

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
PRODUÇÃO**

**GERAÇÃO DE SUBSÍDIOS PARA A TOMADA DE
DECISÃO NA CADEIA PRODUTIVA DA
BOVINOCULTURA DO BRASIL**

DISSERTAÇÃO DE MESTRADO

Lorena Vicini

**Santa Maria, RS, Brasil
2006**

GERAÇÃO DE SUBSÍDIOS PARA A TOMADA DE DECISÃO NA CADEIA PRODUTIVA DA BOVINOCULTURA DO BRASIL

por

Lorena Vicini

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Engenharia de Produção, Área de Concentração em Qualidade e Produtividade, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Engenharia de Produção.**

Orientador Prof. Adriano Mendonça Souza

Santa Maria, RS, Brasil

2006

Vicini, Lorena, 1974-

V635g

Geração de subsídios para a tomada de decisão na cadeia produtiva da bovinocultura do Brasil / por Lorena Vicini ; orientador Adriano Mendonça Souza. – Santa Maria, 2006.
156 f. : il.

Dissertação (mestrado) – Universidade Federal de Santa Maria, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia de Produção, RS, 2006.

1. Engenharia de produção 2. Análise multivariada 3. Metodologia de séries temporais 4. Previsões e Incentivos fiscais 5. Estatística I. Souza, Adriano Mendonça II. Título

CDU: 658.5:311.2

Ficha catalográfica elaborada por
Luiz Marchiotti Fernandes – CRB 10/1160
Biblioteca Setorial do Centro de Ciências Rurais/UFSM

© 2006

Todos os direitos autorais reservados a Lorena Vicini. A reprodução de partes, ou do todo, deste trabalho, só poderá ser com autorização, por escrito, do autor.

Fone (0xx) 55 91166701

End. Eletr: lorenavicini@pop.com.br

**Universidade Federal de Santa Maria
Centro de Tecnologia
Programa de Pós-Graduação em Engenharia de Produção**

A Comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**GERAÇÃO DE SUBSÍDIOS PARA A TOMADA DE DECISÃO NA
CADEIA PRODUTIVA DA BOVINOCULTURA DO BRASIL**

elaborada por
Lorena Vicini

como requisito parcial para obtenção do grau de
Mestre em Engenharia de Produção

COMISSÃO EXAMINADORA:

Adriano Mendonça Souza, Dr.
(Presidente/Orientador)

Leandro Cantorski da Rosa, Dr. (UFSM)

Lindolfo Storck, Dr. (UFSM)

Santa Maria, 23 de março de 2006.

AGRADECIMENTOS

A Deus por ter estado presente em todos os momentos, sempre transmitindo-me a segurança necessária para que eu conseguisse prosseguir o meu caminho, e tornando os meus sonhos em realidade.

Ao Professor Dr. Adriano Mendonça Souza, deixo a minha sincera gratidão, não apenas por este trabalho, mas pelo conhecimento compartilhado durante todo este período em que tive a oportunidade de trabalhar sob sua orientação. Para mim será sempre um mestre, o qual deve ser seguido, pelo seu conhecimento e dedicação dispensada à concretização deste trabalho;

Ao meu namorado, o qual sempre foi o meu porto seguro, incentivando-me para que este trabalho se concretizasse.

A minha família, que, mesmo à distância, sempre torce para que este sonho se transforma em realidade;

Aos meus colegas e amigos, que sempre foram parte fundamental desta caminhada, deixo a minha sincera gratidão, pois em muitos momentos de tropeços fizeram com que o caminho se tornasse menos difícil de ser seguido, e que a palavra desistir não fazia parte dessa caminhada.

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Engenharia de Produção
Universidade Federal de Santa Maria, RS, Brasil

GERAÇÃO DE SUBSÍDIOS PARA A TOMADA DE DECISÃO NA CADEIA PRODUTIVA DA BOVINOCULTURA DO BRASIL

AUTORA: Lorena Vicini

ORIENTADOR: Adriano Mendonça Souza, Dr

Data e Local da Defesa: Santa Maria, 23 de março de 2006

A busca por produtos, que atendam determinado padrão de qualidade tem crescido significativamente, nos últimos anos, forçando, assim, os produtores a inserirem seu produto em determinados padrões de qualidade exigidas pelo mercado consumidor, sob pena de serem excluídos do mesmo. Observa-se, hoje, que, devido as exigências feitas pelo mercado consumidor, cada vez mais busca-se qualidade no produto, redução de custos, preços baixos, fazendo com que haja a necessidade de se iniciar o processo de mudanças que estabeleça melhorias das condições de criação. O avanço tecnológico exige que o produtor rural se ajuste à nova situação, e, para isso, é necessário estudar a situação da bovinocultura, que é o objetivo deste trabalho. Inicialmente, procedeu-se com a estatística descritiva, a qual possibilitou verificar o comportamento das variáveis analisadas. Em um segundo momento, utilizou-se as técnicas da análise multivariada, análise de agrupamentos que possibilitaram identificar as variáveis que possuíam características semelhantes, através dos grupos formados. Após, efetuou-se a análise fatorial, que reduziu a dimensionalidade do problema, utilizando-se apenas alguns fatores, que melhor explicarão as variáveis originais, e, determinando, assim, o número de fatores a serem utilizados, aplicou-se a análise de componentes principais, a qual possibilitou identificar os estados que se destacaram ao longo de todo o período analisado, bem como aqueles estados nos quais a bovinocultura não possui destaque. Em um terceiro momento fez-se uso da metodologia de séries temporais, a qual foi aplicada apenas naqueles estados que possuíam destaque na produção de bovinos. Monitorou-se o preço do boi gordo, em arroba no período analisado, o qual afeta diretamente na produção, e a partir disso fez-se a previsão para os anos de 2005 e 2006, a qual servirá de suporte para os pecuaristas traçarem suas futuras metas. Os resultados obtidos servirão de base para gerar políticas de incentivos fiscais para o desenvolvimento deste setor econômico. Esta pesquisa é relevante, pois, a partir da geração de subsídios, poderão ser traçadas metas a serem atingidas em todo o país.

Palavras-chave: Análise multivariada, metodologia de séries temporais, previsões e incentivos fiscais.

ABSTRACT

Masters Course Dissertation
Post Graduation Program in Production Engineering
Federal University of Santa Maria

GENERATING SUBSIDIES FOR DECISION MAKING IN LIVESTOCK'S PRODUCTIVE CHAIN IN BRAZIL

AUTHOR: Lorena Vicini

SUPERVISOR: Dr. Adriano Mendonça Souza

Date and Place of defence: Santa Maria, March 23rd 2006.

The search for products that would reach a certain level of quality has been significantly growing lately, thus forcing producers to insert their product in certain quality levels elected by the consumer market, under the risk of being excluded of it. Nowadays, due to those demands, quality, low prices, costs reduction are wanted. It all makes necessary to start the process for changes that would improve farming conditions. The advance of technology also wants the farmer to adapt to the new scenery, to do so it is needed a study of the livestock situation in Brazil from 1997 to 2004, which is the objective of this work. Initially, a descriptive statistics was applied, which allowed to verify the behaviour of the variables analyzed. Secondly, the techniques of the multivariate analyses were used, that is, group analysis that allow to identify similar variables according to each group. Later, a factors analysis was applied, and reduced the dimension of the problem by using some factors only to better explain the original variables. Then, after determining the number of factors to be used, was applied the principal components analysis method, which allowed to identifying in what States the livestock is not expressive. Thirdly, it was applied the temporal series method only to those States with expression in bovine production. It was also monitoring the live cattle prices in the period analyzed 1997 the 2004 which affects directly the production. From that data, it was possible to make the prevision for 2005 and 2006, in order to support farmers to draw their objectives. The results will work as support to create taxes incentive policies for the development of this economic sector. Therefore, this research is relevant as in the whole country will be possible to draw aims to be reached based on the future development of subsidies.

Key-words: Multivariate analysis, temporal series method, prevision and taxes incentives

LISTA DE QUADROS

QUADRO 1 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.....	102
QUADRO 2 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado de MG, no período de janeiro de 1997 a dezembro de 2004.....	107
QUADRO 3 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado do MS, no período de janeiro de 1997 a dezembro de 2004.....	112
QUADRO 4 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado do MT, no período de janeiro de 1997 a dezembro de 2004.....	117
QUADRO 5 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado do GO, no período de janeiro de 1997 a dezembro de 2004.....	122
QUADRO 6 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado do BA, no período de janeiro de 1997 a dezembro de 2004.....	127

LISTA DE FIGURAS

FIGURA 1 – Fases da produção de bovinos.....	24
FIGURA 2 – Categorias em que esta dividida a bovinocultura.....	26
FIGURA 3 - Etapas para a realização da análise de agrupamentos.....	32
FIGURA 4 – Classificação dos processos de aglomeração	35
FIGURA 5 – Distância mínima entre os grupos	36
FIGURA 6 - Distância máxima entre grupos.....	36
FIGURA 7 – Esquema da aplicação da análise de componentes principais	38
FIGURA 8 – Elipsóide de densidade constante	38
FIGURA 9 – Representação gráfica dos autovalores e autovetores	44
FIGURA 10 – Ciclo PDCA.....	63
FIGURA 11 – Gráfico referente à categoria dos bois com mais de 4 anos no período de 1997 a 2004.....	71
FIGURA 12 – Verificação do número de vacas no período de 1997 a 2004, nos estados brasileiros.....	72
FIGURA 13 – Gráfico referente à categoria representada pelos bezerras, no período de 1997 a 2004.....	73
FIGURA 14 – Gráfico referente à categoria representada pelos bezerras, no período de 1997 a 2004.....	73
FIGURA 15 – Gráfico referente à categoria representada pelas bezerras, no período de 1997 a 2004.....	74
FIGURA 16 – Gráfico referente à categoria representada pelos garrotes, de 1 a 2 anos, no período de 1997 a 2004.....	74
FIGURA 17 – Gráfico referente à categoria representada pelos garrotes, de 2 a 3 anos, no período de 1997 a 2004.....	75
FIGURA 18 – Gráfico referente à média de produção de cada estado, no período de 1997 a 2004.....	76
FIGURA 19 - Dendograma da matriz de distâncias, pelo método de agrupamento por ligação simples, no período de 1997 a 2000.....	77
FIGURA 20 – Dendograma com as variáveis no período de 1997 a 2000, pelo método de agrupamento por ligação simples.....	79

FIGURA 21 - Dendograma referente aos estados, utilizando o método de agrupamento de ligação simples, no período de 1997 a 2000.....	80
FIGURA 22 - Gráfico de explicação da proporção de variação de cada componente principal, no período de 1997 a 2000.....	83
FIGURA 23 - Gráfico representando a relação entre fatores fator 1 x fator 2, no período de 1997 a 2000.....	87
FIGURA 24 - Gráfico da distribuição da nuvem de variáveis, no círculo de correlações, no período de 1997 a 2000.....	88
FIGURA 25 - Gráfico da distribuição da nuvem de pontos (os estados), no período de 1997 a 2000.....	89
FIGURA 26– Dendograma com as variáveis, no período de 2000 a 2004, pelo método de agrupamento por ligação simples.....	90
FIGURA 27 - Dendograma referente aos estados, utilizando o método de agrupamento de ligação simples, no período de 2001 a 2004.....	91
FIGURA 28 - Gráfico de explicação da proporção de variação de cada componente principal, no período de 2001 a 2004.....	94
FIGURA 29 - Gráfico representando a relação entre fatores fator 1 x fator 2 no período de 2001 a 2004.....	96
FIGURA 30 - Gráfico da distribuição da nuvem de variáveis, no círculo de correlações, no período de 2001 a 2004.....	97
FIGURA 31 - Gráfico da distribuição da nuvem de pontos (os estados), no período de 2001 a 2004.....	98
FIGURA 32 – Série original referente ao preço do boi gordo, em arroba, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.	100
FIGURA 33 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do RS, no período de janeiro de 1997 a dezembro de 2004.....	101
FIGURA 34 – Sazonalidade do preço do boi gordo, em arroba, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.....	101
FIGURA 35 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo SARIMA (2,1,0)(1,0,1) ₁₂ , no estado do RS, no período de janeiro de 1997 a dezembro de 2004.....	103

