésio, molibdênio e zinco (CAMPOS, 1980).

No Agreste de Pernambuco, sua importância é decorrente da grande necessidade de se buscar alternativas para a região com a finalidade de minimizar os problemas socioeconômicos da população através da comercialização de produtos advindos da leucena (carvão vegetal, estaca), alimentação animal (forragem). Além disso, o fato de não se ter conhecimento científico sobre o desenvolvimento da espécie nas condições específicas da região, com relação ao clima e ao solo, são necessárias pesquisas sobre fatores fisiológicos (crescimento) bem como a sua interação com o meio ambiente (sítio), resultando num prognóstico de produção viável.

Uma ferramenta muito utilizada em Ciências Florestais é a modelagem do crescimento/mortalidade pelo fato de que permite ao pesquisador planejar técnicas de manejo florestal, desbastes, produtividade, quantificação de sítios florestais.

A teoria da modelagem continua envolvente e aumentando em sofisticação na medida em que, novas técnicas de estatísticas e avanços na área da computação vêm sendo incorporadas aos problemas de crescimento e produção (BURKHART, 1987).

Neste contexto, é de grande importância a utilização de sistemas de predição do crescimento e da produção. Dentre as possibilidades pra modelar o crescimento e a produção dos povoamentos florestais se destacam os modelos mecanicistas e os biométricos. As duas linhas de abordagem são complementares, e não conflitivas.

Uma técnica que é considerada em estudos em que se tem informações ao longo do tempo em períodos eqüidistantes é a de séries temporais que tem sido pouco usada em Ciências Florestais.

Entre os procedimentos usados em séries temporais, destacam-se os modelos ARIMA (modelos autoregressivos integrados de médias móveis) que permitem predições das mais variadas ao longo do tempo, podendo ser aplicado em estudos de crescimento/mortalidade

Sendo assim, o principal objetivo deste estudo é modelar a série em estudo, através de modelos ARIMA, eliminando a tendência, através de modelos não-lineares.

2 Revisão de Literatura

2.1 Leucena

A *Leucaena leucocephala* (Lam.) de Wit, também conhecida por leucena ou esponjeira, tem despertado grande interesse científico devido a sua grande versatilidade dentre as leguminosas de clima tropical. Leucena é a mais comum dentre as espécies de plantas do gênero *Leucaena* (VIEIRA, 1992).

Salermo e Seiffert (1986) apresentaram o gênero *Leucaena* como pertencente a família Leguminosae, subfamília Mimosoidae e tribo Euminosae.

Embora não se saiba precisar o local de sua origem, supõe-se que seu mais antigo centro de dispersão tenha sido a península de Yucatan, no México, distribuindo-se, naturalmente na costa do Oceano Pacífico, mais amplamente, na faixa compreendida entre Peru e o Texas.

A leucena apresenta folhas bipinadas, de 15 a 20 cm de comprimento, ráquis pubescentes, tendo de 4 a 10 pares de pinas, de 5 a 10 cm de comprimento, cada pina pode apresentar de 5 a 20 pares de folíolos oblongo-lineares, agudos e inequiláteros, de 7 a 15 mm de comprimento por 3 a 4 mm de largura (SANTOS, 1991).

As flores são brancas e numerosas, possuindo cada uma 5 sépalas, 5 pétalas e 10 estames, formando inflorescências arredondadas, e chegam a medir de 2 a 2,5 cm de diâmetro. De uma maneira geral é de autopolinização, o que resulta em cachos de vagens.

As vagens são finas, achatadas, em formas de tiras, acuminadas, medindo de 1,4 a 2,0 cm de largura, geralmente agrupadas em 15 a 60 vagens, cada uma com 15 a 25 sementes, as quais possuem formas elípticas, e com, aproximadamente, 2 mm de espessura.

As sementes possuem ainda, uma película cerosa e resistente, que impede a sua germinação imediata, mesmo em solos úmidos e férteis (SANTOS, 1991).

A leucena é uma leguminosa perene, arbórea e com sistema radicular profundo, que lhe

permite a obtenção de umidade das camadas mais profundas do solo, mantendo-a verde durante as estações secas, só perdendo os folíolos em secas prolongadas ou em geadas muito fortes (FINGER, 1992).

A leucena é adequada para o cultivo em área com precipitação pluviométrica entre 600 mm e 1700 mm, embora na Indonésia ela venha sendo cultivada ao longo dos anos sob precipitação superior a 3900 mm anuais (SALERMO e SEIFFERT, 1986).

Em áreas com precipitação pluviométrica média anual de apenas 250 mm, a leucena pode se apresentar como a espécie dominante (FRANCO e SOUTO, 1986).

Apesar de sua resistência à estiagem, em período de seca prolongado ela pode perder suas folhas (SALERMO e SEIFFERT, 1986). O melhor crescimento da leucena, é obtido em áreas onde chove de 600 a 1.700 mm/ano, sendo uma espécie que suporta grandes diferenças de regimes de precipitação, luminosidade, salinidade do solo, fogo, geada leve, e épocas curtas de estiagem (FINGER, 1992).

Segundo Brewbaker (1978), a leucena é raramente encontrada em altitudes acima de 2000 metros. No Hawaii, desenvolve-se bem a uma altitude de 150 m em regiões chuvosas, e até 300 m nos locais mais secos (VILELA e PEREIRA, 1976).

Na Indonésia, encontram-se populações entre 500 m e 1.200 m de altitude (OAKES, 1968). No Ceilão florescem em altitudes de até 1.200 m, e apresentam um desenvolvimento insatisfatório em altitudes acima de 1.300 m.

A temperatura média anual ideal para o cultivo gira em torno de 22° e 30° C, abaixo deste intervalo ela apresenta rendimento insatisfatório, em temperaturas inferiores a 15,5° C, a planta paralisa seu crescimento. Em solos ácidos, latossolos, com alto teor de alumínio, deficiência de cálcio, molibdênio e zinco, a leucena não se desenvolve, e permanece arbustiva, sendo nesse caso, necessário a aplicação de calcário e fosfato (OLIVEIRA, 1995).

Em relação ao pH, Salermo e Seiffert (1986) afirmam que a leucena cresce melhor em solos com pH próximo ao neutro, e que a nodulação e o crescimento são afetados quando o PH se situa abaixo de 5,5. Segundo Hutton (1984), a leucena é uma leguminosa bastante exigente em relação à adubação com fósforo, molibdênio e zinco, em solos tropicais, o que a torna, na presença destes elementos, uma planta de raízes profundas.

Grande parte dos pecuaristas tem se preocupado com o alto custo dos alimentos concentrados, o que faz gerar um grande interesse por fontes alternativas de alimentação. Dentre essas fontes, destaca-se o uso de leguminosas, não só para alimentação humana e animal, como também para reflorestamento e adubação verde, uma vez que esta alternativa

vem minimizar o uso de adubo nitrogenado, que pode ser substituído pela fixação biológica do nitrogênio (SOUZA, 1991).

A leucena apresenta alto teor de proteína bruta, o que ressalta sua importância como forrageira, possuindo uma excelente palatabilidade e alto valor nutricional, sendo apresentada como fonte potencial de alimento para o gado (HILL, 1971).

O valor dessa leguminosa não está apenas no complemento alimentar, mas também no seu grande número de variedades com diferentes tamanhos e hábitos de crescimento, o que proporciona uma ampla diversificação de usos. Dentre eles, podemos citar: produção de madeira para lenha (MENDES, 1985); construção civil, postes, dormentes, carvão, mourões (CUNHA, 1979); popa, celulose, tabuas, taco, parque, chapa de partículas (BREWBAKER, 1989); produção de forragens para animais (KLUTHCOUSKI, 1982); sombreamento de culturas, quebra-vento, proteção contra o vento (FREITAS, 1991).

2.2 Conceitos de crescimento florestal

Os estudos de crescimento e produção de florestas iniciaram na Europa, no final do século XVIII (SPURR, 1952).

É comum se encontrar na literatura pertinente a crescimento e produção, muitos conceitos que assemelham entre si. Para Husch et al. (1982) o crescimento de uma árvore ou de um povoamento é o mais importante fenômeno na floresta, consistindo no alongamento e engrossamento das raízes, troncos e galhos, causando mudanças na árvore, influenciando o seu peso, volume e forma (SCOLFORO, 1993).

A palavra crescimento recorre a um aspecto de mudanças em um sistema orgânico que representa em um aumento em tamanho, e que pode ser medido em comprimento, diâmetro ou peso (MACHADO, 1978).

O crescimento linear de todas as partes de uma árvore, é resultante das atividades do meristema primário, enquanto o crescimento radial, é resultante das atividades do meristema secundário ou câmbio, os quais produzem uma nova camada de tecidos entre a madeira e a casca antiga (SPURR, 1952).

O crescimento é influenciado pela capacidade genética das espécies e interação com o ambiente, que incluem fatores climáticos (temperatura, precipitação, vento e insolação); fatores edáficos (características físicas, químicas e microorganismo); características topográficas (declividade, elevação, aspecto) e competição (influência de outras árvores,

vegetação rasteira e animais)(OLIVEIRA, 1995).

As medidas de crescimento em uma floresta mais comumente utilizadas são: altura total e comercial, diâmetro à altura do peito (DAP) e diâmetro em vários pontos do fuste (OLIVEIRA, 1995).

2.3 Modelagem matemática

Modelo é uma formulação matemática baseada em hipóteses que tentam representar fenômenos físicos ou biológicos, com a finalidade de gerar uma equação que possa estimar quantitativamente tal(is) fenômeno(s) a um determinado nível de probabilidade α .

Um modelo matemático é uma representação ou interpretação simplificada da realidade, ou uma interpretação de um fragmento de um sistema, segundo uma estrutura de conceitos experimentais.

Um modelo apresenta apenas uma visão ou cenário de um fragmento do todo. Normalmente, para estudar um determinado fenômeno complexo, criam-se vários modelos.

Os modelos matemáticos são utilizados praticamente em todas as áreas científicas, como, por exemplo, na biologia, química, física, economia, engenharia e na própria matemática pura.

Entretanto, dependendo da área de atuação, um mesmo tipo de modelo pode ter diferentes nomes.

Segundo Ferreira (2004), modelo é uma representação informal baseado em descrições formais de objetos, e que permite, por meio de variações, simular os efeitos de mudanças do fenômeno que representa. O mesmo autor afirma que a modelagem conceitual é a etapa da análise de um "sistema", na qual são definidos os recursos, itens de dados e suas inter-relações.

Modelo é um "sistema" que pode ser definido como um conjunto de elementos que se interrelacionam (BERTALANFFY, 1969)

Interrelação significa que alguns elementos (p) mantêm relações com R , de forma tal que o comportamento de um elemento p é diferente de seu comportamento em outra relação R . Dependendo da natureza dos dados, existem vários tipos de modelos:

Modelo físico: conjunto de hipóteses sobre a estrutura ou comportamento de um sistema físico pelo qual se procuram explicar ou prever, dentro de uma teoria científica as pro-

priedades de um sistema (FERREIRA, 2004).

Modelo dinâmico: é aquele que utilizam equações diferenciais que podem ser resolvidas de forma analítica ou numérica, sendo que as estimativas e/ou simulações são realizadas em função do tempo (SPATHELF e NUTTO, 2000).

Modelo estático: não incorpora a variável tempo, pois é representativo de uma situação em um determinado momento (SPATHELF e NUTTO, 2000).

Os modelos empíricos ou de prognose não consideram hipóteses associadas à casualidade. No pressuposto dessa abordagem, é que se encontram padrões comuns no crescimento de indivíduos. Nesse tipo de modelo a floresta é vista como um sistema para a produção (SPATHELF e NUTTO, 2000).

Um modelo empírico tende, principalmente, a descrever um fenômeno, enquanto que um modelo mecanicístico trata de dar uma descrição com compreensão de um determinado fenômeno (FRANCE e THORNLEY, 1984).

Os modelos mecanicísticos têm uma hipótese associada ao fenômeno descrito e servem para o entendimento de processos.

Nesse tipo de modelo a floresta é vista como um sistema com ligações complexas entre os seus elementos, introduzindo mais funcionalidade na representação do crescimento e competição entre as populações, embora seu uso esclareça mais sobre a competição entre as espécies, por outro lado, são mais difíceis suas aplicações (SPATHELF e NUTTO, 2000).

Tais modelos tentam representar, indiretamente, o efeito do ambiente e das práticas silviculturais no desenvolvimento das árvores de um povoamento florestal, utilizando fontes de variação como o sítio, a área basal e a idade.

Modelos com propriedades biológicas e amparados numa amostragem adequada, com informações precisas e remedições, propiciam prognoses do crescimento e da produção florestal com elevado grau de confiabilidade. A partir deles se pode prescrever regimes de manejo adequados para cada espécie, em cada sítio, que visem à qualidade do produto final como desbastes, rotação econômica ótima e planejamento da colheita (SPATHELF e NUTTO, 2000).

Modelos baseados em processos constituem uma ciência em desenvolvimento crescente vinculada à fisiologia vegetal. Esse ramo de conhecimento tem evoluído, consideravelmente, nos estudos de interação planta x solo x atmosfera, conferindo uma visão mais generalista e, fisiologicamente, embasada para estimativa de crescimento (SPATHELF e

NUTTO, 2000).

O modelo em sua forma analítica é composto de variáveis (dependentes e independentes), parâmetros e um erro associado, quando se trata de modelo probabilístico, tem-se por exemplo,

$$V_i = \beta_0 + \beta_1 DAP_i + \beta_2 H_i + \varepsilon_i,$$

em que V_i é o volume da árvore i , DAP_i é o diâmetro a altura do peito da árvore i , H_i altura total da árvore i , β_0 e β_1 são os parâmetros do modelo, ε_i é o erro aleatório.

O erro aleatório ε confere ao modelo a natureza probabilística, isto é, esse é influenciado por situações não controláveis, que quanto mais ocorrerem menor será a precisão do modelo.

Os modelos matemáticos somente representam um conjunto de suposições físicas ou biológicas. Essas suposições são sempre simplificações da realidade (FRANCE e THORNLEY, 1984).

Na maioria das vezes esses modelos denominados de matemáticos não se preocupam com a natureza do fenômeno, buscam apenas conseguir bons ajustes.

Nas ciências, os modelos matemáticos são muito importantes e usam uma ferramenta especial: as equações, sendo que um modelo de crescimento pode abranger um sistema de equações, sendo capaz de prognosticar o crescimento e a produção sob várias condições (SPATHELF e NUTTO, 2000).

Existem também os modelos determinísticos que são aqueles em que o erro aleatório não existe, pois caso o fenômeno se repita n vezes o resultado será sempre o mesmo, tem-se como exemplo a seguinte expressão,

$$V_f = V_i(1 + i)^t,$$

em que, V_f é o valor final, V_i é o valor inicial, i é a taxa de juros, t é o período de aplicação.

A análise de regressão tem sido utilizada com muita frequência no campo da modelagem florestal, principalmente, porque permite se fazer estimativas de parâmetros da floresta com baixos custos, em curto espaço de tempo e precisas, desde que o conjunto de dados que gerou tal equação represente a população em consideração (SCHNEIDER, 1998).

Uma regressão quando utiliza apenas uma variável independente para explicar um fenômeno é denominada de regressão simples podendo ser linear ou não linear em função da natureza dos parâmetros no modelo. É considerada múltipla quando a variável dependente é estimada (explicada) por duas ou mais variáveis independentes.

As técnicas de regressão linear múltipla são muito úteis, mas também muito perigosas se usadas e interpretadas de forma imprópria.(PRODAN, 1968).

Existem vários modelos de crescimentos que se adaptam a diferentes situações. Há modelos lineares e não-lineares. Em muitas situações, mais de um modelo se adapta ao que se quer modelar; então o problema passa a ser a escolha do melhor modelo, o que nem sempre é tarefa fácil. Muitos modelos já foram desenvolvidos com essa finalidade, geralmente, levando em consideração dois critérios gerais de avaliação: 1) o ajustamento da função aos dados; 2) e o afastamento dos dados com relação ao modelo. A forma mais simples de escolha de modelos baseia-se, portanto, no uso do coeficiente de determinação R^2 da regressão, que estima o ajustamento, e no quadrado médio dos resíduos $QMRes$ que estima o afastamento (SIT, 1994).

2.4 Curvas de Crescimento

O desenvolvimento das árvores depende de duas variáveis principais: o tempo e o ambiente. Sendo os fatores ambientais favoráveis, as árvores podem apresentar, em determinada época, o máximo rendimento, ficando implícitas, suas potencialidades genéticas.

A Biometria avalia esse desenvolvimento pela relação crescimento/tempo, considerando os fatores ambientais e genéticos constantes.

Sob este ponto de vista é possível construir os fundamentos matemáticos para estimar o crescimento das árvores em determinado sítio. Quando se plota uma dimensão de uma árvore, como V, DAP, H, etc, em relação a idade, constrói-se uma curva denominada curva de crescimento, que é caracterizada pela forma de sigmóide ou S e mostra o tamanho acumulado até uma certa idade. Essa curva, tem a mesma conformação tanto para células individuais, como tecidos, órgãos, animais ou vegetais durante toda a vida, como também é o padrão para pequenos períodos como uma estação de crescimento.

De acordo com a variável que estiver sendo considerada ocorrerão mudanças na forma exata da curva, permanecendo entretanto, para qualquer caso, as suas características básicas. A curva de crescimento mostra o tamanho alcançado em cada idade, começando no

ponto zero e subindo primeiramente de forma paulatina e depois mais íngrememente até um ponto de inflexão. Após este ponto, a curva muda de concavidade, a inclinação diminui e aproxima-se assintoticamente do valor final.

2.5 Modelos de crescimento em florestas

Nas Ciências Florestais, tais como Biologia, Sociologia e Economia, um grande número de tentativas tem sido feito para exprimir o crescimento como fórmulas matemáticas (PRODAN, 1968).

Os estudos de crescimento e produção de florestas iniciaram-se na Europa, no final do século XVIII, muitos trabalhos têm sido desenvolvidos (SPURR, 1952).

Nas ciências biológicas, as funções de crescimento têm sido usadas por muitos anos, principalmente, para prover dados do crescimento de um organismo ou parte de um organismo em um determinado tempo. O termo função de crescimento é geralmente usado para denotar uma função analítica que pode ser descrita em uma equação simples (SANTOS, 2005).

Entre os modelos usados em Ciências Florestais se destacam os seguintes.

2.5.1 Modelo de Von Bertalanffy

Bertalanffy (1957) desenvolveu um modelo baseado no metabolismo dos seres vivos, considerando que a taxa anabólica é proporcional à área da superfície de um organismo enquanto a taxa catabólica é proporcional ao volume do corpo do organismo. Ele expressou esse fato em fórmula matemática. A forma conhecida desse modelo é expressa por

$$\frac{dW}{dt} = \eta W^m - KW^n,$$

em que W é o peso do corpo, t é o tempo, η é a constante do fator anabólico, K é a constante do fator catabólico, m e n são constantes.

Bertalanffy aplicou a expressão em uma experiência em mamíferos, e encontrou o valor $2/3$ para o m e geralmente o n é igual à unidade. Assim, generalizou o seu modelo como:

$$\frac{dW}{dt} = \eta W^{2/3} - KW.$$

2.5.2 Modelo de Richard's

Richards(1959) e Chapman(1961) aplicaram o modelo de Bertalanffy em estudos com cogumelos e peixes, respectivamente e encontraram que o expoente 2/3 é muito restritivo. Eles passaram a usar o modelo da seguinte forma:

$$H = A(1 - \beta \exp(-kt))^{\frac{1}{1-m}},$$

em que H é a altura da variável na idade t, A é o valor da assíntota e β , k, m são parâmetros. Este modelo apresenta a vantagem de permitir uma interpretação lógica e ser bastante flexível, adaptando-se bem ao desenvolvimento de diferentes variáveis dendrométricas.

Segundo o autor, o modelo transforma-se em diferentes modelos segundo a variação do coeficiente "m". Quando "m" for zero, o modelo chama-se Monomolecular, este modelo não tem ponto de inflexão e sua taxa de crescimento decresce linearmente com o aumento da variável dependente. Por esses motivos o Modelo Monomolecular é empregado principalmente para representar a porção da vida após o culminio do incremento corrente anual.

Quando "m" assume valor 1, na equação diferenciável de Von Bertalanffy o modelo transforma-se no Modelo Gompertz. Este modelo é assimétrico em relação ao ponto de inflexão e apresenta uma relação linear entre a taxa de crescimento relativa e o logaritmo da variável dependente e também entre o logaritmo da taxa de crescimento relativa e a idade. Esse modelo é mais apropriado para estudos de crescimento de animais.

Quando "m" for igual a 2, apresenta-se como a curva logística cuja característica é simétrica em relação ao ponto de inflexão, sendo que a taxa de crescimento relativa também decresce linearmente como o incremento de H (MACHADO, 1978).

2.5.3 Modelo de Silva-Bayley

Silva (1986) estudando o crescimento de *Pinus elliottii* Engelm var. *elliottii* em função de diferentes fontes de nutrientes, considerou a idéia de dinâmica de populações proposta por Leslie (1945) que diz que para qualquer distribuição arbitrária no tempo i, a estrutura de uma população pode ser expressa na forma de n+1 equações lineares, em que de n para n+1 é o ultimo grupo de idade na tabela de distribuição etária.

O modelo expressa o tamanho da população W_{i+1} no tempo $i+1$ como função do tamanho da população no tempo $i(W_i)$, da taxa de sobrevivência (S), do número médio de recruta-

mento (F) produzido no tempo i . A expressão é $W_{i+1} = W_i(SF)^n$.

Em populações florestais, o problema de recrutamento (nascimento ou migração) no período de investigação é desprezível, bem como migração que são fenômenos comuns em populações animais (móveis). Ocorre mortalidade, mas esta pode ser modelada sem maiores problemas. Dessa forma, Silva (1986) considerou o plantio de *Pinus* como uma população fechada e o parâmetro F pode ser retirado do modelo.

O parâmetro S é desconhecido e depende da taxa de mortalidade que pode ser regular ou irregular (BURKHART, 1987). Mortalidade regular é resultante de fatores tais como densidade de plantio, qualidade do sítio, tratamentos silviculturais, etc, enquanto que mortalidade irregular é consequência de eventos catastróficos como fogo, doenças, pragas, ventos, etc.

2.6 Estacionariedade

Uma das suposições a respeito de uma série temporal é a de que ela se desenvolve aleatoriamente no tempo em torno de uma média constante, demonstrando alguma forma de equilíbrio ou estabilidade (MORETTIN e TOLOI, 2004). Todavia, a maior parte das séries apresentam alguma forma de não-estacionariedade. Uma condição importante para avaliar o regime estacionário da série temporal y_t , é a rapidez da função de autocorrelação convergir para zero. Nas séries infinitamente longas, pode-se calcular as autocorrelações r_k para um número de defasagens $k \leq N/4$ e, nesse caso, a condição de estacionariedade é que a série $\sum_{k=1}^{\infty} \widehat{r}_k^2$ seja convergente. Sabe-se, que a condição de convergência da série está vinculada a queda rápida de \widehat{r}_k^2

2.7 Processos estocásticos

Definição 2.7.1. *Seja T um conjunto arbitrário. Um processo estocástico é uma família $\{y_t, t \in T\}$, tal que, para cada $t \in T$, y_t é uma variável aleatória.*

Nessas condições, um processo estocástico é uma família de variáveis aleatórias (*v.a.*), que se supõem definidas num mesmo espaço de probabilidades (Ω, A, P) . Supõe-se que as variáveis aleatórias envolvidas sejam reais (MORETTIN e TOLOI, 2004).

O conjunto T é normalmente considerado como o conjunto dos inteiros $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ ou o conjunto dos reais \mathbb{R} .

Como para cada $t \in T$, y_t é uma variável aleatória definida sobre Ω na realidade, y_t é uma função de dois argumentos $y(t, \omega)$, $t \in T$, $\omega \in \Omega$.

A figura 1 ilustra essa interpretação de um processo estocástico. Observa-se na figura 1 que para cada $t \in T$, tem-se uma variável aleatória $y(t, \omega)$, com uma função densidade de probabilidade $f_Y(y)$.

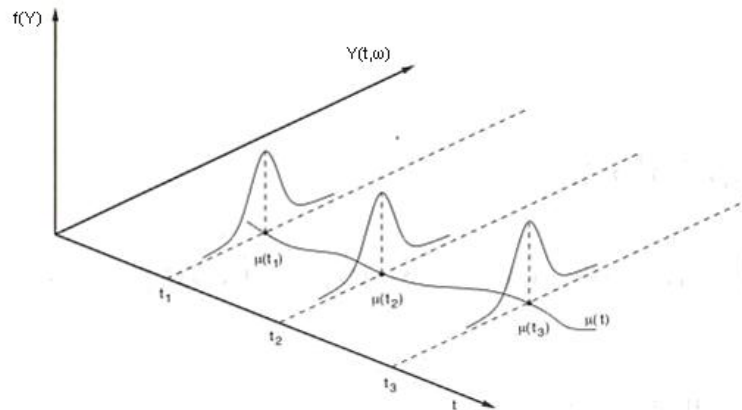


Figura 1: Processo estocástico interpretado como uma família de variáveis aleatórias

Por outro lado, para cada $\omega \in \Omega$, fixado, tem-se uma função de t , ou seja, uma realização ou trajetória do processo.