FIGURA 36 - Gráfico da distribuição normal, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.....	103
FIGURA 37 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do RS, de janeiro a dezembro de 2005.....	104
FIGURA 38 - Série original referente ao preço do boi gordo, em arroba, no estado do MG, no período de janeiro de 1997 a dezembro de 2004.	106
FIGURA 39 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do MG, no período de janeiro de 1997 a dezembro de 2004.....	106
FIGURA 40 – Sazonalidade do preço do boi gordo, em arroba, no estado de MG, no período de janeiro de 1997 a dezembro de 2004.....	107
FIGURA 41 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo SARIMA $(1,1,0)(1,0,1)_{12}$, no estado do MG, no período de janeiro de 1997 a dezembro de 2004.....	108
FIGURA 42 - Gráfico da Distribuição Normal, do estado de MG, no período de janeiro de 1997 a dezembro de 2004.....	108
FIGURA 43 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do MG, de janeiro a dezembro de 2005.....	109
FIGURA 44 - Série original referente ao preço do boi gordo no estado do MS, no período de janeiro de 1997 a dezembro de 2004.	111
FIGURA 45 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do MS, no período de janeiro de 1997 a dezembro de 2004.....	111
FIGURA 46 – Sazonalidade do preço do boi gordo em arroba, no estado do MS, no período de janeiro de 1997 a dezembro de 2004.....	112
FIGURA 47 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo SARIMA $(0,1,1)(1,0,1)_{12}$, no estado do MS, no período de janeiro de 1997 a dezembro de 2004.....	113
FIGURA 48 - Gráfico da Distribuição Normal, do estado do MS, no período de janeiro de 1997 a dezembro de 2004.....	113
FIGURA 49 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do MS, de janeiro a dezembro de 2005.....	114

FIGURA 50 - Série original referente ao preço do boi gordo, em arroba, no estado do MT, no período de janeiro de 1997 a dezembro de 2004.	116
FIGURA 51 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do MT, no período de janeiro de 1997 a dezembro de 2004.....	116
FIGURA 52 – Sazonalidade do preço do boi gordo, em arroba, no estado do MT, no período de janeiro de 1997 a dezembro de 2004.....	117
FIGURA 53 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo SARIMA (0,1,1)(1,0,1) ₁₂ , no estado do MT, no período de janeiro de 1997 a dezembro de 2004.....	118
FIGURA 54 - Gráfico da Distribuição Normal, do estado do MT, no período de janeiro de 1997 a dezembro de 2004.....	118
FIGURA 55 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do MT, de janeiro a dezembro de 2005.....	119
FIGURA 56 - Série original referente ao preço do boi gordo, em arroba, no estado do GO, no período de janeiro de 1997 a dezembro de 2004.	121
FIGURA 57 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do GO, no período de janeiro de 1997 a dezembro de 2004.....	121
FIGURA 58 – Sazonalidade do preço do boi gordo em arroba, no estado de GO, no período de janeiro de 1997 a dezembro de 2004.....	122
FIGURA 59 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo ARIMA (1,1,2), no estado do RS, no período de janeiro de 1997 a dezembro de 2004.....	123
FIGURA 60 - Gráfico da Distribuição Normal, do estado de GO, no período de janeiro de 1997 a dezembro de 2004.....	123
FIGURA 61 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do GO, de janeiro a dezembro de 2005.....	124
FIGURA 62 - Série original referente ao preço do boi gordo, em arroba, no estado do BA, no período de janeiro de 1997 a dezembro de 2004.	126
FIGURA 63 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do BA, no período de janeiro de 1997 a dezembro de 2004.....	126

FIGURA 64 – Sazonalidade do preço do boi gordo, em arroba, no estado da BA, no período de janeiro de 1997 a dezembro de 2004.....	127
FIGURA 65 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo ARIMA (0,2,2), no estado do BA, no período de janeiro de 1997 a dezembro de 2004.....	128
FIGURA 66 - Gráfico da Distribuição Normal, do estado da BA, no período de janeiro de 1997 a dezembro de 2004.....	128
FIGURA 67 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do BA, de janeiro a dezembro de 2005.....	129

LISTA DE TABELAS

TABELA 1 – Componentes principais obtidas da análise de p variáveis X_1, X_2, \dots, X_p	45
TABELA 2 – Escores relativos a n objetos (indivíduos), obtidos em relação aos k primeiros componentes principais	46
TABELA 3 - Padrões Teóricos das FAC e FACP.....	59
TABELA 4 - Estatística descritiva referente à média do número de cabeças de bovinos vivos, representadas por cada categoria, no período de 1997 a 2004...	70
TABELA 5 – Matriz de correlação entre as variáveis, no período de 1997 a 2000.....	81
TABELA 6 – Autovalores e percentual da variância explicada de cada componente, no período de 1997 a 2000.....	82
TABELA 7 – Cargas fatoriais, para composição dos fatores, no período de 1997 a 2000.....	84
TABELA 8 – Cargas fatoriais, para composição dos fatores, no período de 1997 a 2000.....	86
TABELA 9 – Matriz de correlação entre as variáveis, no período de 2001 a 2004.....	92
TABELA 10 – Autovalores e percentual da variância explicada de cada componente, no período de 2001 a 2004.....	93
TABELA 11 – Cargas fatoriais, para composição dos fatores no período de 2001 a 2004.....	95
TABELA 12 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.....	100
TABELA 13 - Previsões para o preço do boi gordo, em arroba, no estado do RS, nos meses de janeiro a dezembro de 2005.....	104
TABELA 14 - Estatísticas descritivas, referentes ao preço do boi gordo em arroba, no estado do MG, no período de janeiro de 1997 a dezembro de 2004.....	105

TABELA 15 - Previsões para preço do boi gordo, em arroba, no estado de MG, nos meses de janeiro a dezembro de 2005.....	109
TABELA 16 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado do MS, no período de janeiro de 1997 a dezembro de 2004.....	110
TABELA 17 - Previsões para preço do boi gordo, em arroba, no estado do MS, nos meses de janeiro a dezembro de 2005.....	114
TABELA 18 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado do MT, no período de janeiro de 1997 a dezembro de 2004.....	115
TABELA 19 - Previsões para o preço do boi gordo, em arroba, no estado do MT, nos meses de janeiro a dezembro de 2005.....	119
TABELA 20 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado do GO, no período de janeiro de 1997 a dezembro de 2004.....	120
TABELA 21 - Previsões para preço do boi gordo, em arroba, no estado do GO, nos meses de janeiro a dezembro de 2005.....	124
TABELA 22 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado do BA, no período de janeiro de 1997 a dezembro de 2004.....	125
TABELA 23 - Previsões para preço do boi gordo, em arroba, no estado da BA, nos meses de janeiro a dezembro de 2005.....	129
TABELA 24 – Total da produção do rebanho bovino brasileiro, por estado e por categorias, no período de 1997 a 2004	138
TABELA 25 - Número de bois com mais de 4 anos de idade, nos estados brasileiros, no período de 1997 a 2004.....	145
TABELA 26 - Número de vacas, nos estados brasileiros, no período de 1997 a 2004.....	146
TABELA 27 - Número de bezerros, nos estados brasileiros, no período de 1997 a 2004.....	146
TABELA 28 - Número de bezerras, nos estados brasileiros, no período de 1997 a 2004.....	147

TABELA 29 - Número de novilhos de 2 a 3 anos, nos estados brasileiros, no período de 1997 a 2004.....	147
TABELA 30 - Número de garrotes de 1 a 2 anos, nos estados brasileiros, no período de 1997 a 2004.....	148
TABELA 31 - Número de garrotes de 2 a 3 anos, nos estados brasileiros, no período de 1997 a 2004.....	148
TABELA 32 – Matriz de correlação entre as variáveis, no período de 1997 a 2000.....	149
TABELA 33 – Matriz de correlação entre as variáveis, no período de 2000 a 2001.....	153

LISTAS DE REDUÇÕES

Símbolos com letras gregas

Σ = Matriz de variância-covariância populacional

$\hat{\Lambda}$ = Autovalores

ϕ = coeficiente do modelo autoregressivo AR(p);

ϕ_1 = coeficiente do modelo autoregressivo AR(1)

ϕ_p = coeficiente do modelo autoregressivo AR(p)

θ = coeficiente do modelo de médias móveis MA(q)

Θ = coeficiente do modelo de médias móveis do sazonal

Φ = coeficiente do modelo autoregressivo sazonal

ε_t = Erro aleatório

ε_t = Vetor dos erros

ε_{t-1} = Erro aleatório no instante t -1

y_{t-1} = Termo da série no instante t-1.

ρ_k = Autocorrelação

ρ_{kk} = Autocorrelação Parcial;

$\Delta^d y$ = diferença

$\Delta^D y$ = diferença sazonal

Símbolos com letras romanas

C = Matriz cofenética

$d(P_1, P_2)$ = Distância entre dois pontos

d_{ii} = Distância entre dois indivíduos

\hat{e} = Autovetores da matriz de correlação

F = Matriz fenética

I = Matriz identidade

R = Matriz de correlação

r_{ij} = Correlação entre variáveis

S = Matriz de variância-covariância amostral

\bar{x} = Autovetores da matriz de variância-covariância amostral

X_{ij} = i-ésimo objeto e j-ésima variável

X_i = i-ésima variável

\vec{X} = Vetor aleatório

$\vec{\bar{x}}$ = Vetor média

Z_{ij} = Valor padronizado

Y_i = i-ésimo componente principal

Siglas

AA = Análise de Agrupamentos

ACP = Análise de Componentes

AC = Acre

AIC = Critério de Informação de Akaike Principais

AF = Análise Fatorial

AL = Alagoas

AM = Amazonas

AM = Análise Multivariada

AP = Amapá

AR = Modelo Autoregressivo

ARMA = Modelo Autoregressivo de Médias Móveis

ARIMA = Modelo Autoregressivo Integrado de Médias Móveis

BA = Bahia

BIC = Critério de Informação Bayesiano

BST = Medida de adequação dos dados

CE = Ceará

CP = Componentes Principais

Côv – Covariância

CV = Coeficiente de Variação

DF = Distrito Federal

ES = Espírito Santo

FACr = Função de autocorrelação residual

FACPr = Função de autocorrelação parcial residual

FAC = Função de autocorrelação

FACP = Função de autocorrelação parcial

GO = Goiás

IBGE = Instituto Brasileiro de Geografia e Estatística

MA = Médias móveis

MA = Maranhão

MG = Minas Gerais

MS = Mato Grosso do Sul

MT = Mato grosso

PA = Pará

PB = Pernambuco

PDCA = Planejar, Executar, Verificar e Agir

PI = Piauí

PR = Paraná

RJ = Rio de Janeiro

RN = Rio Grande do Norte

RO = Rondônia

RR = Roraima

RS = Rio Grande do Sul

SAR = Modelo Autoregressivo Sazonal

SARIMA = Modelo Autoregressivo Integrado de Médias Móveis Sazonal

SC = Santa Catarina

SE = Sergipe

SP = São Paulo

ST = Série Temporal

TO = Tocantins

UE = União Européia

VAR = Variável

Var = Variância

KMO = Medida de adequação dos dados

SUMÁRIO

1 INTRODUÇÃO	19
1.1 Tema da pesquisa	20
1.2 Justificativa e importância da pesquisa	20
1.3 Objetivos	20
1.3.1 Objetivo geral	20
1.3.2 Objetivos específicos	21
1.4 Delimitação da pesquisa	21
1.5 Organização do trabalho	21
2 REVISÃO DE LITERATURA	23
2.1 A bovinocultura no Brasil e o cenário mundial	23
2.2 Análise descritiva dos dados	27
2.3 Técnicas da análise multivariada	28
2.3.1 Análise de agrupamentos.....	29
2.3.2 Análise de componentes principais	36
2.3.3 Análise fatorial relacionada à análise de componentes principais.....	48
2.4 Séries temporais	55
3 METODOLOGIA	63
3.1 Elaboração do banco de dados	64
3.2 Análise descritiva dos dados	64
3.3 Análise multivariada	65
3.3.1 Análise de agrupamentos	65
3.3.2 Análise de componentes principais e análise fatorial.....	66
3.3.3 Séries temporais.....	66
3.4 Geração de resultados	68
4 APLICAÇÃO DA METODOLOGIA E RESULTADOS	69
4.1 Análise descritiva	69
4.2 Análise de agrupamentos de 1997 a 2000	76

4.3 Análise fatorial e componentes principais de 1997 a 2000.....	80
4.4 Análise de agrupamentos de 2001 a 2004.....	90
4.5 Análise fatorial e componentes principais de 2001 a 2004.....	92
4.6 Séries temporais.....	99
4.6.1 Rio Grande do Sul.....	99
4.6.2 Minas Gerais.....	105
4.6.3 Mato Grosso do Sul.....	110
4.6.4 Mato Grosso.....	115
4.6.5 Goiás.....	120
4.6.6 Bahia.....	125
5 CONCLUSÕES E RECOMENDAÇÕES	131
6 BIBLIOGRAFIA	134
7 ANEXOS.....	137

1 INTRODUÇÃO

Em um país que tem mais gado do que gente, conforme censo realizado pelo IBGE, que mostra que somos 179 milhões de pessoas contra 195 milhões de bovinos, nada mais justo do que se ter a vantagem competitiva no cenário mundial (GUIA AG...). A agroindústria brasileira é responsável pela movimentação de 458 bilhões de reais, anual, o que equivale a um terço do produto interno bruto, e gera em torno de 18 milhões de empregos, sendo que 42%, de tudo o que o Brasil exporta, vem do agronegócio (GUIA AG, 2005).