2.8 Processos estacionários

Nas situações em que se pretende utilizar modelos para descrever processos físicos, torna-se necessário introduzir suposições que permitam a obtenção de propriedades probabilísticas e estatísticas desejáveis (MORETTIN e TOLOI, 2004).

No caso de processos estocásticos, uma suposição normalmente feita é a de estacionariedade. Intuitivamente, um processo y_t é estacionário se ele se desenvolve no tempo, de modo que a escolha de uma origem dos tempos não seja importante. Em outras palavras, as características probabilísticas de $y_{t+\tau}$, para todo $\tau \in T$, são as mesmas de y_t .

2.9 Séries temporais

Informalmente, uma série temporal pode ser vista como um conjunto de observações estocásticas y_t , cada uma observada num instante particular t . Assim, contrariamente ao que se passa em relação a amostra aleatórias, em que o instante em que é efetuada a observação não desempenha nenhum papel especial e é, em geral, ignorado nas séries

temporais o instante em que é efetuada cada uma das observações necessita ser registrado e é usado na análise/modelagem da série (BOX e JENKINS, 1970).

As séries temporais são observadas em domínios variados, por exemplo: Agronomia (produções e áreas cultivadas anuais e produção média anual por hectare de um dado produto agrícola), Economia (rendimento anual per capita, vendas trimestrais de um dado produto), Engenharia (intensidade do som num determinado local, intensidade da corrente elétrica num dado ponto), Medicina (registros de eletroencefalogramas e eletrocardiogramas) e nas ciências sociais (taxas de mortalidade, taxas de acidentes de aviação, taxas de criminalidades).

Uma série temporal é uma coleção de observações realizadas seqüencialmente ao longo do tempo. A característica mais importante deste tipo de dados é que as observações vizinhas são dependentes e estamos interessados em analisar e modelar esta dependência. Enquanto em modelos de regressão, por exemplo, a ordem das observações é irrelevante para a análise, em séries temporais a ordem dos dados é crucial. Vale notar também que o tempo pode ser substituído por outra variável como espaço, profundidade, etc.

Séries de dados representando o crescimento das árvores ao longo do tempo podem ser obtidas com bom nível de precisão por parcelas permanentes, ou medição de anéis de crescimento (SCHNEIDER, 1998).

Amostras de altura em série temporal podem ser consideradas como amostra em pares de dados de altura e idade, que apresentam a vantagem de ter menor flutuação que amostras independentes (WONNACOTT e WONNACOTT, 1980).

2.10 Funções de Autocovariância e Autocorrelação

Em uma série temporal a função de autocorrelação é uma ferramenta importante, pois avalia se existe dependência entre as observações da série (CORDEIRO, 2002).

Seja $\{y_t, t \in T\}$ uma sequência tal que $Var(y_t) = \sigma^2 < \infty, \forall t \in T$. Então, a função de autocovariância é definida por

$$\gamma(t, t+h) = Cov(y_t, y_{t+h}) = E[(y_t - \mu_t)(y_{t+h} - \mu_{t+h})].$$

Quando se trata de um processo estacionário, tem-se:

- i) $\gamma_0 \geq 0$ e $Var(y_t) = \sigma_y^2$;
- ii) $|\gamma_h| \leq \gamma_0 \forall h$;
- iii) a função $\gamma(\cdot)$ é par, isto é, $\gamma_h = \gamma_{-h} \forall h$.

A autocovariância depende da unidade de medida de y . Contudo, é conveniente normalizar as autocovariâncias dividindo-se pelo produto dos respectivos desvios padrão, obtendo-se, assim, a função de autocorrelação (FAC).

A FAC de um processo estacionário, caso em que a variância é constante, é da forma

$$\rho_h = \frac{Cov(y_t, y_{t+h})}{\sigma_y^2} = \frac{\gamma_h}{\gamma_0}.$$

A função de autocorrelação de modo geral é dada por

$$\rho_h = \frac{Cov(y_t, y_{t+h})}{\sigma_t \cdot \sigma_{t+h}} = \frac{\gamma_h}{\sigma_t \cdot \sigma_{t+h}},$$

e tem as seguintes características básicas:

- i) $\rho_0 = 1$;
- ii) $|\rho_h| \leq 1 \forall h$;
- iii) a função $\rho(\cdot)$ é par, isto é $\rho_h = \rho_{-h} \forall h$.

Os valores da autocorrelação calculados com base numa mostra $\hat{\rho}_h$ é chamada de autocorrelação amostral. O gráfico dos valores de $\hat{\rho}_h$ versus os valores não-negativos de h é chamado de correlograma ou função de autocorrelação amostral.

Obtém-se, então, uma ferramenta útil para uma análise preliminar dos dados observados de uma série, pois o correlograma permite verificar se o processo é estacionário ou não, salvo alguns casos.

O correlograma de um processo estacionário decai bruscamente para "zero", ou melhor dizendo, para um intervalo em que admite-se ρ_h nulo.

A significância estatística de qualquer $\hat{\rho}_h$ pode ser julgada por seu erro-padrão. Bartlett (1946) mostrou que se uma série temporal for puramente aleatória, ou seja, se exibir ruído,

os coeficientes de autocorrelação amostral são, aproximadamente, distribuídos normalmente com média zero e variância $1/n$, em que n é o tamanho da amostra. Então, seguindo as propriedades da distribuição normal padrão, o intervalo de confiança de 95% para qualquer $\hat{\rho}_h$ será $\pm 1,96 \cdot \sqrt{1/n}$. Assim, se um $\hat{\rho}_k$ estimado se situar no intervalo $(-1,96/\sqrt{n}, +1,96/\sqrt{n})$, não rejeitamos a hipótese de que o verdadeiro ρ_k seja zero.

Mas se ele se encontrar fora desse intervalo de confiança, então podemos rejeitar a hipótese de que o verdadeiro ρ_k seja zero (GUJARATI, 2000).

A autocorrelação parcial pode ser considerada como uma correlação entre duas variáveis quaisquer y_{t-k} e y_t , separados por k períodos com os efeitos das variáveis intermediárias $y_{t-k+1}, y_{t-k+2}, \dots, y_{t-1}$ eliminadas, tendo-se, portanto

$$\text{Corr}(y_{t-k}, y_t | y_{t-k+1}, y_{t-k+2}, \dots, y_{t-1}),$$

que é a influência direta de y_{t-k} em y_t .

A função de autocorrelação parcial (FACP) é outra ferramenta que também contribui na escolha de um modelo adequado, sendo denotada por

$$\phi = \frac{|P_k^*|}{|P_k|},$$

em que P_k é a matriz de autocorrelações e P_k^* é a matriz P_k com a última coluna substituída pelo vetor de autocorrelações, Obtém-se, assim,

$$\phi_{11} = \rho_1,$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}},$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}.$$

2.11 Modelos Auto-Regressivos - AR(p)

De acordo com esse modelo, y_t é descrito apenas por seus valores passados e pelo ruído branco ϵ_t . A versão mais simples de um modelo AR é aquela em que y_t depende somente de y_{t-1} e de ϵ_t . Diz-se, nesse caso, que o modelo é auto-regressivo de ordem 1, o que se indica compactamente por AR(1). A representação algébrica desse modelo é

$$y_t = \phi y_{t-1} + \epsilon_t,$$

em que ϕ é um parâmetro e $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma^2$ e $E(\epsilon_t, \epsilon_s) = 0$ para $t \neq s$.

Por se tratar de um modelo fracamente estacionário, a variância de y_t (γ_0) deve ser constante e as autocovariâncias (γ_k) devem ser independentes de t .

- Variância do AR(1)

$$\gamma_0 = E[y_t^2] = E[\phi y_{t-1} + \epsilon_t]^2 = \phi^2 V(y_{t-1}) + V(\epsilon_t),$$

$$\gamma_0 = \phi^2 \gamma_0 + \sigma^2,$$

$$\gamma_0 = \frac{\sigma^2}{1 - \phi^2}.$$

Suponha $\{y_t, t \in T\}$ uma série temporal sem movimentos sazonais. Portanto, y_t pode ser escrita como função linear da variável imediatamente anterior mais um ruído branco, ou seja,

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t, \quad (2.1)$$

em que $u_t \sim RB(0, \sigma^2)$ e ϕ_1, \dots, ϕ_p e c são parâmetros desconhecidos, e ainda ϕ_1, \dots, ϕ_p são parâmetros auto-regressivos. É razoável também que (2.1) seja expresso como

$$c + u_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) y_t, \quad (2.2)$$

em que $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, ou ainda

$$\phi(B) y_t = c + u_t, \quad (2.3)$$

com $B y_t = y_{t-1}$.

Um processo AR(p) é estacionário quando todas as raízes da equação $1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0$, no plano complexo, encontram-se fora do círculo unitário. Com relação à invertibilidade, não existe restrições, todo processo AR(p) é invertível.

- Média

$$E(y_t) = E(c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t)$$

$$c = (1 - \phi_1 - \phi_2 - \dots - \phi_p) \mu$$

$$\mu = \frac{c}{1 - \phi_1 - \phi_2 - \dots - \phi_p}.$$

- Autocovariância

A Autocovariância de ordem j é expressa por

$$E[(y_t - \mu)(y_{t-j} - \mu)] = \phi_1 E[(y_{t-1} - \mu)(y_{t-j} - \mu)] + \dots + \\ + \phi_p E[(y_{t-p} - \mu)(y_{t-j} - \mu)] + E[u_t(y_{t-j} - \mu)].$$

Assim,

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \cdots + \phi_p \gamma_{j-p}, \quad j = 1, 2, \dots \quad (2.4)$$

- Autocorrelação

Dividindo (2.4) por γ_0 , obtém-se

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p}, \quad j = 1, 2, \dots$$

2.12 Modelos de Médias Móveis - MA(q)

Suponha $\{y_t, t \in T\}$ uma série temporal estacionária. Portanto, pode ser escrita como função linear do valor atual de um processo ruído branco mais o valor deste mesmo processo no instante imediatamente anterior multiplicado por uma constante, (GUJARATI, 2000), ou seja,

$$y_t = \mu + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q},$$

ou

$$y_t = \mu + \theta(B)u_t,$$

para

$$\theta(B) = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q \cdot B^q),$$

em que $u_t \sim RB(0, \sigma^2)$.

2.13 Etapas da Metodologia de Box-Jenkins

Os modelos lineares para séries temporais, introduzidos por Box e Jenkins (1976) têm importância fundamental na Econometria, uma vez que possibilitaram as primeiras análises e previsões do domínio do tempo.

Pode-se fazer uma analogia dos modelos lineares para séries temporais com modelos de regressão. Ambos assumiram a forte hipótese da linearidade na estrutura dos dados,

hipótese esta que na maioria dos casos práticos não é satisfeita, mas que após transformações adequadas permite uma modelagem adequada.

Recentemente, modelos mais gerais têm sido propostos, os quais explicam melhor o comportamento de muitas séries conhecidas, assim como permite realizar previsões mais confiáveis.

Quando se aplica a metodologia de Box-Jenkins a uma determinada série, o primeiro passo consiste em descobrir qual as várias versões do modelo ARIMA descreve o comportamento da série, esta etapa é denominada identificação. O segundo passo consiste em estimar os parâmetros do modelo identificado: os parâmetros ϕ se houver um componente autorregressivo, os parâmetros θ se houver o filtro de médias móveis e a variância do ruído branco σ_ε^2 , esta etapa é chamada de estimação. Segue-se a ela a etapa de verificação, que consiste em avaliar se o processo de estimação foi bem-sucedido. Se a conclusão for negativa, repete-se o procedimento desde a etapa de identificação com o objetivo de encontrar uma especificação de modelo mais adequada. Quando se obtém um modelo satisfatório, passa-se para a última etapa da metodologia, a previsão, que se constitui no objetivo primordial do enfoque de Box-Jenkins.

2.13.1 Identificação

Um dos maiores obstáculos à utilização da metodologia de Box-Jenkins, está na identificação. Vários pesquisadores, usando a mesma série, podem identificar modelos diferentes (MORETTIN e TOLOI, 2004).

Enfim, o processo de identificação consiste em determinar quais os filtros AR, I e MA que compõem o processo gerador da série, bem como quais são suas respectivas ordens.

A observação do gráfico da série não será de grande valia para distinguir um AR de um MA ou de um ARMA. O gráfico indicará, no máximo, se ela é ou não estacionária. A realização do processo de identificação necessita, portanto, de outros instrumentos. São eles a função de autocorrelação (FAC) e a função de autocorrelação parcial (FACP).