De acordo com dados divulgados na revista Globo Rural (2005), o Brasil, no ano de 2003 exportou 1,36 milhão de toneladas de carne bovina *in natura* e industrializada, gerando 1,49 bilhão de dólares, já no ano de 2004 obteve um aumento significativo, atingindo 1,8 milhão de toneladas exportadas, representando, para a economia brasileira, 2,4 bilhões de dólares. Em contrapartida, o preço pago pelo boi gordo, no mercado interno, não pára de cair, sendo que a cotação teve uma queda de 2,13%, enquanto que os custos de produção tiveram um aumento médio em torno de 9,04%, comprometendo, dessa forma, a continuidade do crescimento das exportações.

Há nove anos o Brasil busca a adaptação do agronegócio brasileiro às novas regras sanitárias e de qualidade, elaboradas pela Organização Mundial do Comércio, como fórmula de agregar valor, abrir novos mercados e consolidar mercados existentes. Entre essas regras, destaca-se a certificação de produtos e a rastreabilidade dos rebanhos. A certificação abrange um conjunto de procedimentos que permite aferir um determinado produto atendendo às especificações estabelecidas, impondo requisitos de qualidade e padronização dos produtos, enquanto a rastreabilidade é a característica de um produto que dispõe de registro de todas as etapas envolvidas, desde a genealogia, passando por datas importantes, a partir do nascimento, as vacinações, os medicamentos, as transferências, o manejo e os principais fatos nutricionais.

Sendo a bovinocultura muito importante para a economia nacional, ela deve ser tratada como um processo produtivo. Dessa forma, os problemas devem ser detectados, solucionados, implementados e, também planejados, para que, no futuro, esse processo seja produtivo, lucrativo e traga um bom retorno para os seus gestores e consumidores, melhorando, de forma geral, a economia do país.

1.1 Tema da pesquisa

O tema da pesquisa abrange um estudo em escala nacional, com ênfase na bovinocultura, e aplicação de técnicas do sistema de qualidade para detectar as principais regiões produtoras a fim de monitorá-las.

1.2 Justificativa e importância da pesquisa

O mercado de alimentos tem se mostrado muito dinâmico, com alterações no comportamento da população, implicando em mudanças estruturais na demanda de produtos derivados de bovinos. Além disso, a crescente expansão da exportação de carnes e derivados faz com que haja incentivo à pesquisa e ao desenvolvimento tecnológico no setor.

Devido à bovinocultura ser um importante segmento do agronegócio brasileiro, que representa 42% de tudo o que o Brasil exporta, é importante identificar os estados que, num determinado período, destacaram-se, ou ainda destacam-se, na produção nacional de bovinos, bem como aqueles que buscam investir nessa atividade econômica, mas ainda não estão apontados como destaque nacional.

A presente pesquisa trará significativa contribuição, à medida que se saiba existir uma heterogeneidade de produção entre as regiões do Brasil, a qual deve-se, em especial, às diferenças climáticas, culturais e de incentivos fiscais, entre outros fatores.

1.3 Objetivos

1.3.1 Objetivo geral

Gerar subsídios para uma política de incentivo à bovinocultura, determinando os estados que são responsáveis por essa atividade econômica, assim como a previsão de seus valores futuros, por meio da análise multivariada (AM) e séries temporais (ST).

1.3.2 Objetivos específicos

Para que se cumpra esse objetivo geral, os seguintes objetivos específicos serão realizados:

- Identificar a categoria de bovinos, que cada região produz, bem como as regiões que se destacam em relação a essas, mediante a análise de agrupamentos, análise fatorial e análise de componentes principais;
- Traçar um perfil da bovinocultura nacional, em relação ao número de bovinos em cada estado, através da estatística descritiva e análise de agrupamentos;
- Analisar o crescimento histórico da bovinocultura e efetuar uma projeção para um ano, mediante séries temporais.

1.4 Delimitação da pesquisa

O tema abrange o estudo em escala nacional, referente a bovinocultura, sendo o período de análise, de 1997 a 2004, na qual a amostra coletada refere-se ao número total de cabeças de bovinos.

1.5 Organização do trabalho

Este trabalho está organizado em 5 capítulos.

No capítulo 1, aborda-se a importância do trabalho, a forma na qual este será desenvolvido, sua justificativa, a metodologia utilizada e a delimitação do estudo proposto.

No capítulo 2, apresenta-se revisão de literatura, abordando-se as técnicas que servem de suporte para o desenvolvimento desse trabalho.

No capítulo 3, aborda-se a metodologia do trabalho, bem como as técnicas utilizadas no mesmo.

No capítulo 4, apresenta-se a aplicação das técnicas e os resultados obtidos com a mesma.

No capítulo 5, apresenta-se a conclusão do trabalho desenvolvido e sugestões para trabalhos futuros.

2 REVISÃO DE LITERATURA

Os itens abordados neste capítulo servirão de embasamento para o desenvolvimento desta pesquisa. Inicialmente, é traçado um histórico da bovinocultura no Brasil, logo após aborda-se as técnicas estatísticas que serão utilizadas.

2.1 A bovinocultura no Brasil e o cenário mundial

A bovinocultura, no Brasil, sempre desempenhou um importantíssimo papel no contexto da sociedade e na economia do país, desde o período colonial. Essa atividade fornecia à população, da Colônia, não apenas o alimento fundamental representado pela carne, mas também a força motriz para os engenhos, o couro, com suas múltiplas utilidades e os animais de transporte para as zonas agrícolas e mineradoras.

Do ponto de vista técnico, a pecuária implantada no Brasil era uma atividade bastante rudimentar. A princípio, havia criação de gado nos engenhos de açúcar, onde o boi era utilizado para mover as moendas e fornecer carne para alimentação. Mas criado à solta, esse **gado de quintal** foi considerado contraproducente, pois ocupava extensos pastos que seriam mais lucrativos se cultivados para agricultura canavieira.

Como o eixo fundamental da economia colonial tinha por base as atividades exportadoras, o rei de Portugal (por intermédio de uma Carta Régia em 1701) proibiu a criação de gado numa faixa de 10 léguas a partir do litoral. Assim, o gado foi empurrado, oficialmente, para o sertão, relegado a ocupar áreas inadequadas para a agricultura exportadora. Condicionada pelo modelo de colonização implantado no Brasil, a pecuária, desde o início, teve como destino desbravar o sertão e realizar uma tarefa monumental, em termos de conquista e ocupação do vasto território brasileiro.

No escudo da pecuária colonial brasileira pode-se distinguir duas grandes zonas criatórias: as caatingas do Nordeste e as Campinas do Sul. Esse cenário manteve-se até o início do século XX.

Tendo em vista a expansão dessa atividade, o Governo Federal, através do Ministério da Agricultura e EMBRAPA, incentivou a pesquisa para melhoramento genético, e também o desenvolvimento de programas de criação intensiva. Esses programas foram inseridos, principalmente, na região do Mato Grosso, onde, nas décadas de 70 e 80, conforme Silva (2001, p.11), “esta região tinha na pecuária bovina de corte sua principal atividade econômica”, devido as constantes inundações ocorridas na região do pantanal mato-grossense, inviabilizando a exploração da agricultura.

Hoje, pode-se relacionar os estados, que possuem a bovinocultura como sua principal atividade econômica: MS, MT, MG e GO.

Apresentando o maior rebanho bovino do mundo, o Brasil possui baixo consumo *per capita/ano*. Pesquisas realizadas na década de 90 constataram um aumento das taxas de rentabilidade e produtividade do rebanho, e melhora nas carcaças brasileiras. Conforme revista do ABCZ n. 12 (2003), estima-se que no ano de 2005 o país tenha um rebanho de 208,1 milhões de cabeças, com desfrute de 22,0% e produção de 7,7 milhões de toneladas de carcaça.

Hoje, os produtores podem ser agrupados em dois sistemas de produção: um subsistema tradicional e o outro um subsistema melhorado, onde o primeiro caracteriza-se por manter uma taxa de natalidade em torno de 60% e a idade de abate em torno de 4 anos. O segundo caracteriza-se por manter a natalidade em torno de 70%, e a idade de abate em torno de 3 anos, podendo haver variações, de acordo com a adoção de tecnologias. Entretanto, as maiores mudanças ocorrem nas regiões do Centro Oeste, Sul e Sudeste.

De acordo com ANUALPEC (1998, p.18) "o mercado exige sistemas de produção voltados para produzir a custos unitários cada vez menores. Os preços pecuários tendem a permanecer apontando para a queda real a longo prazo, em função dos ganhos de produtividade previstos para os próximos anos. Produzir a custos cada vez mais baixos é, pois, a única saída”.

O sistema de produção de bovinos, hoje, está disposto em três fases, conforme Figura 1.

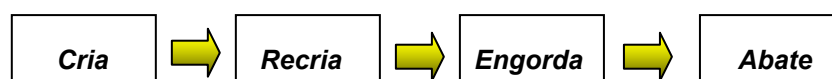


FIGURA 1 – Fases da produção de bovinos.

A 1ª fase da cria compreende da reprodução à desmama do bezerro. Essa fase de cria do gado de corte é a mais complexa do sistema de produção, pois envolve, especialmente, a produção de terneiros, e as maiores variações nas exigências envolvidas no processo. Várias alternativas têm sido estudadas, com o objetivo de aumentar a produtividade e a eficiência na produção de terneiros, visando a uma melhor utilização dos recursos de manejo e a minimização dos custos de produção. A adoção de sistemas de alimentação diferenciada para vacas de cria, através de pastagens naturais, e a utilização de suplementos e feno, têm sido alternativas aplicadas para o aumento das taxas de prenhez (BARCELOS, 1999, p.119).

A 2ª fase consiste na recria, que vai da desmama ao início da reprodução, ou da engorda. Na recria, os animais permanecem no pasto por mais tempo. No subsistema tradicional os animais são abatidos por volta de 4 anos, e a recria pode estender-se por 30 meses, podendo, nos casos de criação de novilhos precoces, diminuir este prazo para 12 a 15 meses.

A 3ª fase consiste na engorda, que vai da recria até o abate dos animais.

A 4ª fase consiste no abate de animais, ou seja, é a última etapa da produção, quando os animais estão prontos para atender as exigências do mercado consumidor. Aportes tecnológicos estão sendo inseridos com maior ênfase, nesta fase, especialmente no que se refere à nutrição dos animais. Não basta o criador investir em tecnologia, é fundamental que o mesmo tenha estímulo para melhorar a qualidade do seu produto. Para isso, é necessário que se coloque preços diferenciados para carnes de melhor, ou pior qualidade, como salienta a Revista do Agrobusiness, n. 107 (1996, p.33) “a classificação de carcaças e o estímulo ao produtor para redução da idade de abate, são condicionantes para que se deflagre um processo de modernização do setor de carne bovina no país”.