Após uma breve fundamentação teórica, é possível definir precisamente a etapa de identificação do modelo. Há várias estratégias que podem ser utilizadas. Em modelos AR(p), o gráfico das autocorrelações parciais amostrais deve apresentar um corte na ordem p. Já em processos MA(q), o corte deve ocorrer na ordem q no gráfico das autocorrelações amostrais. Contudo, sendo o processo misto a identificação torna-se difícil (MORETTIN e TOLOI, 2004).

2.13.2 Estimação

Após a etapa de identificação da ordem do modelo que melhor representa a série observada, torna-se necessária a obtenção de estimativas para os parâmetros do modelo identificado.

O processo de estimação do modelo geral ARIMA (p,d,q) conforme expressão,

$$\phi(B)\nabla^d y_t = \theta(B)\epsilon_t,$$

em que $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ e $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ e significa, portanto, a obtenção de "p" estimativas dos parâmetros $\phi_1, \phi_2, \dots, \phi_p$ e "q" estimativas para os parâmetros $\theta_1, \theta_2, \dots, \theta_q$, além da variância do ruído, σ_ϵ^2 .

Considere um modelo ARIMA(p,d,q) e um vetor $\xi = (\phi, \theta, \sigma_\epsilon^2)$, em que $\phi = (\phi_1, \dots, \phi_p)$ e $\theta = (\theta_1, \dots, \theta_q)$, onde serão colocados os $p + d + 1$ parâmetros do modelo. E considere, também, $\eta(\phi, \theta)$. Supondo aqui, quando $d > 0$ e $\mu_w = 0$, em caso contrário, o número de parâmetros será $p + q + 2$, pois μ_w será incluído como um parâmetro do modelo.

Considere-se as N observações y_1, \dots, y_N e a função de máxima verossimilhança $L(\xi|y_1, \dots, y_N)$ expressa como função de ξ . Os valores que maximizam L ou $l = \log L$ serão os estimadores de máxima verossimilhança (EMV) de ξ .

Trabalhando com a suposição de que o processo ϵ_t tem distribuição normal com média zero e variância σ_ϵ^2 para cada t, então os EMV serão, aproximadamente, estimadores de mínimos quadrados (EMQ).

Alcançada a estacionariedade, após d diferenças, obtém-se $n = N - d$ observações W_1, \dots, W_n , em que $W_t = \nabla^d y_t$. O modelo ARMA(p,q) obtido é estacionário e invertível, expresso por

$$\epsilon_t = \widetilde{W}_t - \phi_1 \widetilde{W}_{t-1} - \dots - \phi_p \widetilde{W}_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}, \quad (2.5)$$

em que $\widetilde{W}_t = W_t - \epsilon_w$.

a) Método dos Momentos

A idéia, neste método, é substituir os momentos teóricos pelos seus correspondentes momentos amostrais, nas equações em que estes se relacionam como os parâmetros do

modelo. Os estimadores dos parâmetros obtidos desta forma serão utilizados, na maioria das vezes, como estimativas iniciais em procedimentos iterativos de estimação não-linear.

b) Método de Máxima Verossimilhança

Embora os parâmetros auto-regressivos sejam lineares, permitindo o uso de métodos simples de estimação como, por exemplo, o de mínimos quadrados ordinários, os parâmetros média móvel não são lineares, o que torna bastante complexos os processos de estimação. Por essa razão a escolha da técnica utilizada recai no método de máxima verossimilhança.

Esse método é uma técnica de estimação bastante usual na inferência estatística, na medida em que as estatísticas assim obtidas apresentam muitas das propriedades desejáveis para a escolha dos estimadores.

O princípio de verossimilhança sustenta que, se o modelo está corretamente identificado, todas as informações que os dados observados fornecem sobre os parâmetros estão contidas na função de verossimilhança, e todos os demais aspectos informativos das observações são irrelevantes.

Para grandes amostras, as estimativas da máxima verossimilhança possuem as propriedades assintóticas tais como: eficiência com menor variância que qualquer outro estimador, consistência, caso não haja um estimador suficiente, ele será produzido pelo estimador de máxima verossimilhança, ser normalmente distribuído, com facilidade de cálculo para obtenção dos parâmetros média e variância.

A idéia do método de máxima verossimilhança é achar grandezas populacionais que geram os valores que mais se assemelham aos da amostra observada, ou seja, o método busca estabelecer os valores populacionais hipotéticos que maximizam a verossimilhança da amostra observada. Dito de outra maneira, o método consiste em selecionar aqueles estimadores que maximizam a probabilidade de se obter a mostra realmente observada.

Este método necessita de valores iniciais para os \tilde{W}'_s e para os ϵ'_s da expressão (2.5). Para tanto, são conhecidos dois procedimentos, sendo um condicional e outro não condicional.

i) Procedimento Condicional.

A função de densidade conjunta de $\epsilon_1, \dots, \epsilon_n$, sob a suposição de $\epsilon_t \sim N(0, \sigma_\epsilon^2)$, é expressa por

$$f(\epsilon_1, \dots, \epsilon_n) = (2\pi)^{-\frac{n}{2}} (\sigma_\epsilon)^{-n} \exp \left\{ - \sum_{t=1}^n \frac{\epsilon_t^2}{2\sigma_\epsilon^2} \right\}. \quad (2.6)$$

Serão denotados de W_t^* e ϵ_t^* os p valores supostos de W_t e os q valores, também, supostos de ϵ_t usados para calcular $\epsilon_1, \dots, \epsilon_n$.

A partir das equações (2.5) e (2.6), será obtida a função de verossimilhança condicional à escolha de W_t e ϵ_t . Tem-se,

$$L(\xi|W, W^*, \epsilon^*) = (2\pi)^{-\frac{n}{2}} (\sigma_\epsilon)^{-n} \\ \times \exp \left\{ - \frac{1}{2\sigma_\epsilon^2} \sum_{t=1}^n (\widetilde{W}_t - \phi_1 \widetilde{W}_{t-1} - \dots - \phi_p \widetilde{W}_{t-p} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-1})^2 \right\}.$$

Usando um asterisco para denotar I e S condicionais a $W = (W_1, \dots, W_n)$, $W^* = (W_1^*, \dots, W_p^*)$, $(\epsilon_1^*, \dots, \epsilon_q^*)$ e, em seguida, aplicando o logaritmo, obtém-se,

$$l_* = (\xi|W, W^*, \epsilon^*) \propto -n \log(\sigma_\epsilon) - \frac{S_*(\eta)}{2\sigma_\epsilon^2},$$

em que $S_*(\eta) = S(\eta|W, W^*, \epsilon^*) = \sum_{t=1}^n \epsilon_t^2(\eta|W, W^*, \epsilon^*)$ é denominada soma dos quadrados (SQ) condicional. Existem duas opções para escolha de W^* e ϵ^* :

- a) Primeiro é supor que seus elementos são iguais às suas esperanças $E(\epsilon_t) = 0$ e se $E(W_t) \neq 0$ substitui-se todos os elementos de W^* por \overline{W} .
- b) A segunda é indicada quando alguma raiz de $\phi(B) = 0$ estiver próxima do círculo unitário. O indicado então, é calcular $\epsilon_{p+1}, \dots, \epsilon_{p+n}$ a partir de (2.5) e assumindo os valores anteriores de ϵ_t iguais a zero. E, assim, estarão sendo usados os valores observados de w_t , obtendo-se

$$\epsilon_{p+1} = \widetilde{W}_{p+1} - \phi_1 \widetilde{W}_p - \dots - \phi_p \widetilde{W}_1 + \theta_1 \epsilon_p + \dots + \theta_q \epsilon_{p-q+1}.$$

ii) Procedimento não condicional.

Dada a função de verossimilhança não condicional

$$l(\xi) \cong -n \log(\sigma_\epsilon) - \frac{S(\eta)}{2\sigma_\epsilon^2}, \quad (2.7)$$

em que

$$S(\eta) = S(\phi, \theta) = \sum_{-\infty}^n [\epsilon_t(\eta, W)]^2 \quad (2.8)$$

é denominada soma de quadrados não-condicional, em que

$$[\epsilon_t(\eta, W)] = E(\epsilon_t | \eta, W). \quad (2.9)$$

Pode-se encontrar esta demonstração em Box, *et al* (1994).

Então, como boa aproximação para a EMV são os EMQ obtidos após minimização de (2.8). Dado, então, η deseja-se calcular a SQ e, para tanto, deve-se calcular as esperanças condicionais (2.9) por meio de (2.5). Porém, um processo chamado "backforecasting" será necessário para inicializar o processo, obtendo-se, assim, os valores $[W_{-j}]$ e $[\epsilon_{-j}]$, $j = 1, 2, \dots$, em fim valores antes do início da série.

Suponha um modelo ARIMA usual

$$\phi(B)W_t = \theta(B)\epsilon_t. \quad (2.10)$$

Então, pode-se demonstrar, usando a função geradora de autocovariância que a estrutura probabilística de W_1, \dots, W_n é igualmente explicada pelo modelo (2.10) ou pelo modelo

$$\phi(F)W_t = \theta(F)u_t, \quad (2.11)$$

em que F é o operador de transição para o futuro e u_t é um ruído branco com a mesma variância que ϵ_t . A equação (2.11) é chamada forma "backward" do processo e fornece uma representação estacionária e invertível na qual W_t é expressa somente em termos de valores de W_t e de u_t .

Desta maneira, o valor W_j tem a mesma relação probabilística com W_1, \dots, W_n que W_{n+j+1} tem com W_n, W_{n-1}, \dots, W_1 , ou seja, fazer previsão antes que a série se inicie é equivalente a prever a série inversa.

c) Função de verossimilhança exata

Através da simulação via Monte Carlo, a estimação por máxima verossimilhança exata é geralmente mais confiável em amostras pequenas (ANSLEY e NEWBOLD, 1980). Para o modelo ARMA(p,q) a função de verossimilhança é bem complexa (NEWBOLD, 1974).

Considere um modelo AR(1) e W_t uma série estacionária e gerada por

$$\widetilde{W}_t = \phi \widetilde{W}_{t-1} + \epsilon_t \quad |\phi| < 1$$

em que $\widetilde{W}_t = W_t - \mu$ e $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ são independentes.

Pode-se, também, escrever a expressão (2.11) da forma

$$\widetilde{W}_t = \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j}, \quad (2.12)$$

em que $\widetilde{W}_t \sim N(0, \sigma_\epsilon^2(1 - \phi^2))^{-1}$.

Com as duas últimas expressões, obtém-se

$$\widetilde{W}_1 = \sum_{j=0}^{\infty} \phi^j \epsilon_{1-j} = v_1,$$

$$\widetilde{W}_t = \phi \widetilde{W}_{t-1} + \epsilon_t, \quad t = 2, \dots, n. \quad (2.13)$$

em que $(v_1, \epsilon_2, \dots, \epsilon_n) \sim (0, \Sigma)$ e $\Sigma = \text{diag}((1 - \phi^2)^{-1} \sigma_\epsilon^2, \sigma_\epsilon^2, \dots, \sigma_\epsilon^2)$.

Logo, a função de distribuição conjunta de $(v_1, \epsilon_2, \dots, \epsilon_n)$ é expressa como

$$\begin{aligned} f(v_1, \epsilon_2, \dots, \epsilon_n) &= \left[\frac{(1 - \phi^2)}{2\pi\sigma_\epsilon^2} \right]^{1/2} \exp \left[-\frac{\widetilde{W}_1^2(1 - \phi^2)}{2\sigma_\epsilon^2} \right] \left[\frac{1}{2\pi\epsilon_\epsilon^2} \right]^{\frac{(n-1)}{2}} \\ &\times \exp \left[-\frac{1}{2\sigma_\epsilon^2} \sum_{t=2}^n (\widetilde{W}_t - \phi \widetilde{W}_{t-1})^2 \right]. \end{aligned}$$

E, conseqüentemente, o logaritmo da função de verossimilhança exata de $(\widetilde{W}_1, \dots, \widetilde{W}_n)$

é

$$l(W|\xi) = -\frac{n}{2} \ln 2\pi + \frac{1}{2} \ln(1 - \phi^2) - \frac{n}{2} \ln l\sigma_\epsilon^2 - \frac{S(\eta)}{2\sigma_\epsilon^2}, \quad (2.14)$$

em que

$$S(\eta) = (W_1 - \mu)^2(1 - \phi^2) + \sum_{t=2}^n [(W_t - \mu) - \phi(W_{t-1} - \mu)]^2. \quad (2.15)$$

Os estimadores de "máxima verossimilhança exata" são obtidos com a derivação de 2.14 com respeito a μ , ϕ e σ_ϵ^2 .

2.13.3 Verificação

Essa etapa da metodologia de Box-Jenkins consiste em verificar se o modelo identificado e estimado é adequado. Em caso positivo, pode-se adotá-lo para fazer previsão, caso negativo, outra especificação deve ser escolhida para modelar a série, o que implica refazer as etapas de identificação e estimação (GUJARATI, 2000).

Após estimar o modelo tem-se que verificar se ele representa, ou não, adequadamente, os dados, diante de qualquer insuficiência revelada pode-se sugerir um modelo alternativo como sendo adequado (GUJARATI, 2000).

Uma técnica utilizada é o superajustamento, ou seja, estima-se um modelo com parâmetros extras e verifica se estes são estatisticamente significativos e se sua inclusão diminui significativamente a variância residual (MORETTIN e TOLOI, 2004).

Segundo Gujarati (2000) a verificação do modelo pode ser feita analisando os resíduos, suponha que o modelo ajustado seja

$$\phi(B)y_t = \theta(B)\epsilon_t.$$

Se este modelo for verdadeiro, então os erros verdadeiros $\epsilon_t = \theta^{-1}(B)\phi(B)y_t$ constituirão um ruído branco. Nesta seção serão apresentados o teste de autocorrelação residual e o Teste de Box-Pierce, estes testes são baseados, em geral, nas autocorrelações estimadas dos resíduos (MORETTIN e TOLOI, 2004).

2.13.4 Teste de autocorrelação residual

Estima-se ϕ e θ , as quantidades

$$\hat{\epsilon}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)y_t,$$

são chamadas resíduos estimados ou simplesmente resíduos. Se o modelo for adequado, os $\hat{\epsilon}_t$ deverão estar próximos dos ϵ_t e, portanto, deverão ser aproximadamente não-correlacionados. Se indicarmos por \hat{r}_k as autocorrelações dos resíduos $\hat{\epsilon}_t$, então deve-se ter $\hat{r}_k \simeq 0$. Em particular, tem-se, aproximadamente,

$$\hat{r}_k \sim N\left(0, \frac{1}{n}\right),$$

sempre sob a suposição que o modelo ajustado é apropriado. As autocorrelações \hat{r}_k são calculadas por

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{\epsilon}_t \hat{\epsilon}_{t-k}}{\sum_{t=1}^n \hat{\epsilon}_t^2}.$$

Contudo, o desvio padrão de \hat{r}_k pode ser consideravelmente menor que $1/\sqrt{n}$, especialmente para pequenos valores de k , como mostrou Durbin (1970). Ele provou que para um AR(1), $Var[\hat{r}_k] \simeq \phi^2/n$, que pode ser bem menor que $1/n$. Box *et al* (1994) provaram que, para um modelo AR(1), tem-se

$$Var[\hat{r}_k] \simeq \frac{1}{n} [1 - \phi^{2(k-1)(1-\phi^2)}],$$

$$Cov\{\hat{r}_i, \hat{r}_j\} \simeq \frac{1}{n} [\delta_{ij} - \phi^{i+j-2}(1 - \phi^2)],$$

em que δ_{ij} é o delta de Kronecker, tem-se que, para k grande ou moderado, a variância de \hat{r}_k é, aproximadamente, $1/n$ e as autocorrelações estimadas são não-correlacionadas. De qualquer modo, a comparação de \hat{r}_k com os limites $\pm 2/\sqrt{n}$ fornece uma indicação geral de possível quebra de comportamento de ruído branco com ϵ_t com a condição de que seja lembrado que, para pequenos valores de k , estes limites subestimarão a significância de qualquer discrepância (MORETTIN e TOLOI, 2004).

2.13.5 Teste de Box-Pierce

Box e Pierce (1970) sugeriram um teste para as autocorrelações dos resíduos estimados que, apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos. Uma variação desse teste foi sugerida por Ljung

e Box (1978).

Se o modelo for apropriado, a estatística

$$Q(K) = n(n + 2) \sum_{k=1}^K \frac{r_k^2(\hat{\epsilon})}{n - k},$$

terá, aproximadamente, uma distribuição χ^2 com $K - p - q$ graus de liberdade. A hipótese de ruído branco para os resíduos é rejeitada para valores grandes de $Q(K)$, em geral basta utilizar as 10 ou 15 primeiras \hat{r}_k .

2.13.6 Previsão

Após o término do processo iterativo de identificação, estimação e checagem do diagnóstico, o qual proporciona uma estimativa do modelo gerador da série que se ajusta aos dados reais, pode-se utilizá-lo para prever valores futuros da variável (ANDERSON, 1976).

Entretanto, deve-se sempre ter presente que, na prática, nunca se tem o processo gerador verdadeiro da realização amostral. Dispõe-se apenas de uma aproximação desse processo gerador, a qual está sujeita a erros tanto no que se refere a identificação quanto à estimação. O bom desempenho do processo de previsão dependerá, fundamentalmente, da qualidade representativa do modelo teórico estimado, a qual depende do trabalho nas etapas de identificação e estimação (SOUZA, 1974).

Os processos de previsão com modelos de séries temporais são procedimentos que visam estender a valores futuros o modelo descrito e ajustado aos valores passados e ao valor presente da variável.

Dessa maneira, o estudo dos processos de previsão que aqui se analisam envolvem um conjunto de técnicas, que busca estabelecer uma base teórica que norteia o estudo de previsões de séries univariadas de tempo, proposto por Box e Jenkins, bem como os fundamentos da inferência estatística dos erros de previsão e os conseqüentes intervalos de confiança para aqueles procedimentos.

Ao estender a valores futuro prováveis o modelo de série temporal construído de forma a descrever os valores passados e o valor presente da variável, se está diante de uma análise que busca obter as características comportamentais sistemáticas da série, vale dizer, que busca apresentar uma descrição dos mecanismos da natureza aleatória do processo estocástico que gerou a série de valores passados e do valor presente.

O estabelecimento da estrutura probabilística da realização amostral da variável proporcionará os meios para inferir conclusões sobre a probabilidade associadas a cada possível valor futuro da variável.

Esta etapa visa utilizar o modelo identificado e estimado para fazer previsões. O objetivo agora é prever um valor y_{t+h} , $h \geq 1$, supondo que tem observações $\dots, y_{t-2}, y_{t-1}, y_t$ até o instante t , que é chamado origem das previsões.

É possível expressar a previsão $\hat{y}_t(h)$ da seguinte forma:

$$E(y_{t+h}) = \phi_1 E(y_{t+h-1}) + \dots + \phi_{p+d} E(y_{t+h-p-d}) - \theta_1 E(\epsilon_{t+h-1}) - \dots - \theta_q E(\epsilon_{t+h-1}) + E(\epsilon_{t+h}), \quad (2.16)$$

para $h \geq 1$. Aqui, deve-se utilizar as seguintes equações:

$$\begin{aligned} E(y_{t+h}) &= \hat{y}_t(h) & h > 0 \\ E(y_{t+h}) &= y_{t+h} & h \leq 0 \\ E(y_{t+h}) &= 0 & h > 0 \\ E(y_{t+h}) &= \epsilon_{t+h} & h \leq 0 \end{aligned}$$

Pode-se notar que:

- i) os termos de médias móveis desaparecem para $h > q$;
- ii) para calcular $\hat{y}_t(h)$ precisa-se de $\hat{y}_t(h-1), \hat{y}_t(h-2), \dots$ que são calculados recursivamente;
- iii) existe uma certa aproximação quando utilizamos esse procedimento, pois, na prática, só se conhece um número finito de dados passados. Portanto, na realidade, a previsão é $E(y_{t+h}|y_t, \dots, y_1)$ que é diferente da previsão ótima $E(y_{t+h}|y_t, \dots, y_{t-1})$. Entretanto as duas fórmulas forcecem resultados semelhantes para um valor grande de t . Essa aproximação é introduzida quando atribuímos valores iniciais para calcular os valores da seqüência ϵ_t ;
- iv) As previsões para um AR(p) são exatas, uma vez que pode ser demonstrado que esse modelo,

$$E(y_{t+h}|y_t, y_{t-1}, \dots) = E(y_{t+h}|y_t, \dots, y_{t+1-p}),$$

v) as inovações ϵ_t são calculadas recursivamente.

Para fazer previsões de valores futuros é necessário avaliar a precisão das previsões futuras oriundas de método selecionado. Considere uma série temporal com N observações e retire desta a n últimas observações, então gere previsões n passos à frente a partir do modelo estimado com $N - n$ observações e obtenha os erros de previsão

$$e_j = y_j - \hat{y}_j, \quad (2.17)$$

em que \hat{y}_j denota a previsão de y_j

A capacidade preditiva do modelo estimado pode ser avaliada utilizando as medidas de previsão a seguir, em que k é o número de passos à frente $k = N - n$

i) Erro quadrático médio (EQM):

$$EQM = \frac{1}{n} \sum_{j=k+1}^N e_j^2.$$

ii) Erro médio percentual (EMP):

$$EMP = \left(\frac{1}{n} \sum_{j=k+1}^N \frac{e_j}{y_j} \right) 100.$$

Obviamente, os modelos são tanto mais adequados segundo estes critérios quanto menor for o respectivo valor.

3 Material e Métodos

3.0.1 Caracterização do experimento

O experimento se localizou na Estação Experimental do Instituto Agrônomo Pernambuco (IPA), no município de Caruaru, no Estado de Pernambuco, com as seguintes coordenadas geográficas: latitude 08°14'18"S; 38°00'00" W. e altitude de 537 m, com classificação climática do tipo Dd'a' (semi-árido megatérmico) segundo Thornthwaite, sendo o solo caracterizado pela associação de Neossolo (EMBRAPA, 1999).

O experimento foi implantado com sementes de *Leucaena leucocephala* (Lam.) de Wit., tipo Hawaii, variedade K8, de origem australiana, coletada em plantio experimental no município de Floresta - PE, semeadas em agosto de 1989, sendo o plantio definitivo realizado em dezembro de 1989, 121 dias após o início do experimento (MEUNIER, 1991).

Foram utilizadas para este trabalho a altura média de 544 árvores da espécie *Leucaena leucocephala* (Lam.) de Wit., sobreviventes das 576 árvores plantadas no início do experimento. A variável mensurada foi altura média (Y, essas medições foram realizadas ao longo de 5 anos (60 meses), a partir do mês de janeiro de 1990, conforme verifica-se na Tabela 1.

t	altura	t	altura	t	altura	t	altura	t	altura
1	0,580	13	3,571	25	4,583	37	5,101	49	5,445
2	0,998	14	3,692	26	4,638	38	5,134	50	5,466
3	1,387	15	3,804	27	4,690	39	5,165	51	5,487
4	1,734	16	3,907	28	4,740	40	5,195	52	5,487
5	2,041	17	4,003	29	4,788	41	5,224	53	5,507
6	2,313	18	4,093	30	4,833	42	5,279	54	5,526
7	2,554	19	4,176	31	4,877	43	5,305	55	5,544
8	2,770	20	4,255	32	4,918	44	5,330	56	5,563
9	2,963	21	4,328	33	4,958	45	5,355	57	5,580
10	3,137	22	4,397	34	4,996	46	5,379	58	5,598
11	3,295	23	4,463	35	5,033	47	5,402	59	5,614
12	3,439	24	4,524	36	5,068	48	5,424	60	5,657

Tabela 1: Altura da leucena nos 60 meses de medição

Na estimativa dos parâmetros dos modelos não-lineares foi utilizado o procedimento NON-LINEAR do Systat versão 12 for Windows sendo utilizado o método de Marquardt. Na modelagem da série temporal utilizou-se o programa computacional Minitab Versão 14.

Para a seleção do melhor modelo de regressão não linear, foram observadas em conjuntos as seguintes estatísticas: índice de ajuste (IA %), e o erro padrão da estimativa (Syx%) (SCHNEIDER, 1998). O índice de ajuste (IA%), usado para a escolha da melhor equação não-linear, é descrito da seguinte forma

$$IA\% = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} 100,$$

em que

$IA\%$ = Índice de ajuste em percentagem;

Y_i = Valor observado (Valor real);

\hat{Y}_i = Valor ajustado;

\bar{Y} = Médias dos valores observados.

O erro padrão da estimativa percentual $S_{yx}\%$ foi obtido através da fórmula (SCHNEIDER, 1998).

$$S_{yx}\% = \frac{\sqrt{QMRes}}{\bar{Y}} 100,$$

em que

$S_{yx}\%$ = erro padrão da estimativa em percentagem;

\bar{Y} = Médias dos valores observados;

QMRes = quadrado médio do resíduo, obtido na análise da variância.

O erro padrão da estimativa é uma medida de dispersão entre os valores observados e estimados pela regressão, sendo desejável aquele que tenha o menor valor. Na comparação entre equações, a melhor por este critério é aquela que apresentar o menor valor de erro padrão da estimativa. Entretanto, essa estatística só pode ser usada como comparador entre várias equações com variáveis dependentes de mesma unidade (SCHNEIDER, 1998).

O erro padrão em percentagem é uma alternativa para a comparação de equações com

variáveis dependentes de diferentes unidades (MEYER, 1938).