Segundo ANUALPEC (2003, p.14), “a maioria dos investimentos na pecuária leva anos para atingir resultados esperados”. Portanto, é de fundamental importância saber qual a realidade que o setor está vivendo, quando tais resultados aparecerem.

A bovinocultura, nos estados, está dividida por categorias de animais, o que pode ser observado na Figura 2. Essas categorias possibilitarão identificar o tipo de produção que ocorre em cada estado brasileiro.

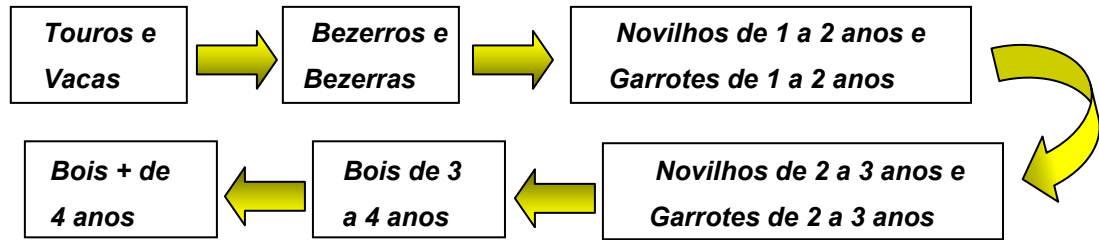


FIGURA 2 – Categorias em que esta dividida a bovinocultura.

A visão menos elaborada, das perspectivas do setor, está no insucesso de grande número de investimentos pecuários. Isso ocorre devido ao fato de a maioria das decisões estratégicas serem tomadas com base na realidade vivida no momento, que não será a mesma, quando surgirem os resultados esperados.

Num contexto mundial, ocorre uma reforma na política agrícola comum na União Européia (UE), que iniciou sua implementação no ano de 2003, levando a uma diminuição da produção local de carne bovina. O novo dispositivo eliminou as ajudas ligadas à produção (subsídios por cabeça no setor da pecuária) e os substituiu por um pagamento único por propriedade, independente do volume da produção e condicionada a respeito de normas ambientais, de segurança alimentar, de sanidade animal e vegetal e de bem-estar dos animais. Conforme ANUALPEC (2003. p.15), " a partir de 2006 o volume total de subsídios será congelado e sua distribuição passará a ser feita proporcionalmente à área de terras possuídas por produtor, independente de sua produção". Em consequência disso, algum dos quinze países membros já estão tomando a iniciativa de redução de subsídios prevista para 2006. Com isso, nos próximos dez anos a pecuária europeia tenderá a tornar-se menor e menos intensiva, o que resultará em menores exportações subsidiadas e maiores importações. Devido a esse fato, teve-se uma redução no rebanho de corte europeu e na oferta doméstica de carne. Enquanto isso, o consumo doméstico deve manter-se acima de 8,1 milhão de toneladas, até 2010, na UE ampliada, composta, hoje, por 25 países. Dessa forma, à medida em que a indústria de carne local reduz sua produção, com o aumento na demanda de consumo, a UE vai eliminar a superprodução, registrada até 2001, e passar, de uma situação de excedente, para uma situação de déficit.

Sendo o segundo maior mercado consumidor de carne bovina do planeta, em 2005 a UE apresentou 16,3% do consumo mundial, a auto-suficiência na produção (98,3% em 2004) pode contrair-se a níveis jamais observados. A comissão

européia prevê uma taxa de auto-suficiência de 96,7%, em 2010, para o bloco de 25 países (REVISTA ABCZ, 2003). O déficit, entre a decrescente produção doméstica e o consumo, deve gerar uma nova oportunidade para países exportadores de carne.

Dessa forma, o Brasil se insere para obter vantagens na nova conjuntura européia, já que se cria o gado, em vastos espaços, o que ajuda a evitar doenças deflagradas, constantemente, em países europeus. Nos últimos anos, as alíquotas alfandegárias, elevadíssimas, cobradas pela UE, não impediram o crescimento das importações de carne bovina brasileira, um produto altamente competitivo e apropriado ao perfil do consumidor europeu.

Esse mercado, de carnes certificadas e rastreadas, é abastecido por grupos de pecuaristas especializados, que exploram o diferencial de qualidade e a imagem de excelência dos seus produtos, exaltando que os seus animais são criados em harmonia com a natureza (criação em regime de pasto).

Atualmente, poucos são os países que apresentam uma estrutura fundiária de grandes propriedades em regiões de terras produtivas, como a que o Brasil possui no Centro-Oeste, Norte e Nordeste, que podem tirar proveito dos benefícios proporcionados pela economia de escala.

É esse fator que, sem desmerecer os outros, com excelentes centros de pesquisa nacionais, o clima e a mão-de-obra tornou possível o país transformar-se no mais competitivo exportador dos principais produtos agropecuários do planeta. Devido a isso, torna-se importante inserir a bovinocultura nos sistemas de qualidade, desenvolvidos, até o presente momento, nas linhas de produção das empresas.

2.2 Análise descritiva dos dados

O desenvolvimento e aperfeiçoamento de técnicas estatísticas de obtenção e de análise de informações permitem o controle e o estudo adequado de fenômenos, fatos, eventos e ocorrências, em diversas áreas do conhecimento. A estatística tem por objetivo fornecer métodos e técnicas para tratar, racionalmente, com situações sujeitas a incertezas.

Ao longo do século XX, os métodos estatísticos foram desenvolvidos como uma mistura de ciência, tecnologia e lógica, para solução e investigação de

problemas em várias áreas do conhecimento humano. Foi reconhecida como um campo da ciência, nesse período, mas sua história tem início bem anterior a 1900.

A estatística moderna é uma tecnologia quantitativa, que permite avaliar e estudar as incertezas e os efeitos no planejamento e interpretação de experiências, e de observações de fenômenos da natureza e da sociedade.

A estatística descritiva é um ramo que aplica várias das muitas técnicas usadas para sumarizar um conjunto de dados. De certa forma, descreve-se, ou sumariza-se, as características dos dados que pertencem a esse conjunto. Dentre as principais medidas descritivas, que serão utilizadas, estão: a média aritmética, que reflete o comportamento das variáveis, o cálculo do desvio-padrão, que possibilitará avaliar a dispersão de produtividade entre os estados e o coeficiente de variação, que é uma medida de dispersão relativa utilizada para mostrar o comportamento relativo entre as variáveis.

Além dessas medidas sintetizadoras, medidas de avaliação visual serão utilizadas, tais como gráficos, facilitando a comparação entre as variáveis em análise.

Embora a análise descritiva retrate o processo que está sendo estudado, muitas vezes uma análise mais específica é exigida.

Como a análise descritiva é uma técnica que vem sendo estudada por vários autores como Bussab et al. (1986), Costa Neto (2002), Fonseca (1996), Toledo (1985) e Triola (1999), não serão descritas, aqui, nessa pesquisa, a metodologia da análise descritiva.

2.3 Técnicas da análise multivariada

Em quase todas as áreas do conhecimento pesquisas são realizadas, e várias características (variáveis) são observadas. Essas variáveis, em geral, não são independentes, e por isso devem ser analisadas conjuntamente. Análise Multivariada é a área da Estatística que trata desse tipo de análise. Várias são as técnicas que podem ser aplicadas aos dados. Sua utilização depende do tipo de dado que se deseja analisar e dos objetivos do estudo. Aqui, utiliza-se as seguintes técnicas multivariadas: Análise de Agrupamentos, Análise Fatorial e Análise de Componentes Principais.

2.3.1 Análise de agrupamentos

Todos nós acreditamos que qualquer população é composta de segmentos distintos. Se trabalhamos com as variáveis adequadas, a análise de agrupamentos nos ajudará a ver se existem grupos que são mais semelhantes entre si do que com membros de outros grupos (Tom Myers, consultor Burke Customer, Satisfaction Associates).

A análise de agrupamentos (AA), em sua aplicação, engloba uma variedade de técnicas e algoritmos, sendo que o objetivo é encontrar e separar objetos em grupos similares. Essa técnica pode ser observada, por exemplo, se se tiver vários produtos em uma determinada prateleira de um supermercado, e distribuir esses produtos, na prateleira, segundo suas características, de um mesmo composto, ou o mesmo princípio ativo, por exemplo. Aí está-se a praticar AA. Agora, se esses produtos estiverem espalhados por toda a prateleira, significa que se terá mais de uma característica, e, para que se possa uní-los por características comuns, será muito trabalhoso, exigindo conceitos mais sofisticados de semelhança, e procedimentos mais científicos para juntá-los. É em relação a esse procedimento multidimensional que se trabalhará.

Em alguns estudos torna-se necessário conhecer algumas características de determinado grupo de um conjunto de elementos amostrais, principalmente quando é resultante de uma ou mais variáveis. Quando se obtém mensuração de diferente natureza, pode-se observar se há similaridades no conjunto de dados. Um dos métodos a AA, que poderá ser utilizado para os objetivos acima descritos.

A AA estuda todo um conjunto de relações interdependentes. Não faz distinção entre variáveis dependentes e independentes, isto é, variáveis do tipo causa e efeito, como na regressão.

Conforme Everitt (1974 apud BUSSAB et al., 1990), a AA pretende resolver o seguinte problema: “dada uma amostra de n objetos (ou indivíduos), cada um deles medindo segundo p variáveis, procurar um esquema de classificação que agrupe os objetos em g grupos. Deve ser determinado, também, o número de variáveis desses grupos”. Portanto, a finalidade dessa técnica é reunir os objetos (indivíduos, elementos) verificados nos grupos em que haja homogeneidade dentro do grupo e heterogeneidade entre os grupos, objetivando propor classificações. Os objetos em um grupo são relativamente semelhantes, em termos dessas variáveis, e diferentes de objetos de outros grupos. Quando utilizada dessa forma, a AA é o

inverso da análise de fatores, pelo fato de reduzir o número de objetos, e não o número de variáveis, concentrando-os em um número muito menor de grupos.

A AA constitui uma metodologia numérica multivariada, com o objetivo de propor uma estrutura classificatória, ou de reconhecimento da existência de grupos, objetivando, mais especificamente, dividir o conjunto de observações em um número de grupos homogêneos, segundo algum critério de homogeneidade. Muitas vezes, nessa técnica, são feitas afirmativas empíricas, que nem sempre têm respaldo teórico. Muitas técnicas são propostas, mas não há, ainda, uma teoria generalizada e amplamente aceita (REGAZZI, 2001). Devido a isso, deve-se utilizar vários métodos e comparar os resultados, para que a análise dos dados seja realizada pela técnica mais adequada.

A AA é um método simples, baseada nos cálculos de distância, no entanto não requer conhecimento estatístico para a sua aplicação, como é o caso quando se aplica análise de variância, de regressão, ou fatorial. AA não requer o uso de um modelo, os demais casos necessitam. Para a aplicação da AA, as estatísticas e os conceitos, a seguir, serão utilizados:

Esquema de aglomeração: Informa sobre objetos, ou casos a serem combinados em cada estágio de um processo hierárquico de aglomeração.

Composição de um Agrupamento: Indica o agrupamento ao qual pertence cada objeto, ou caso (MALHOTRA, 2001, p.528).

Dendograma ou Fenograma: Também chamado de gráfico em árvore. Este representa uma síntese gráfica do trabalho desenvolvido, sintetizando a informação, ocasionando uma pequena perda da mesma, pelo fato de ser uma síntese. Embora aconteça essa perda de informação, esse gráfico é de grande utilidade para a classificação, comparação e discussão de agrupamentos.

Há duas formas de se representar um dendograma: horizontal e vertical.

No dendograma horizontal, as linhas verticais, ou o eixo das ordenadas, representam os grupos unidos por ordem decrescente de semelhança, e a posição da reta, na escala ou o eixo das abscissas, indica as distâncias entre os grupos que foram formados. O dendograma é lido de cima para baixo, quando for feito na forma horizontal. Como hoje, ainda, não existe uma teoria que diga em qual altura deve-se fazer um corte no gráfico, é o pesquisador quem decide.

No dendograma vertical, a leitura é feita da direita para esquerda, no qual as linhas verticais, ou o eixo das ordenadas, indicam as distâncias entre os grupos foram formados, e a posição da reta na escala, ou o eixo abscissas, representa os grupos unidos por ordem decrescente de semelhança.

Distância entre centros de agrupamentos. Indica a distância que separa os pares individuais de grupos. Sendo que os grupos que se apresentam bem separados são distintos. São esses os desejáveis para a análise (MALHOTRA, 2001, p.528).

O primeiro passo, para realizar a AA, consiste em formular o problema de aglomeração, definindo as variáveis sobre as quais se baseará o agrupamento. Logo após, faz-se a coleta dos dados, que serão reunidos numa tabela com m colunas (variáveis) e n linhas (objetos). Antes de escolher a medida de distância para a análise dos dados, é necessário verificar se os mesmos encontram-se com a mesma unidade de medida. Caso contrário, deve-se fazer a padronização dos mesmos. Escolhe-se, então, uma medida apropriada de distância, que irá determinar o quão semelhantes, ou diferentes, são os objetos que estão sendo agrupados. Dentre vários processos de aglomeração, o pesquisador deve escolher aquele que for mais apropriado ao problema estudado.

A decisão sobre o número total de conglomerados, a constarem na análise, caberá ao pesquisador, pois esta dependerá de cada pesquisa.

Acredita-se que a formulação do problema seja a parte mais importante da análise de agrupamentos, ou seja, a escolha das variáveis nas quais se baseará o processo de aglomeração. A inclusão de uma, ou duas variáveis, sem importância, poderá vir a distorcer o resultado final da análise. O conjunto de variáveis escolhido deve descrever a semelhança entre objetos, em termos relevantes para o problema em pesquisa. Esta fase é importante para a AA, pois é onde se fixa o critério de homogeneidade. Segundo Bussab & Morettin (1986, p.2), “critérios distintos levam a grupos homogêneos distintos, e o tipo de homogeneidade depende dos objetivos a serem alcançados”.

A estrutura básica da aplicação da AA pode ser representada em etapas, conforme mostra a Figura 3. Deve-se observar que essas etapas não são independentes. Algumas vezes será necessário voltar a etapas anteriores, para corrigir e aprimorar algumas etapas posteriores. Considera-se que as etapas

descritas na Figura 3 formam um procedimento metodológico muito útil para a realização da AA.

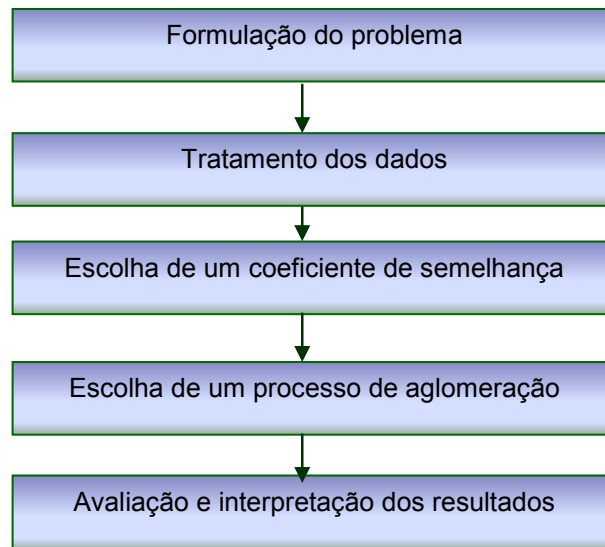


FIGURA 3 - Etapas para a realização da análise de agrupamentos.

Ao analisar os dados, em primeiro lugar deve-se verificar se eles devem ser tratados. Por exemplo, deve-se observar se as variáveis foram medidas em unidades muito diferentes entre si. A solução por aglomerado será influenciada pelas unidades de medida. Nesse caso, deve-se, antes de aglomerar as amostras, padronizar os dados. Embora a padronização possa remover a influência da unidade de medida, poderá também reduzir as diferenças entre grupos em variáveis que melhor descrevam os conglomerados, pois as unidades associadas às variáveis podem, arbitrariamente, afetar o grau de similaridade entre os objetos, e a padronização dos dados faz com que esse efeito da arbitrariedade seja eliminado, fazendo com que as variáveis possuam a mesma contribuição no cálculo do coeficiente de similaridade entre os objetos.

Para que seja possível padronizar as variáveis, é necessário ter-se uma matriz de dados com p variáveis ($j = 1, 2, \dots, p$) e n objetos ($i = 1, 2, \dots, n$). Sendo que, na matriz de dados, o valor do i -ésimo objeto e j -ésima variável será denotado por X_{ij} , no qual o valor padronizado será representado por Z_{ij} . Onde as variáveis padronizadas terão média 0 e variância constante 1, sendo esta a mais utilizada na prática, e é representada pela seguinte função:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j} \quad (2.1)$$

sendo cada i fixo, no qual $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, p$.

Para aplicar a AA, em um conjunto de dados, é muito importante a escolha de um coeficiente que quantifique o quão parecidos dois objetos são. Esse coeficiente pode ser dividido em duas categorias, que dizem respeito à estimação de uma medida de similaridade, ou dissimilaridade, entre os indivíduos, ou populações, a serem agrupados. Na medida de similaridade, quanto maior for o valor observado, mais parecido serão os objetos. Já na medida de dissimilaridade, quanto maior for o valor observado, menos parecido serão os objetos. Um exemplo de medida de similaridade é o coeficiente de correlação, pois quanto maior seu valor, maior a associação e de dissimilaridade a distância euclidiana, pois quanto menor o valor mais próximo os objetos estão uns dos outros (REGAZZI, 2001).

Como o objetivo da análise de agrupamento é reunir objetos semelhantes, torna-se necessário alguma medida para avaliar o quão semelhantes, ou diferentes são os objetos. Geralmente, costuma-se avaliar a semelhança em termos de distância entre pares de objetos. Os objetos que possuem a menor distância entre si são mais semelhantes do que os objetos com a maior distância. Essa medida de semelhança é fornecida pela distância euclidiana.

Um grande problema da AA é a escolha da medida de proximidade mais adequada, sendo que as técnicas são baseadas em diferentes medidas de proximidade, e nem sempre chegam ao mesmo resultado. Devido a isso, é importante testar mais de uma medida de distância, para que possa ser utilizada a mais adequada para a análise.

Conforme Malhotra (2001, p. 529), “a utilização de diferentes medidas de distância pode levar a resultados diferentes de aglomeração. Assim, é conveniente utilizar medidas diferentes e comparar os resultados”.

As medidas de distância consideram que, se dois indivíduos são similares, estão próximos um do outro, ou seja, são comuns ao conjunto de variáveis e vice-versa.

A distância euclidiana é, sem dúvida, a medida de distância mais utilizada para a análise de agrupamentos.

Considerando o caso mais simples, no qual existem n indivíduos, onde cada um dos quais possuem valores para p variáveis, a distância euclidiana entre eles é obtida mediante o teorema de Pitágoras, para um espaço multidimensional.

Segundo Manly (1986), “a distância euclidiana, quando for estimada a partir das variáveis originais, apresenta a inconveniência de ser influenciada pela escala, de medida, pelo número de variáveis e pela correlação existente entre as mesmas”. Para contornar as escalas, faz-se a padronização das variáveis em estudo, para que possuam a variância igual à unidade, conforme item (2.1).

Considerando dois indivíduos i e i' , a distância entre eles é dada por:

$$d_{ii'} = \sqrt{\sum_{j=1}^p (X_{ij} - X_{i'j})^2} \quad (2.2)$$

A escolha do método de agrupamento é tão difícil quanto a escolha do coeficiente de associação. Dessa escolha dependerá a correta classificação de uma amostra estar dentro de um grupo, ou de outro, que já tenha sido formado. Os métodos de agrupamento foram desenvolvidos com base nos modelos e dados diversos.

Há grande quantidade de métodos de agrupamento. As diferenças entre os métodos existem em função de diferentes formas de definir proximidade entre um indivíduo em um grupo, contendo vários indivíduos, ou entre grupos de indivíduos.

Na AA, não se pode dizer que exista um método que seja melhor para se aplicar. O pesquisador deve decidir qual será o mais adequado para o desenvolvimento do seu trabalho, pois cada método leva a um resultado. Os métodos de agrupamento mais utilizados são os hierárquicos.

Como se pode observar na Figura 4, os processos de agrupamento podem ser divididos em dois grupos: hierárquicos ou não-hierárquicos. Conforme Malhotra (2001, p. 529), a **aglomeração hierárquica** caracteriza-se pelo estabelecimento de uma hierarquia, ou estrutura em forma de árvore, sendo esta a mais utilizada. Os métodos hierárquicos são divididos em *aglomerativos* e *divisivos*. Como nesse trabalho será utilizado apenas o métodos aglomerativos, não serão enunciados os demais.

O *agrupamento aglomerativo* tem início em um grupo separado. Formam-se os grupos reunindo-se os objetos em grupos cada vez maiores. O processo continua até que todos os objetos sejam membros de um único grupo, sendo esse método seqüencial, onde os objetos são reunidos um após o outro, respeitando uma determinada seqüência de aglomeração. O critério básico da fusão entre um objeto e um grupo, ou entre dois grupos, é sempre o mesmo: serão reunidos os grupos que tenham maior similaridade entre si. O problema é: como calcular esta similaridade? O método de cálculo depende do método de aglomeração escolhido (VALENTIN, 2000).

A Figura 4 mostra os dois processos de agrupamento, hierárquicos e não-hierárquicos.

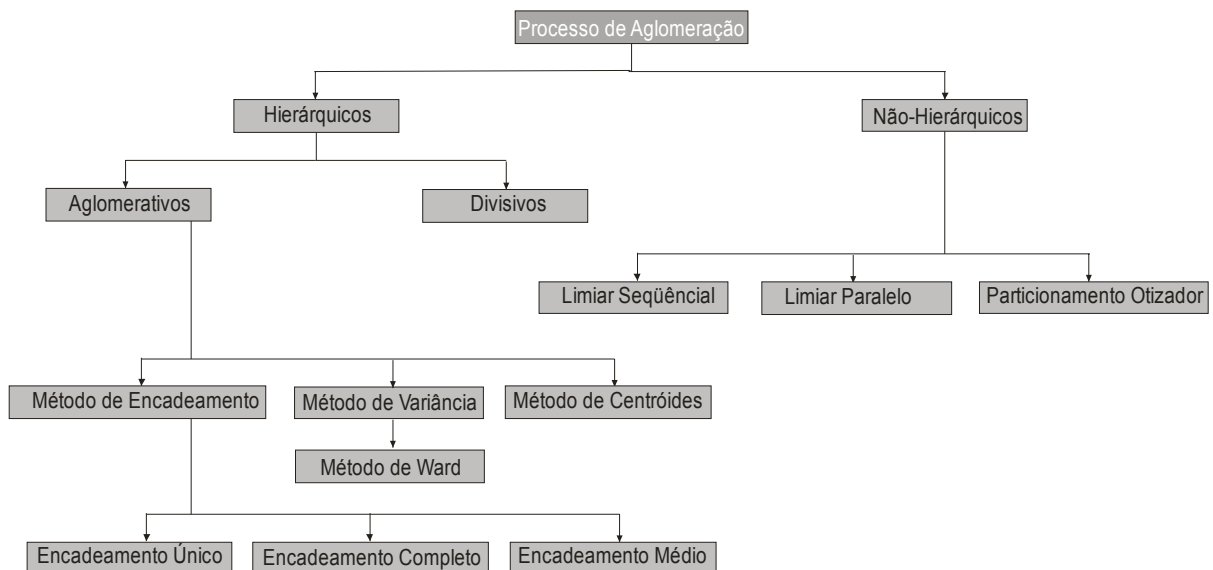


FIGURA 4 - Classificação dos processos de aglomeração.
Fonte: MALHOTRA (2001, p. 531).