Os modelos não lineares que serão utilizados para remover a tendência da séries são descritos na Tabela 2.

Modelo	Equação
Chapman-Richards	$H = \beta_0 [1 - \exp(-\beta_1 t)]^{\beta_2} \varepsilon$
Bertalanffy	$H = \beta_0 [1 - \exp(-\beta_1 t)]^3 \varepsilon$
Monomolecular	$H = \beta_0 [1 - \exp(-\beta_1 t)] \varepsilon$
Schumacher	$H = \beta_0 \exp(\frac{\beta_1}{t}) \varepsilon$
Silva-Bailey	$H = \beta_0 \exp(\beta_1 \beta_2^t) \varepsilon$

Tabela 2: Modelos não-lineares utilizados para remover a tendência

Após a remoção da tendência da série, utilizando-se o modelo de Chapman-Richards, a nova série designada S_t , foi modelada aplicando-se a metodologia de Box-Jenkins. Foram utilizados os modelos ARIMA(1,1,0), (1,1,1), (1,1,2) e (1,1,3).

4 Resultados e Discussão

Na Tabela 3 são apresentados as estimativas dos parâmetros dos modelos não-lineares testados, bem como seus índices de ajuste (IA%) e erro padrão da estimativa (S_{yx} %) obtidos através da modelagem de crescimento de leucena na Estação Experimental da Empresa Pernambucana de Pesquisa Agropecuária (IPA) no município de Caruaru-PE, no período 60 meses, a partir de janeiro de 1990.

Analisando os valores das estimativas dos parâmetros de todas as equações observa-se que a equação de Chapman-Richards, após as análises estatísticas, obteve o melhor resultado nas medidas de precisão, obtendo para o índice de ajuste 99,84% e um erro padrão da estimativa de 0,39% este o melhor modelo testado. Analisando as mesmas medidas de precisão o segundo melhor modelo foi o Monomolecular obtendo 98,93% e 0,41%, para índice de ajuste e erro padrão da estimativa respectivamente.

O modelo de Chapman-Richards é muito utilizado para estimar o crescimento de florestas artificiais que possui um desenvolvimento uniforme como o pinus e o eucalipto, tendo pouca tortuosidade do fuste, pouca ramificação e uma copa bem definida, levando em consideração a uniformidade das alturas entre os indivíduos, o mesmo não acontecendo com a leucena, que apresenta forma bem diferente, possuindo alta variabilidade de forma de fuste ramificação e copa, e desuniformidade entre as alturas dos indivíduos, mesmo sendo um plantio artificial. O modelo de Silva-Bailey aparece como o terceiro melhor modelo, apresentando um índice de ajuste de 98,39% e um erro de estimativa de 0,52%.

Os modelos de Chapman-Richards, Silva-Bailey, por considerarem o ponto de inflexão da curva de crescimento das espécies, os seus valores assintóticos e a sua forma sigmóide, obtiveram uma maior acuracidade dos modelos resultando maiores valores para as medidas de precisão, já o modelo de Bertalanffy não possui um dos princípios citados por Oliveira (1995), que leva em conta o ponto de inflexão da curva de crescimento ($m = 0$), resulta numa menor precisão dos índices estimados bem como obtenção de estimativas menos precisa.

Modelo	Predidor	Valor	StDev	IA%	S_{yx} %
Chapman-Richards	β_0	5,71937	0,48695	99,84	0,39
	β_1	0,05344	0,02849		
	β_2	0,70723	0,25642		
Bertalanffy	β_0	5,13711	0,16506	88,64	0,57
	β_1	0,18392	0,01973		
Monomolecular	β_0	5,47307	0,24273	98,93	0,41
	β_1	0,07972	0,01295		
Schumacher	β_0	6,02942	0,28353	97,46	0,46
	β_1	-6,15864	1,07285		
Silva-Bailey	β_0	5,44755	0,19028	98,39	0,42
	β_1	-1,65957	0,27469		
	β_2	0,90709	0,01735		

Tabela 3: Modelos não-lineares e as medidas de precisões

A Figura 2 apresenta as observações da série (altura da leucena) ao longo do tempo. O gráfico temporal de uma série permite visualizar seu comportamento, sendo muitas vezes possível identificar, por exemplo, a existência de sazonalidade, tendência, descontinuidades e não-estacionariedade. Observar-se na Figura 2 que a série em estudo apresenta uma forte tendência e não estacionariedade. Apesar da análise visual da Figura 2, é mais confiável fazer uma avaliação da função de autocorrelação da série, apresentada na Figura 3, na qual percebe-se que a FAC decai suavemente ao longo das defasagens, demonstrando assim, que a série possui uma memória longa e que não é estacionária.

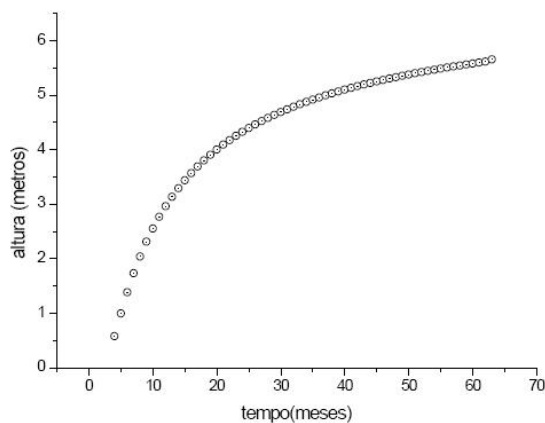


Figura 2: Altura da leucena

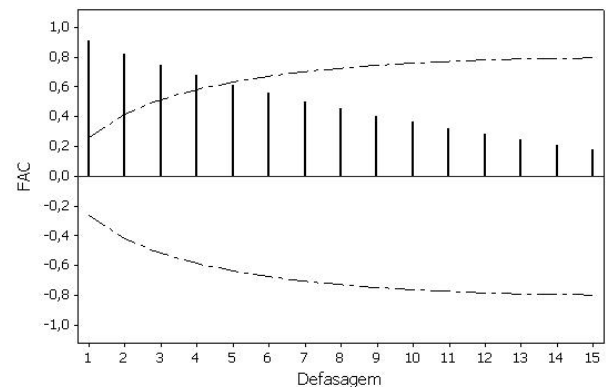


Figura 3: FAC da série original

Por conta da não-estacionariedade da série em estudo, utilizou-se o modelo de Chapman-Richards para remover a tendência, encontrando assim uma nova série denominada S_t . Essa nova série é constituída pelos erros provenientes do ajuste dos dados originais através do modelo de Chapman-Richards.

Na Figura 4 observa-se a nova série S_t , que será modelada utilizando a metodologia de Box-Jenkins.

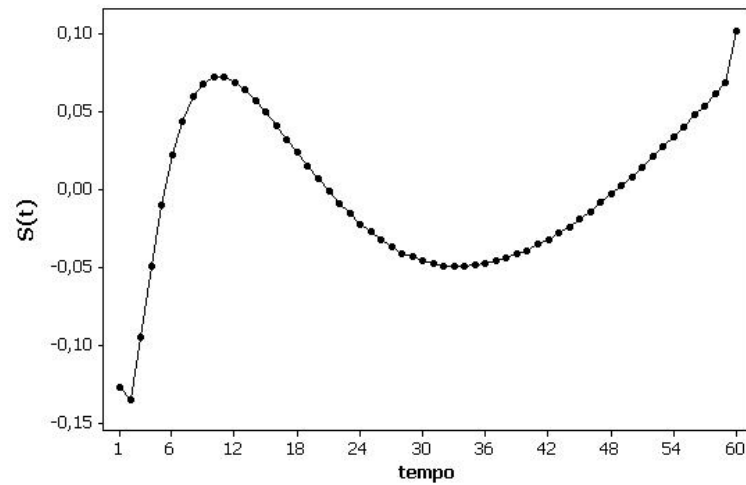


Figura 4: Erros provenientes do Modelo de Chapman-Richards

Após a remoção da tendência e uma diferenciação na série S_t , procedeu-se à análise do gráfico da função de autocorrelação parcial (Figura 5) da série S_t , a fim de determinar a ordem máxima da parte autoregressiva (AR).

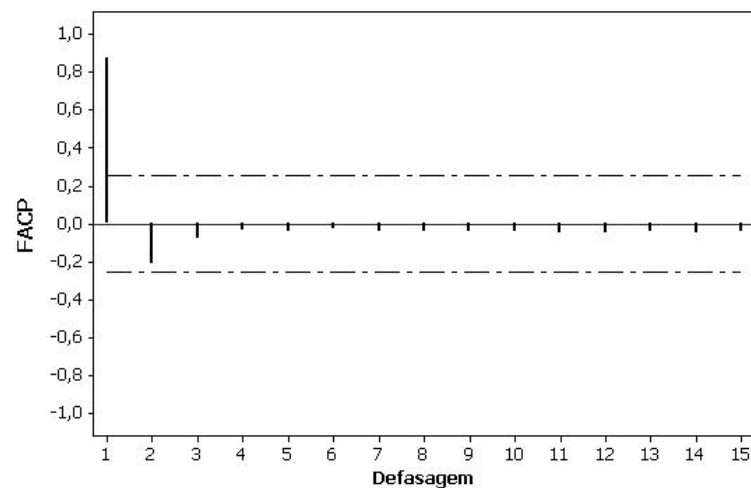


Figura 5: Função de autocorrelação parcial da Série S_t

Percebe-se na Figura 5, que só o primeiro valor da FACP está fora do intervalo de confiança, portanto, pode-se afirmar que o valor máximo para p será 1.

Analisando a Figura 6, observa-se que os três primeiros valores da FAC estão fora do intervalo de confiança, indicando o valor máximo de q igual a 3.

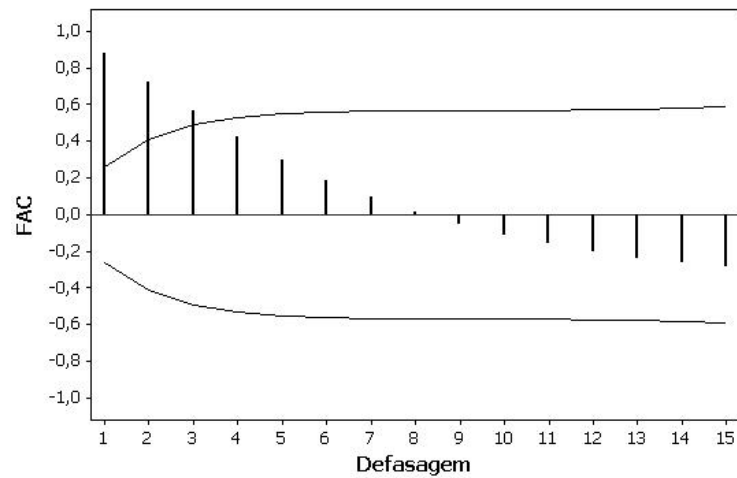


Figura 6: Função de autocorrelação da série S_t

Determinou-se que o modelo terá uma diferença, com no máximo um termo autoregressivo e no máximo três termos de média móvel. Foram estimadas todas as possíveis combinações, ou seja, ARIMA(1,1,0), ARIMA(1,1,1), ARIMA(1,1,2) e ARIMA(1,1,3) no cálculo das estimativas do parâmetros foi utilizado o software MINITAB Release 14.1 for Windows.

4.0.1 Modelo Arima(1,1,0)

A Tabela 4 apresenta o valor do parâmetro estimado e os respectivo valor da estatística t , além dos nível descritivo (p -valores) do modelo ARIMA(1,1,0).

Modelo	Coeficiente	Desvio Padrão	t	p
ARIMA(1,1,0)	$\phi_1 = 0,8664$	0,0805	10,77	0,000

Tabela 4: Modelo ARIMA(1,1,0) ajustado à série S_t

Analisando o p -valor na Tabela 4 observa-se que o parâmetro ϕ_1 é significativamente diferente de zero, o modelo apresenta uma variância residual de 0,00005625.

Para que o modelo seja considerado adequado é necessário que os resíduos do modelo estimado seja estimativas do ruído branco, ou seja, suas autocorrelações devem se comportar de maneira aleatória e devem ser não significativas.