O procedimento básico, de todos os métodos aglomerativos de agrupamento, é similar. Inicia-se com o cálculo de uma matriz de distâncias entre as variáveis e finaliza-se com um dendograma, no qual é possível verificar as fusões sucessivas dos indivíduos, até os indivíduos formarem um único grupo (REGAZZI, 2001).

Os métodos aglomerativos são de uso comum. Esses são constituídos de métodos de encadeamento, métodos de erros de somas de quadrados, ou métodos de variância e métodos centróides.

Os **métodos de encadeamento** compreendem:

O método do encadeamento único (*Single Linkage*), que se baseia na distância mínima, regra do vizinho mais próximo. Os dois primeiros objetos agrupados são os que apresentam menor distância entre si. Identifica-se a menor distância agrupando-se o terceiro objeto com os dois primeiros, ou formando um novo grupo de dois objetos. Em cada estágio a distância entre dois grupos é definida como a distância entre seus dois pontos mais próximos. Dois grupos podem incorporar-se em cada estágio por meio do encadeamento mais curto entre eles. Continua-se o processo até que todos os objetos estejam em um único grupo, conforme Figura 5.

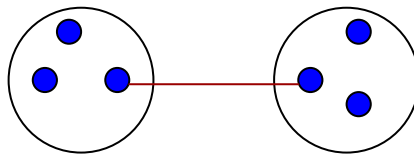


FIGURA 5 - Distância mínima entre os grupos.

O método do encadeamento completo (*Complete Linkage*) é semelhante ao encadeamento único, embora se baseie na distância máxima entre os objetos ou o método do vizinho mais afastado. Neste, a distância entre dois grupos é calculada entre seus dois pontos mais afastados, conforme Figura 6.

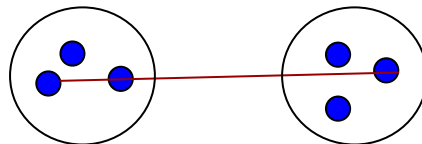


FIGURA 6- Distância máxima entre grupos.

Neste trabalho, será abordado apenas o método, ou algoritmo de agrupamento, do encadeamento único (*Single Linkage*), ou, ainda, método do vizinho mais próximo.

2.3.2 Análise de Componentes Principais

A análise de componentes principais (ACP) tem por objetivo descrever os dados contidos num quadro indivíduos-variáveis numéricas: p variáveis serão medidas com n indivíduos. Esta é considerada um método fatorial, pois a redução do

número de variáveis não se faz por uma simples seleção de algumas variáveis, mas pela construção de novas variáveis sintéticas, obtidas pela combinação linear das variáveis iniciais, por meio dos fatores (BOUROCHE & SAPORTA, 1982).

A ACP é uma técnica matemática da análise multivariada, que possibilita investigações com um grande número de dados disponíveis. Possibilita, também, a identificação das medidas responsáveis pelas maiores variações entre os resultados, sem perdas significativas de informações. Além disso, transforma um conjunto original de variáveis em outro conjunto: os componentes principais (CP) de dimensões equivalentes. Essa transformação, em outro conjunto de variáveis, ocorre com a menor perda de informação possível, sendo que esta também busca eliminar algumas variáveis originais que possuam pouca informação. Essa redução de variáveis só será possível se as p variáveis iniciais não forem independentes e possuírem coeficientes de correlação não-nulos.

A meta da análise de componentes principais é abordar aspectos como a geração, a seleção e a interpretação das componentes investigadas. Ainda pretende-se determinar as variáveis de maior influência na formação de cada componente, que serão utilizadas para estudos futuros, tais como de controle de qualidade, estudos ambientais, estudos populacionais entre outros.

A idéia central da análise baseia-se na redução do conjunto de dados a ser analisado, principalmente quando os dados são constituídos de um grande número de variáveis inter-relacionadas. Procura-se redistribuir a variação nas variáveis (eixos originais) de forma a obter o conjunto ortogonal de eixos não correlacionados, essa redução é feita transformando-se o conjunto de variáveis originais em um novo conjunto de variáveis que mantém, ao máximo, a variabilidade do conjunto (REGAZZI 2001, p.1). Isto é, com a menor perda possível de informação.

Na prática, o algoritmo baseia-se na matriz de variância-covariância, ou na matriz de correlação, de onde são extraídos os autovalores e os autovetores.

A análise de componentes principais tem a finalidade de substituir um conjunto de variáveis correlacionadas por um conjunto de novas variáveis não-correlacionadas, sendo essas combinações lineares das variáveis iniciais, e colocadas em ordem decrescente por suas variâncias, $VAR CP_1 > VAR CP_2 > \dots > VAR CP_p$ (VERDINELLI, 1980).

As novas variáveis geradas denominam-se CP, possuem independência estatística e são não correlacionadas. Isso significa que, se as variáveis originais

não estiverem correlacionadas, as ACP não oferecerão vantagem alguma. Variáveis dependentes querem dizer que o conhecimento de uma variável importa para o conhecimento da outra (SOUZA, 2000).

Para a determinação das componentes principais é necessário calcular a matriz de variância-covariância (Σ), ou a matriz de correlação (R), encontrar os autovalores e os autovetores e, por fim, escrever as combinações lineares, que serão as novas variáveis, denominadas de componentes principais.

O esquema descrito na Figura 7 servirá de base para a aplicação da ACP.

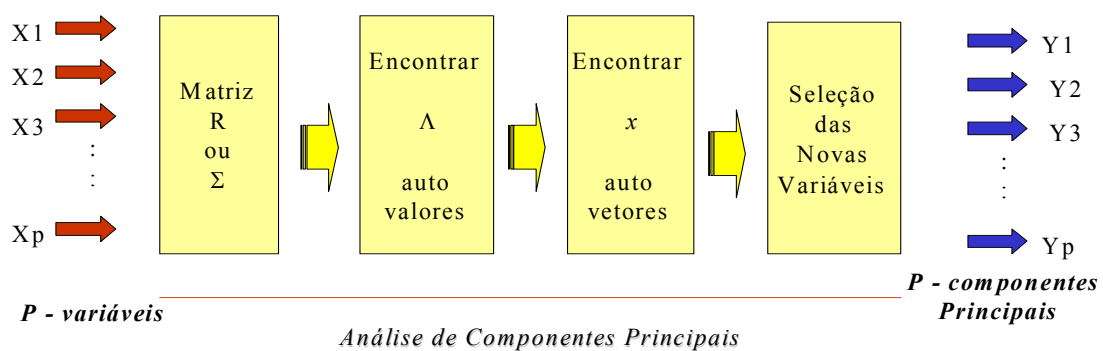


FIGURA 7 - Esquema da aplicação da análise de componentes principais.
Fonte: SOUZA (2000, p.25).

Supondo-se que, na análise que se está realizando, existam apenas duas variáveis X_1 e X_2 , conforme a Figura 8, observa-se o elipsóide de densidade de probabilidade constante.

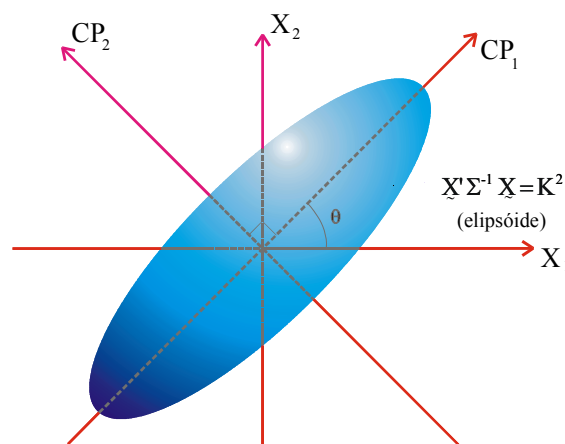


FIGURA 8 - Elipsóide de densidade constante.
Fonte: LOPES (2001, p.31).

O primeiro componente corresponde ao maior eixo da elipse (CP_1), e o comprimento desse eixo é proporcional a $\sqrt{\Lambda_1}$. O eixo de menor variância (CP_2) é perpendicular ao eixo maior. Esse eixo chama-se segundo componente principal, e seu comprimento é proporcional a $\sqrt{\Lambda_2}$. Assim, a análise das componentes principais toma os eixos X_1 e X_2 e os coloca na direção de maior variabilidade (JOHNSON & WICHERN, 1992).

O primeiro estágio da ACP é a conversão da matriz $n \times p$ de dados em uma matriz quadrada, onde n é o número de indivíduos e p representa um conjunto de variáveis.

Intuitivamente, percebe-se que, quanto maior for o número de variáveis, e quanto mais estas forem interdependentes entre si será mais fácil comparar indivíduos baseando-se nos valores dessas variáveis, originais (REGAZZI, 2001). Essa interdependência é representada pela matriz de variância-covariância Σ , ou pela matriz de correlação R .

Seja Σ a matriz de variância-covariância associada ao vetor aleatório $\vec{X} = [X_1, X_2, \dots, X_p]$. Se Σ possuir o par de autovalores e autovetores estimados da amostra analisada, serão representados por $(\hat{\Lambda}_1, X_1), (\hat{\Lambda}_2, X_2), \dots, (\hat{\Lambda}_p, X_p)$, onde $\hat{\Lambda}_1 \geq \hat{\Lambda}_2 \geq \dots \geq \hat{\Lambda}_p \geq 0$, e fornecerão o i -ésimo componente principal dado por:

$$Y_i = \vec{x}_i' X = \vec{x}_{1i} X_1 + \vec{x}_{2i} X_2 + \dots + \vec{x}_{pi} X_p, \text{ onde } i = 1, 2, \dots, p.$$

Com as escolhas de que:

$$Var(Y_i) = \vec{x}_i' \Sigma \vec{x}_i = \hat{\Lambda}_i \quad i = 1, 2, \dots, p$$

$$Cov(Y_i, Y_k) = \vec{x}_i' \Sigma \vec{x}_k = 0 \quad i, k = 1, 2, \dots, p$$

Se algum $\hat{\Lambda}_i$ é igual, a escolha do coeficiente do vetor correspondente \vec{x}_i também será, e, então, Y_i não é único.

Essa definição mostra que os CP são não correlacionados e possuem variâncias iguais ao autovalor de Σ (JOHNSON & WICHERN, 1992).

Para proceder a ACP, em casos populacionais, utiliza-se a matriz de variância covariância Σ . Porém, quando se tratar de um conjunto de dados amostrais, a matriz será estimada através da matriz de variância-covariância amostral S , e o vetor média por $\vec{\bar{X}} = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p]$ (SOUZA, 2000).

Para aplicar a análise de componentes principais, deve-se seguir algumas etapas, até obter-se o resultado final.

Inicialmente, calcula-se a matriz S , ou a matriz R , e verifica-se se as variáveis estão correlacionadas umas em relação às outras. Caso não estejam, deve-se aplicar o teste do KMO , ou fazer um teste que verifique se as correlações entre as variáveis são significativas, ou não, para verificar se é possível proceder a análise dos dados aplicando essa técnica, conforme item (2.7).

O pesquisador deve verificar, também, se as variáveis foram medidas em escalas diferentes. Deve-se proceder a padronização das mesmas, para evitar erros nos resultados.

Na etapa seguinte, decide-se pelo número total de componentes que melhor explicarão o conjunto de variáveis originais. Existem duas formas de selecionar esses componentes:

- Mediante os autovalores, pelo critério sugerido por Kaiser (1960 apud MARDIA, 1979), que consiste em incluir somente aquelas componentes cujos valores próprios sejam superiores a 1. Esse critério tende a incluir poucas componentes quando o número de variáveis originais é inferior a vinte e, em geral, utiliza-se aquelas componentes que consigam sintetizar uma variância acumulada em torno de 70%.
- Através do método gráfico, este critério considera as componentes anteriores ao ponto de inflexão da curva. Foi sugerido por Cattell (1966) e exemplificado por Pla (1986).

Decidido o número de componentes, passa-se a encontrar os autovetores que irão compor as combinações lineares, os quais irão formar as novas variáveis.