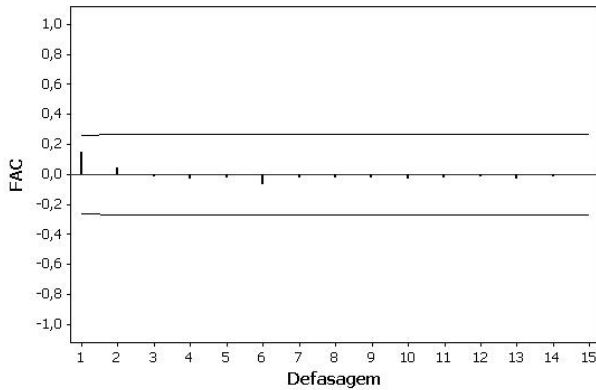


Figura 7: FAC - resíduos ARIMA(1,1,0)

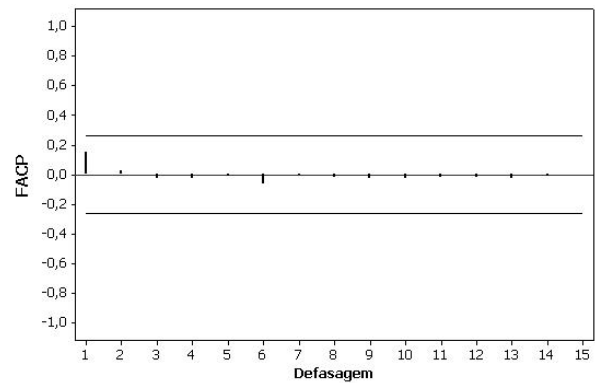


Figura 8: FACP - resíduos ARIMA(1,1,0)

A Figura 7 mostra o comportamento da função de autocorrelação (FAC) dos resíduos estimados do modelo ARIMA(1,1,0). Na Figura 8 observa-se a função de autocorrelação parcial dos resíduos, em ambas funções nota-se o comportamento aleatório dos resíduos, ou seja, não revela padrão específico, confirmando assim, a adequação do modelo.

A Tabela 5 apresenta os resultados do teste de χ^2 de Box-Pierce mostrando as estatísticas, graus de liberdade e níveis descritivos (p-valores).

defasagem	χ^2	gl	p-valor
12	2,2	11	0,998
24	2,7	23	1,000
36	3,1	35	1,000
48	5,0	47	1,000

Tabela 5: Teste de χ^2 de Box-Pierce para resíduos do modelo ARIMA(1,1,0)

Conforme os níveis descritivos apresentados na tabela acima, pode-se concluir que os resíduos podem ser considerados um ruído branco, pois o teste não rejeita para qualquer defasagem k , a hipótese de erros não correlacionados.

4.0.2 Modelo Arima(1,1,1)

A Tabela 6 apresenta os valores dos parâmetros estimados e os respectivos valores da estatística t , além dos níveis descritivos (p-valores) do modelo ARIMA(1,1,1).

Analisando os p-valores na Tabela 6 observa-se que todos os parâmetros são significativamente diferentes de zero, e o modelo ARIMA(1,1,1) apresenta uma variância residual de igual a 0,00004665.

Modelo	Coefficientes	Desvio Padrão	t	p
ARIMA(1,1,1)	$\phi_1 = 0,7587$	0,1153	6,58	0,000
	$\theta_1 = -0,6807$	0,1263	-5,39	0,000

Tabela 6: Modelo ARIMA(1,1,1) ajustado à série S_t

A Figura 9 mostra o comportamento da função de autocorrelação (FAC) dos resíduos estimados do modelo ARIMA(1,1,1). Na Figura 10 observa-se a função de autocorrelação parcial dos resíduos, em ambas funções nota-se o comportamento aleatório dos resíduos, ou seja, não revela padrão específico, confirmando assim, a adequação do modelo.

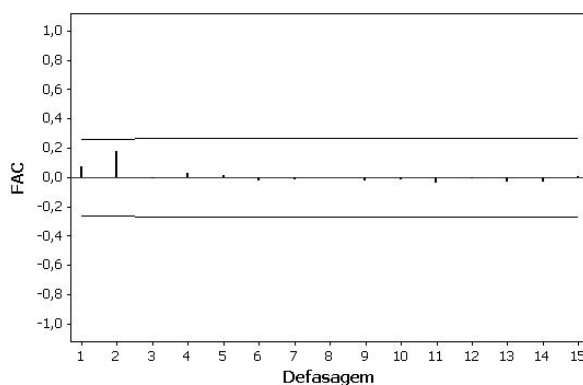


Figura 9: FAC - resíduos ARIMA(1,1,1)

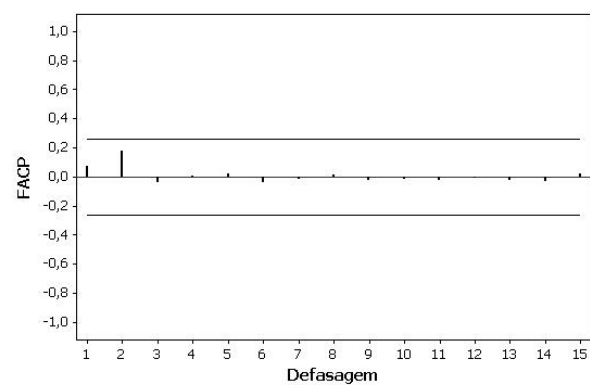


Figura 10: FACP - resíduos ARIMA(1,1,1)

A Tabela 7 apresenta os resultados do teste de χ^2 de Box-Pierce mostrando as estatísticas, graus de liberdade e níveis descritivos (p-valores).

defasagem	χ^2	gl	p-valor
12	2,7	10	0,988
24	3,8	22	1,000
36	5,0	34	1,000
48	9,6	46	1,000

Tabela 7: Teste de χ^2 de Box-Pierce para resíduos do modelo ARIMA(1,1,1)

Conforme os níveis descritivos apresentados na tabela acima, pode-se concluir que os resíduos podem ser considerados um ruído branco, pois o teste não rejeita para qualquer defasagem k , a hipótese de erros não correlacionados.

4.0.3 Modelo Arima(1,1,2)

A Tabela 8 apresenta os valores dos parâmetros estimados e os respectivos valores da estatística t , além dos níveis descritivos (p-valores) do modelo ARIMA(1,1,2).

Modelo	Coeficientes	Desvio Padrão	t	p
ARIMA(1,1,2)	$\phi_1 = 0,6491$	0,1548	4,19	0,000
	$\theta_1 = -0,9594$	0,1507	-6,37	0,000
	$\theta_2 = -0,5063$	0,1492	-3,39	0,001

Tabela 8: Modelo ARIMA(1,1,2) ajustado à série S_t

Analisando os p-valores na Tabela 8 observa-se que todos os parâmetros são significativamente diferentes de zero, e o modelo ARIMA(1,1,2) apresenta uma variância residual igual a 0,00004006.

A Figura 11 mostra o comportamento da função de autocorrelação (FAC) dos resíduos estimados do modelo ARIMA(1,1,1). Na Figura 12 observa-se a função de autocorrelação parcial dos resíduos, em ambas funções nota-se o comportamento aleatório dos resíduos, ou seja, não revela padrão específico, confirmando assim, a adequação do modelo.

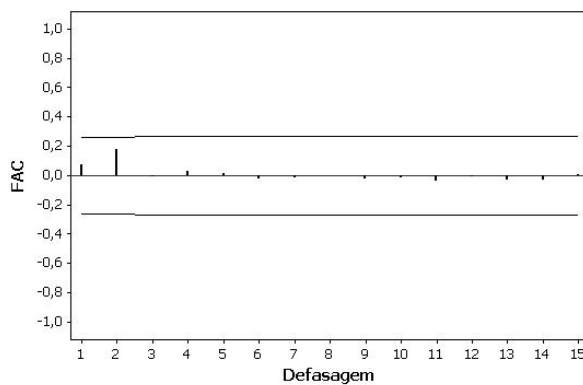


Figura 11: FAC - resíduos ARIMA(1,1,2)

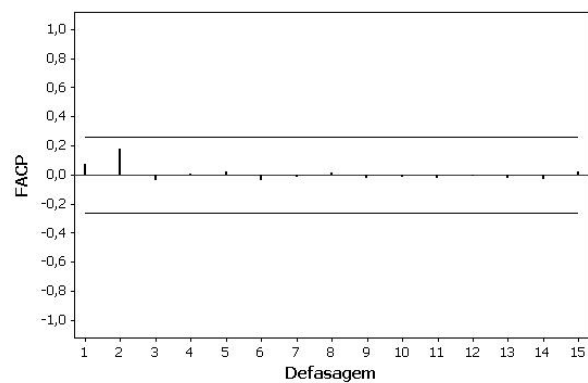


Figura 12: FACP - resíduos ARIMA(1,1,2)

A Tabela 9 apresenta os resultados do teste de χ^2 de Box-Pierce mostrando as estatísticas, graus de liberdade e níveis descritivos (p-valores).

defasagem	χ^2	gl	p-valor
12	3,7	9	0,927
24	6,1	21	0,999
36	8,7	33	1,000
48	12,9	45	1,000

Tabela 9: Teste de χ^2 de Box-Pierce para resíduos do modelo ARIMA(1,1,2)

Observando-se os níveis descritivos apresentados na tabela acima, pode-se concluir que os resíduos do modelo ARIMA(1,1,2) apresentam comportamento de um ruído branco, pois o teste não rejeita para qualquer defasagem k , a hipótese de erros não correlacionados.

4.0.4 Modelo Arima(1,1,3)

A Tabela 10 apresenta os valores dos parâmetros estimados e os respectivos valores da estatística t, além dos níveis descritivos (p-valores) do modelo ARIMA(1,1,3).

Modelo	Coefficientes	Desvio Padrão	t	p
ARIMA(1,1,3)	$\phi_1 = 0,5204$	0,2336	2,23	0,030
	$\theta_1 = -1,1089$	0,2454	-4,52	0,000
	$\theta_2 = -0,8049$	0,2996	-2,69	0,010
	$\theta_3 = -0,1102$	0,2200	-0,50	0,618

Tabela 10: Modelo ARIMA(1,1,3) ajustado à série S_t

Analisando os p-valores na Tabela 10 tem-se que os três primeiros parâmetros são significativamente diferentes de zero, e o modelo ARIMA(1,1,3) apresenta uma variância residual de 0,00003524.

A Figura 13 mostra o comportamento da função de autocorrelação (FAC) dos resíduos estimados do modelo ARIMA(1,1,3). Na Figura 14 observa-se a função de autocorrelação parcial dos resíduos, em ambas funções nota-se o comportamento aleatório dos resíduos, ou seja, não revela padrão específico, confirmando assim, a adequação do modelo.

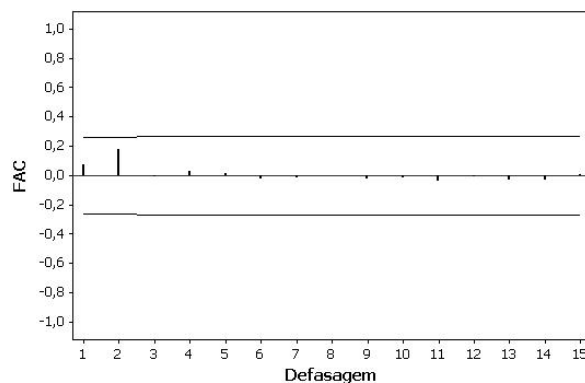


Figura 13: FAC - resíduos ARIMA(1,1,3)

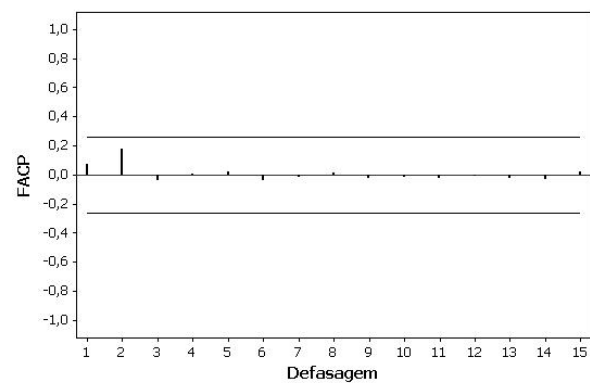


Figura 14: FACP - resíduos ARIMA(1,1,3)

A Tabela 11 apresenta os resultados do teste de χ^2 de Box-Pierce mostrando as estatísticas, graus de liberdade e níveis descritivos (p-valores).

defasagem	χ^2	gl	p-valor
12	8,3	8	0,404
24	11,5	20	0,931
36	15,4	32	0,994
48	21,7	44	0,998

Tabela 11: Teste de χ^2 de Box-Pierce para resíduos do modelo ARIMA(1,1,3)

Observando-se os níveis descritivos apresentados na tabela acima, pode-se concluir que os resíduos do modelo ARIMA(1,1,3) apresentam comportamento de um ruído branco, pois o teste não rejeita para qualquer defasagem k , a hipótese de erros não correlacionados.

Na Tabela 12 encontram-se os Erros Quadráticos Médios dos Modelos ARIMA, ajustados à série S_t , observa-se na referida tabela, que o modelo que apresenta o menor erro, é o ARIMA(1,1,0), à 12 passos à frente.

Modelos	EQM
ARIMA(1,1,0)	0,001
ARIMA(1,1,1)	0,002
ARIMA(1,1,2)	0,022
ARIMA(1,1,3)	0,047

Tabela 12: Erros Quadráticos Médios dos Modelos ARIMA

Na metodologia já mencionada para identificação de modelos ARIMA um dos maiores obstáculos é a seleção do melhor modelo. Muitos pesquisadores preferem utilizar critérios de seleção de modelos construídos com base na variância estimada de ϵ_t , no tamanho da amostra e nos valores de p e q .

A estatística de Schwarz, denotada por BIC (Bayesian Information Criterion) é expressa por

$$BIC(p, q) = \ln(\hat{\sigma}^2) + n \frac{\ln N}{N}.$$

A estatística de Akaike, denotada por AIC (Akaike Information Criterion) é expressa por

$$AIC(p, q) = \ln(\hat{\sigma}^2) + n \frac{2}{N},$$

em que $n = p + q$, N é o tamanho da amostra.

Quanto mais parâmetros são estimados no mesmo período de amostra, menor será o erro estimado, mas isso será penalizado na segunda parcela da estatística. Por isso, deseja-se o menor valor do AIC ou BIC.

Modelo	AIC	BIC
ARIMA(1,1,0)	-7,5163	-7,7755
ARIMA(1,1,1)	-6,7645	-6,8741
ARIMA(1,1,2)	-6,1180	-6,8924
ARIMA(1,1,3)	-5,0340	-5,8488

Tabela 13: Valores do AIC e BIC dos modelos ARIMA

Na Tabela 13 verifica-se os valores do AIC e BIC, para os modelos ARIMA seleccionados, constata-se que o menor valor de AIC e BIC, correspondem ao modelo ARIMA(1,1,0), sinalizado anteriormente pelo EQM.

Na Tabela 14 observa-se os valores reais e preditos da série S_t , à 12 passos à frente, utilizando o modelo ARIMA(1,1,0).

Modelo Arima (1,1,0)		
Período	Valor real	Valor predito
49	0,003	0,001
50	0,008	0,005
51	0,014	0,008
52	0,021	0,014
53	0,028	0,016
54	0,034	0,018
55	0,040	0,019
56	0,048	0,021
57	0,054	0,022
58	0,062	0,022
59	0,069	0,023
60	0,102	0,024

Tabela 14: Previsões da série S_t à 12 passos á frente

5 Conclusões

Aplicar técnicas de Séries Temporais para medidas de altura de árvores de leucena, mensuradas ao longo do tempo em uma série não-estacionária, pode resultar em equações de boas precisões, mas que quando comparadas a equações resultantes de ajustes de modelos tradicionalmente usados em estudos de crescimento em manejo florestal, não apresentou melhores resultados.

Sendo assim, recomenda-se o uso da análise temporal para estudos em que a série seja estacionária, pois mesmo removendo a tendência, os modelos encontrados não superiores aos tradicionais.

Referências

- AKAIKE, H. Maximum likelihood identification of gaussian autoregressive moving average models. *Biometrics*, 60, 255 - 265.
- ANDERSON, O. D. *Time Series Analysis and Forecasting: The Box-Jenkins Approach*. 1976.
- ANSLEY, C.F.; NEWBOLD, P. Finite amostra properties of estimators for autoregressive moving average models. 1980. *Journal of Econometrics*, 13, 159 - 183.
- BARTLETT, M.S. On the theoretical specification of sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society, série B*, vol. 27, 1946, p. 22-41.
- BERTALANFFY, L. V. Quantitative laws for metabolism and growth. *Quaterly Review of Biology*, Stony Brook, v. 32, p. 217-231, 1957.
- BERTALANFFY, L. V. *General system theory: foundations, development, applications*. New York: George Braziller, 1969. 295 p.
- BOX, G.E.P.; JENKINS, G.M. *Time Series Analysis, Forecasting and Control*. San Francisco. 1970. Holden Day.
- BREWBAKER, J. L. Can there is such a thing as a perfect tree? *Agroforestry Today*, Nairobi. v. 1, n. 4, p. 4 - 7, 1989.
- BREWBAKER, J. L. *Guide to the systematic of genus Leucaena (Mimosaceae)*. Cali: CIAT, 1978. 16p.
- BURKHART, H. E. Data collection and modeling approaches for forest growth and yield prediction. In: *PREDICTING FOREST GROWTH AND YIELD CURRENT ISSUES, FUTURE PROSPECTS*, 1987, Seattle. Anais Seattle: University Washington, 1987. p. 3 - 16.
- CAMPO, J.C.C. Análise de crescimento e produção em plantações desbastadas. *Rev. Árvore*, Viçosa, n.4, p.157-169, 1980.
- CHAPMAN, D. E. Statistical problems in dynamics of exploited fishing populations. In: *BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTIC AND PROBABILITY*, 1961, p. 153 - 158, *Procededings*. 1961.
- CORDEIRO, D.M. (2002), *Séries Temporais, Análise Quantitativa, Teoria e Aplicações*, Recife, EDUPE.
- CUNHA, L. S. *Leucaena: a árvore milagrosa de grande futuro energético para o Brasil*. *Jornal dos Reflorestadores*, [s.l.], v. 1, n. 4, p. 17 - 19, 1979.

- EMBRAPA. Centro Nacional de Pesquisa de Solos. Sistemas brasileiros de classificação de solos. Brasília, DF, 1999. 412 p.
- FERREIRA, A. B. de H. Novo dicionário Aurélio da língua portuguesa, 3. ed. Curitiba: Ed. Positivo, 2004. 2120 p.
- FINGER, C.A.G. Fundamentos de biometria florestal. Santa Maria: UFSM. 1992. 269 p.
- FRANCE, J.; THORNLEY, J. H. M. Mathematical models in agriculture. London: Butterworks, 1984. 335 p.
- FRANCO, A. A.; SOUTO, S. M. *Leucaena leucocephala* - uma leguminosa com múltiplas utilidades para os trópicos. Rio de Janeiro: EMBRAPA, 1986. 7 p. (Comunicação técnica, 2).
- FREITAS, A. R. et al. *Leucaena leucocephala* (Lam.) de Wit.: cultura e melhoramento. São Carlos: EMBRAPA - UEPAE, 1991. 93 p. (Documentos, 12).
- FRANCO, A. A.; SOUTO, S. M. *Leucaena leucocephala* - uma leguminosa com múltiplas utilidades para os trópicos. Rio de Janeiro: EMBRAPA, 1986. 7 p. (Comunicação técnica, 2).
- GRANGER, C.W.J.; NEWBOLD, P. Forecasting Economic Times Series. New York, Academic Press. 1986.
- GUJARATI, D.N., *Econometria Básica*, 3ª ed. São Paulo, Makron Books.
- HILL, G. D. *Leucaena leucocephala* for pastures in the tropics. *Herbage Abstracts*, Farnham Royal, v. 41, n. 2, p. 112 - 119, 1971.
- HUSCH, B.; MILLER, C. I.; BEERS, T. W. *Forest mensuration*. New York: J. Willey, 1982. 401 p.
- HUTTON, E. M. Breeding and selecting leucena for acid tropical soils. *Pesquisa Agropecuária Brasileira*, Brasília, v. 19, p. 263 - 74, 1984.
- KLUTHCOUSKI, J. *Leucena: uma alternativa para pequena e média agricultura*. 2. ed., Brasília: EMBRAPA, 1982. 12 p. (Circular técnica, 6).
- LJUNG, G.M.; BOX, G.E.P. On a measure of lack of fit in time séries models. *Biometrika*, 65, p. 297 - 303.
- MACHADO, S. A. *Studies in growth and yield estimation for Pinus taeda L. plantations in the State of Parana - Brazil*. 1978. 170 f. Thesis (Doctor of Philosophy) - University of Washington, 1978.
- MENDES, B. V. *Alternativa tecnológica para a agropecuária do semi-árido*. São Paulo: Nobel, 1985. p. 109 - 112. (Coleção Mossoroense, 266).
- MEUNIER, I. M. J. *Crescimento de mudas de Leucaena leucocephala* (Lam.) de Wit. Em função do uso de composto de resíduo urbano, adubação fosfatada e inoculação de *Rhizobium loti*. 1991. 110 f. Dissertação (Mestrado em Agronomia, Ciência do Solo) - Universidade Federal Rural de Pernambuco, Recife, 1991.

- MEYER, H. A. The standard error of estimate of tree volume from the logarithmic volume equation. *Journal of Forestry*, Bethesda, v. 36, p. 340-342, 1938.
- MORETTIN, P.A. Modelo (1981). Modelo para Previsões de Séries Temporais, volumes I e II. Rio de Janeiro IMPA - CNPq.
- MORETTIN, P.A.; TOLOI, C. M (2004). *Análise de Séries Temporais*. São Paulo. Edgard Blucher, 2004.
- NEWBOLD, P. The exact likelihood function for a mixed autoregressive-moving average process. *Biometrika*, 61, 423 - 426.
- OAKES, A. J. *Leucaena leucocephala*: description. Culture, utilization. *Advancing Frontiers of Plant Sciences*, [s.l.], v. 20, p. 1 - 114, 1968.
- OLIVEIRA, C. G. *Introdução ao estudo e modelos de crescimento e produção florestal*. Recife: UFRPE, 1995. 79 f. Apostila.
- PRODAN, M. *Forest biometrics*. [S. l.]: Pergamon Press, 1968. p.447.
- RICHARDS, F. J. A flexible growth function ofr empirical use. *Journal of Experimental Botany*, Oxford, v. 10, p. 290 - 300, 1959.
- SALERMO, A. R.; SEIFFERT, N. F. *Leucaena*: características agronômicas e recomendações de manejo. Florianópolis: EMPASC, 1986. 9 p. (Comunicado Técnico, 105).
- SANTOS J. R. C. B. Modelagem matemática na estimativa de crescimento em altura de leucena (*Leucaena leucocephala* (Lam.) de Wit)., no Agreste de Pernambuco. Recife: UFRPE, 2005. 95 f. Dissertação (Mestrado em Ciências Florestais) - Universidade Federal Rural de Pernambuco, 2005.
- SANTOS, J. R. Efeito de fosfogesso, superfosfato triplo, composto urbano e inoculações com *Bradyrhizobium* e micorrizas sobre o comportamento de mudas de leucena. 1991, 95 f. Dissertação (Mestrado em Agronomia - Ciência do Solo) - Universidade Federal Rural de Pernambuco, Recife, 1991.
- SCHNEIDER, P. R. *Análise de regressão aplicada à Engenharia florestal*. 2 ed. Santa Maria: UFSM/CEPEF, 1998. 236 p.
- SCHUMACHER, F. X. A new growth curve and its applications to timber-yield studies. *Journal of Forestry*, Bethesda, v. 37, p. 819-820, 1939.
- SCOLFORO, J.R. *Mensuração Florestal 4: avaliação da produtividade florestal através da classificação de sítio*. Lavras: ESAL/FAEPE, 1993. 138p.
- SILVA, J. A. A. Dynamics of stand structure in fertilized slash pine plantations. 1986, 133p. Thesis (Doctor of Philosophy) - University of Georgia, Athens.
- SIT, V. *Catalog of curves for curve fitting - Handbook 4*. Victoria: Ministry of Forests, Forest Science Research Branch. *Biometrics information*, 1994. 110p.
- SOUZA, A.A. Possibilidades de uso da leucena *Leucaena leucocephala* (Lam.) de Wit. como alimento suplementar para ruminantes nas condições do Nordeste brasileiro. Fortaleza: UFC, 1991. 51 p.

- SOUZA, R.C. Identificação e Aplicação de Testes para Modelos Adaptados à Previsão de Séries Temporais. Dissertação de Mestrado, DEE, PUC/RJ, Brasil.
- SPATHELF, P.; NUTTO, L. Modelagem aplicada ao Crescimento e produção florestal. Santa Maria: UFSM, 2000. 70 p.
- SPURR, S. H. Forest inventory. New York: Ronald Press, 1952. 476 p.
- VIEIRA, M. E. Q. O feno de leucena *Leucaena leucocephala* (Lam.) de Wit., como ingrediente não ortodoxo de rações para frangos de corte. Recife: UFRPR, 1992. 184 f. Dissertação (Mestrado em Zootecnia) - Universidade Federal Rural de Pernambuco, Recife, 1992.
- VILELA, E.; PEREIRA, J. V. S. Efeitos de densidade de semeadura e níveis de adubação nitrogenada no estabelecimento de *Leucaena leucocephala* (Lam.) de Wit. Boletim da Indústria Animal, Nova Odessa, v. 33, n. 2, p. 251 - 280, 1976.
- WONNACORR, T. H.; WONNACOTT, R. J. Introdução à estatística. Rio de Janeiro: Livros Técnicos e Científicos, 1980. 589p.

Apêndice

Anexo