A última etapa será fazer normalização e a ortogonalização dos autovetores, para garantir solução única as componentes principais e, também, que estas sejam independentes umas das outras.

- Autovalores e autovetores de uma matriz

A solução desse problema, que está na base das análises ditas fatoriais, consiste em passar de uma matriz A de coeficientes de dependência (correlação ou

covariância) entre variáveis para uma matriz diagonal D , onde todos os coeficientes são nulos, exceto os da diagonal principal $\hat{\Lambda}$ (VALENTIN, 2000).

Segundo Magnusson & Mourão (2003, p.106), “estabelecendo-se algumas premissas importantes e usualmente improváveis, é possível determinar a posição dos eixos no espaço multidimensional usando-se a álgebra de matrizes”.

As análises baseadas nesse princípio são chamadas de análises de auto-vetores, sendo que “Eigen” é uma palavra da língua alemã, que significa “característica”.

O escalar $\hat{\Lambda}$ será chamado de autovalor, e o vetor \vec{x} um autovetor.

Seja S a matriz de variância-covariância quadrada $p \times p$, e I a matriz identidade $p \times p$, então os escalares $\hat{\Lambda}_1, \hat{\Lambda}_2, \dots, \hat{\Lambda}_p$ satisfazem a equação polinomial.

$$|S - \hat{\Lambda}I| = 0 \quad (2.3)$$

são chamados autovalores, ou raízes características, da matriz S .

Seja S a matriz de variância-covariância de dimensão $p \times p$, e seja $\hat{\Lambda}$ um autovalor de S , logo \vec{x} é um vetor não nulo ($x \neq 0$), tal que:

$$S\vec{X} = \hat{\Lambda}\vec{X}, \quad (2.4)$$

no qual, \vec{X} é uma matriz $p \times p$ de todos autovetores, e $\hat{\Lambda}$ é uma matriz $p \times p$ de todos autovalores.

Então, \vec{x} é dito autovetor, ou vetor característico da matriz S , associada com o valor $\hat{\Lambda}$.

Para determinar as componentes principais, a partir da matriz S , procede-se da seguinte forma:

a) Resolve-se a seguinte equação característica, para obter a solução:

$$|S - \hat{\Lambda}I| = 0, \text{ isto é,}$$

$$|S - \hat{\Lambda}I| = 0.$$

Sejam $\hat{\Lambda}_1, \hat{\Lambda}_2, \dots, \hat{\Lambda}_p$ as p soluções, tem-se que a cada autovalor $\hat{\Lambda}_i$ corresponde um autovetor característico.

$$\vec{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \cdot \\ \cdot \\ x_{ip} \end{bmatrix} \text{ com } \sum_{j=1}^p x_{ij}^2 = 1 \quad (\vec{x}_i^t \cdot \vec{x}_i = 1), \text{ sendo esta a condi\c{c}o\~{e}o de normalidade.}$$

e $\sum_{j=1}^p x_{ij}x_{kj} = 0$ para $i \neq k$ ($\vec{x}_i^t \cdot \vec{x}_k = 0$ para $i \neq k$), sendo esta a condi\c{c}o\~{e}o de ortogonalidade dos vetores (REGAZZI, 2001).

A normalidade \u00e9 a primeira restri\c{c}o\~{e}o feita para que o sistema tenha solu\c{c}o\~{e}o \u00fanica, e a segunda restri\c{c}o\~{e}o \u00e9 a ortogonalidade, que garante que as componentes principais s\u00e3o independentes.

Isso significa dizer que cada autovetor \u00e9 normalizado, ou seja, a soma dos quadrados dos coeficientes \u00e9 igual a 1, sendo, ainda, ortogonais entre si.

b) Conforme Regazzi (2001, p.7) para cada autovalor $\hat{\Lambda}_i$ determina-se o autovetor normalizado \vec{x}_i , a partir da solu\c{c}o\~{e}o do sistema de equa\c{c}o\~{e}es dado a seguir:

$$|S - \hat{\Lambda}I| \vec{x}_i = 0$$

$$\vec{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \cdot \\ \cdot \\ x_{ip} \end{bmatrix}, \text{ \u00e9 um autovetor n\u00e3o normalizado.}$$

\vec{o} \u00e9 um vetor nulo, de dimens\u00e3o $p \times 1$.

O autovetor normalizado \u00e9 dado por:

$$\vec{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \cdot \\ \cdot \\ x_{ip} \end{bmatrix} = \frac{1}{\sqrt{x_{i1}^2 + x_{i2}^2 + \dots + x_{ip}^2}} \begin{bmatrix} x_{i1} \\ x_{i2} \\ \cdot \\ \cdot \\ x_{ip} \end{bmatrix} = \frac{\vec{x}_i}{\|\vec{x}_i\|} \quad \vec{x}_i^t \vec{x}_i = 1. \quad (2.5)$$

Ao se estudar um conjunto de n observa\c{c}o\~{e}es de p -vari\u00e1veis, \u00e9 poss\u00edvel encontrar novas vari\u00e1veis denominadas de \hat{Y}_k , $k = 1, \dots, p$, que s\u00e3o combina\c{c}o\~{e}es

lineares (CL) das variáveis originais X_p , não correlacionados, e apresentam um grau de variabilidade diferente umas das outras, também apresentados em ordem decrescente de valores. É importante lembrar que, em componentes principais, a unidade de medida são combinações lineares não correlacionadas, por isso são de difícil interpretação, e também é por esse motivo que as variáveis originais devem estar na mesma unidade de medida.

Conforme Regazzi (2001, p.10), para se fazer uma interpretação correta de quais componentes utilizar no estudo, basta selecionar as primeiras componentes que acumulem uma percentagem de variância explicada, igual ou superior a 70%. Ou seja, fica-se com Y_1, \dots, Y_k tal que:

$$\frac{V\hat{a}r(Y_1) + \dots + V\hat{a}r(Y_k)}{\sum_{i=1}^p V\hat{a}r(Y_i)} .100 \geq 70\% \text{ no qual } k < p. \quad (2.6)$$

A partir da matriz S é possível encontrar os valores $\hat{\Lambda}_1 \geq \hat{\Lambda}_2 \geq \dots \geq \hat{\Lambda}_p \geq 0$, que são as raízes características, todas distintas e apresentadas em ordem decrescente de valores e, como S é positiva definida, todos os autovalores são não negativos.

Os eixos principais são os autovetores das matrizes SI ou RI , sendo que são os autovetores que fornecem a direção dos eixos na análise.

A Figura 9 mostra a elipse que possui dois eixos perpendiculares, cujas coordenadas estão representadas pelos autovetores I e II da matriz S , ou da matriz R . Os elementos desses vetores definem sua posição, isto é, o ângulo que eles formam com os eixos originais de Y_1 e Y_2 . O comprimento desses vetores são os autovalores correspondentes a $\hat{\Lambda}$, dessa matriz, que representa a variância dos novos eixos (VALENTIN, 2000).

A Figura 9 é a representação gráfica dos autovalores e autovetores.

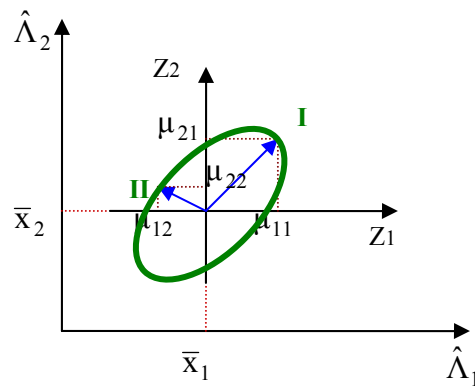


FIGURA 9 - Representação gráfica dos autovalores e autovetores.
Fonte : VALENTIN (2000).

Os eixos fatoriais CP são definidos pela direção e comprimento, através da seguinte equação característica: $|S - \hat{\Lambda}I| = 0$

S = matriz de variância-covariância, ou R a matriz de correlação.

$\hat{\Lambda}$ = autovalor de S , ou R .

I = matriz identidade.

Ao realizar uma análise de componentes principais, é muito importante saber o significado de cada componente no estudo que estiver sendo realizado.

A Tabela 1 mostra um resumo da análise de componentes principais, quais são os componentes principais, seus autovalores, seus autovetores, a correlação das variáveis, a percentagem de variância, explicada por cada componente, e a percentagem total da variância acumulada pelos componentes principais.

TABELA 1 – Componentes principais obtidas da análise de p variáveis X_1, X_2, \dots, X_p .

Componentes Principais	Variância explicada pelos	Coeficientes de ponderação associados às variáveis			Correlação entre X_j e Y_i			Porcentagem da variância de Y_i	Porcentagem acumulada da variância dos Y_i
	Autovalores $\hat{\Lambda}_i$	X_1	X_2	$\dots X_p$	X_1	X_2	$\dots X_p$		
Y_1	$\hat{\Lambda}_1$	x_{11}	x_{12}	$\dots x_{1p}$	$\sqrt{\hat{\Lambda}_1} \frac{x_{11}}{s_1}$	$\sqrt{\hat{\Lambda}_1} \frac{x_{12}}{s_2}$	$\dots \sqrt{\hat{\Lambda}_1} \frac{x_{1p}}{s_p}$	$\left(\hat{\Lambda}_1 / \sum_{i=1}^p \hat{\Lambda}_i \right) \cdot 100$	$\left(\hat{\Lambda}_1 / \sum_{i=1}^p \hat{\Lambda}_i \right) \cdot 100$
Y_2	$\hat{\Lambda}_2$	x_{21}	x_{22}	$\dots x_{2p}$	$\sqrt{\hat{\Lambda}_2} \frac{x_{21}}{s_1}$	$\sqrt{\hat{\Lambda}_2} \frac{x_{22}}{s_2}$	$\dots \sqrt{\hat{\Lambda}_2} \frac{x_{2p}}{s_p}$	$\left(\hat{\Lambda}_2 / \sum_{i=1}^p \hat{\Lambda}_i \right) \cdot 100$	$\left(\hat{\Lambda}_1 + \hat{\Lambda}_2 / \sum_{i=1}^p \hat{\Lambda}_i \right) \cdot 100$
.
.
.
.
Y_p	$\hat{\Lambda}_p$	x_{p1}	x_{p2}	$\dots x_{pp}$	$\sqrt{\hat{\Lambda}_p} \frac{x_{p1}}{s_1}$	$\sqrt{\hat{\Lambda}_p} \frac{x_{p2}}{s_2}$	$\dots \sqrt{\hat{\Lambda}_p} \frac{x_{pp}}{s_p}$	$\left(\hat{\Lambda}_p / \sum_{i=1}^p \hat{\Lambda}_i \right) \cdot 100$	$\left(\hat{\Lambda}_1 + \hat{\Lambda}_2 + \dots + \hat{\Lambda}_p / \sum_{i=1}^p \hat{\Lambda}_i \right) \cdot 100$

Fonte: REGAZZI (2001, p.12)

A Tabela 2 ilustra a substituição da matriz de dados originais (variáveis) por uma nova matriz, gerada após a análise, das componentes principais (escores para os componentes).

TABELA 2 – Escores relativos a n objetos (indivíduos), obtidos em relação aos k primeiros componentes principais.

Objetos (indivíduos)	Variáveis			Escores para os componentes		
	X_1	X_2	$\dots X_p$	Y_1	Y_2	$\dots Y_k$
1	x_{11}	x_{12}	$\dots x_{1p}$	y_{11}	y_{12}	$\dots y_{1k}$
2	x_{21}	x_{22}	$\dots x_{2p}$	y_{21}	y_{22}	$\dots y_{2k}$
.
.
.
n	x_{n1}	x_{n2}	$\dots x_{np}$	y_{n1}	y_{n2}	$\dots y_{nk}$

Fonte: REGAZZI (2001, p.13)

As combinações lineares obtidas através das CP's, segundo Jackson (1980), possuem a característica de que nenhuma combinação linear das variáveis originais irá explicar mais que a primeira componente e, sempre que se trabalhar com a matriz de correlação, as variáveis não sofrerão influência da magnitude de suas unidades medidas.

Resolvendo a matriz de correlação, pode-se observar se existe correlação entre as variáveis; se algumas variáveis iniciais forem linearmente dependentes umas das outras, alguns dos valores próprios serão nulos na matriz de correlação. Nesse caso, a variação total poderá ser explicada pelas primeiras componentes principais.

É difícil encontrar, em um problema, a existência de dependência linear exata, a menos que esta seja introduzida propositalmente nas variáveis redundantes. Na ACP pode ocorrer a dependência linear aproximada entre algumas variáveis. Nesse caso, os valores próprios menores são muito próximos de zero, e sua contribuição, para explicar a variância, será muito pequena (REIS, 1997). Por isso, deve-se retirar da análise aquelas componentes que possuem pouca informação. Isso não implica em uma perda significativa de informação.

Com isso, pode-se reduzir os dados e tornar os resultados mais fáceis de serem interpretados.

Além do uso na redução da dimensionalidade, a técnica de ACP pode ser utilizada como apoio à busca da variável de maior prevalência no sistema responsável, servindo-se do estudo dos coeficientes de correlação entre as componentes e as variáveis originais.

Quando se fala em avaliar a estabilidade de um processo produtivo, as dificuldades, que porventura existam, devem-se à complexidade do processo, e não aos métodos multivariados. A ACP é um recurso adicional de apoio para verificar a estabilidade do sistema Telhada (1995 apud SOUZA, 2000). O problema existente em um conjunto multivariado é que, às vezes, uma observação pode não ser extrema para uma determinada variável, mas pode ser considerada uma observação extrema por não ser semelhante à estrutura de correlação fornecida pelo restante dos dados.

A equação $r_{\hat{Y}_i, X_k} = \sqrt{\hat{\Lambda}_i} \frac{\hat{e}_{ki}}{\sqrt{S_{kk}}}$ deve ser utilizada quando os autovetores são

derivados da matriz de variância S , e a equação $r_{\hat{Y}_i, Z_k} = \hat{e}_{ki} \sqrt{\hat{\Lambda}_i}$ quando os autovetores são derivados da matriz de correlação R .

É importante lembrar que, embora as técnicas multivariadas, que constam na literatura, tenham sido criadas com o objetivo de resolver problemas específicos, como na área de biologia e psicologia, essas podem ser utilizadas para resolverem muitos outros problemas práticos nas diversas áreas do conhecimento. Na maioria das vezes, os objetivos desses problemas práticos só serão resolvidos mediante aplicação de mais de uma técnica multivariada, aplicadas em uma seqüência.

Dessa forma, é interessante ter-se uma visão global de todas, ou quase todas, técnicas multivariadas. Entre as técnicas multivariadas, mais utilizadas, estão: análise de agrupamentos, análise de componentes principais, análise de discriminante, análise de correspondência, dentre outras.

Conforme Reis (1997), a aplicação da ACP e AF deve incluir:

- As variáveis incluídas na análise;
- As percentagens da variância explicadas por cada uma das componentes principais;

- O número de componentes retidas e a proporção de variância total por elas explicada;
- Uma tabela com a contribuição de cada variável para cada componente (*factor loadings*), antes e depois de ser aplicado um método de rotação de fatores.
- Fazer a interpretação de cada componente principal retido.

2.3.3 Análise fatorial, relacionada à análise de componentes principais

A análise fatorial (AF) é formada por um conjunto de técnicas estatísticas, e possui como objetivo reduzir o número de variáveis iniciais com a menor perda possível de informação. Em outras palavras, poder-se-á dizer que AF é aplicada à busca de identificação de fatores num conjunto de medidas realizadas, sendo que esses fatores identificados pela AF serão uma descoberta feita pelo pesquisador. Inicialmente, ele terá várias medidas, e não será possível identificar quais variáveis poderão ser reunidas num fator. A AF é quem irá descobrir isso, pois ela permitirá identificar novas variáveis, em um número reduzido em relação às variáveis iniciais, sem uma perda significativa de informação contida nos dados originais.

A versão clássica da AF determina os fatores ortogonais, que descrevem aproximada e sucessivamente os vetores-resposta de n indivíduos a um conjunto constituído por m testes psicológicos. As primeiras pesquisas realizadas nessa área foram desenvolvidas por Karl Pearson (1901) e por Charles Spearman (1904). Spearman estudou a hipótese da existência de um só fator de inteligência, e da impossibilidade de medi-lo diretamente. Ele desenvolveu essa análise para que fosse possível estudar o fator inteligência, indiretamente, a partir das correlações entre diferentes testes. Em 1947 Thurstone partiu da idéia inicial de Spearman e desenvolveu a AF, por acreditar existir mais de um fator de inteligência. Essa análise permite identificar mais de um fator nos dados iniciais.

A AF não se refere, apenas, a uma técnica estatística, mas a um conjunto de técnicas relacionadas, para tornar os dados observados mais claros para a interpretação. Isso é feito analisando-se os inter-relacionamentos entre as variáveis, de

tal modo que essas possam ser descritas convenientemente por um grupo de categorias básicas, em número menor que as variáveis originais, chamado fatores.

Os fatores podem ser denominados como um constructo, que poderá ser uma variável não observada, escalas, itens, ou uma medida de qualquer espécie. Na análise, fatores explicam a variância das variáveis observadas, tal como se revelam pelas correlações entre as variáveis que estão sendo analisadas.

Um dos métodos mais conhecidos, para a extração dos fatores, é feito por meio da análise de componentes principais, que é baseado no pressuposto que se pode definir \bar{X} vetores estatisticamente não correlacionados, a partir de combinações lineares dos p indicadores iniciais.

A ACP permite transformar um conjunto de variáveis iniciais, correlacionadas entre si, num outro conjunto de variáveis não correlacionadas (ortogonais), que são as componentes principais, que resultam das combinações lineares do conjunto inicial.

Tanto a análise de componentes principais, quanto a análise fatorial, são técnicas da análise multivariada, que são aplicadas a um conjunto de variáveis, para descobrir quais dessas são mais relevantes, na composição de cada fator, sendo estes independentes um dos outros. Os fatores, que são gerados, são utilizados de maneira representativa do processo em estudo e utilizados para análises futuras.

O objetivo da ACP não é explicar as correlações existentes entre as variáveis, mas encontrar funções matemáticas, entre as variáveis iniciais, que expliquem o máximo possível da variação existente nos dados e permita descrever e reduzir essas variáveis. Já a AF explica a estrutura das covariâncias, entre as variáveis, utilizando um modelo estatístico casual e pressupondo a existência de p variáveis não-observadas e subjacentes aos dados. Os fatores expressam o que existe de comum nas variáveis originais (REIS, 1997).

A AF é uma técnica que é aplicada para identificar fatores num determinado conjunto de medidas realizadas, sendo utilizada, também, como uma ferramenta na tentativa de reduzir um grande conjunto de variáveis para um conjunto mais significativo, representado pelos fatores. Esse método determina quais variáveis pertencem a quais fatores, e o quanto cada variável explica cada fator.

Essas duas técnicas, ACP e AF, são sensíveis a correlações pobres entre variáveis, pois, neste caso, as variáveis não apresentarão uma estrutura de ligação entre elas. Logo, a correlação será fraca e prejudicará as análises, inviabilizando o uso da técnica, que tem como objetivo principal o estudo de conjuntos de variáveis correlacionadas.

Quando se trabalhar com AF, dever-se-á levar em consideração que coeficientes de correlação tendem a ser de menor confiança quando se fizer cálculos de estimativas de amostra pequenas. Em geral, para que seja possível trabalhar com análise, é necessário um mínimo de cinco casos, pelo menos, para cada variável observada.

O primeiro passo a ser realizado, quando se aplicar AF, é verificar as relações entre as variáveis, que poderá ser feito utilizando-se o coeficiente de correlação linear como medida de associação entre cada par de variáveis. Conforme Reis (1997), “a matriz de correlação poderá permitir identificar subconjuntos de variáveis que estão muito correlacionadas entre si no interior de cada subconjunto, mas pouco associadas a variáveis de outros subconjuntos”. Nesse caso, utilizar a técnica de AF permitirá concluir se é possível explicar esse padrão de correlações mediante um menor número de variáveis.

A AF é exploratória, pois é utilizada com o objetivo de reduzir a dimensão dos dados, podendo, também, ser confirmatória, se for utilizada para testar uma hipótese inicial de que os dados poderão ser reduzidos a uma determinada dimensão e de qual a distribuição de variáveis, segundo essa dimensão (REIS, 1997).

A ACP e a AF, quando utilizadas na forma direta, servirão para a identificação de grupos de variáveis inter-relacionadas e para a redução do número de variáveis. Em seu uso indireto, é um método que servirá para transformar dados. A transformação de dados ocorrerá através da reescrita dos mesmos, com propriedades que os dados originais não tinham.

Antes de aplicar a AF, dever-se-á levar em consideração certas premissas sobre a natureza dos dados. Primeiramente, o pesquisador deverá analisar a distribuição de freqüência das variáveis através de testes de ajuste da normalidade (Kolmogorov-Smirnov), ou, até, fazer um simples exame de curvas da distribuição. O

pesquisador poderá, ainda, fazer um gráfico de dispersão (*scatterplot*), fazendo um contraste em relação aos valores observados com os esperados numa distribuição normal (PEREIRA, 2001).

Há, também, uma medida de adequação dos dados, muito importante, sugerida por *Kaiser-Meyer-Olkin Measure of Adequacy (KMO)*. O *KMO* serve para avaliar o valor de entrada das variáveis para o modelo, sendo que seu valor possibilita prover resultados no alcance de 0,5 a 0,9. Se se obtiver valores nesse intervalo, então as variáveis poderão ser utilizadas para realizar a AF. Para encontrar o valor do *KMO*, utiliza-se a expressão:

$$KMO = \frac{\sum_i \sum_j r_{ij}^2}{\sum_i \sum_j r_{ij}^2 + \sum_i \sum_j a_{ij}^2}, \quad (2.7)$$

onde:

r_{ij} = é o coeficiente de correlação observado entre as variáveis i e j .

a_{ij} = é o coeficiente de correlação parcial entre as mesmas variáveis, que é, simultaneamente, uma estimativa das correlações entre os fatores. Os a_{ij} deverão estar próximos de zero, pelo fato de os fatores serem ortogonais entre si.

Quando as correlações parciais forem muito baixas, o *KMO* terá valor mínimo próximo a 1 e indicará perfeita adequação dos dados para análise fatorial. Segundo Pereira (2001, p.124 - 125), o teste do *KMO* possui valores que são considerados críticos como se pode observar:

- Para valores na casa dos 0,90: a adequação é considerada ótima para os dados da AF;
- Para valores na casa dos 0,80: a adequação é considerada boa para os dados da AF;
- Para valores na casa dos 0,70: a adequação é considerada razoável para os dados da AF;
- Para valores na casa dos 0,60: a adequação é considerada medíocre para os dados da AF;

- Para valores na casa dos 0,50 ou inferiores: a adequação é considerada imprópria para os dados da AF.

O *KMO* é uma medida de adequação que verifica o ajuste dos dados, utilizando todas as variáveis simultaneamente, e o seu resultado é uma informação sintética sobre os dados.

Outro teste, que poderá ser utilizado para análise fatorial, e também verifica as premissas, é o de *Bartlett Test of Sphericity (BTS)*, que testa a hipótese da matriz de correlação ser uma matriz identidade, ou seja, a diagonal principal igual a 1 e todos os outros valores serem zero, isto é, seu determinante é igual a 1. Isso significa que não há correlação entre as variáveis. A hipótese nula poderá ser rejeitada caso o α adotado será igual a 5% e o valor encontrado seja inferior ao valor de α . O teste de *Bartlett* na aplicação da ACP pressupõe que se rejeite a hipótese nula:

$$H_0 = P = I \text{ ou } H_0 = \hat{\Lambda}_1 = \hat{\Lambda}_2 = \dots = \hat{\Lambda}_p.$$

A análise de correspondência, a análise canônica e a análise fatorial discriminante são, também, métodos fatoriais, que levam a representações gráficas e terão, por isso, traços comuns com ACP. O que diferencia a ACP é que ela trata, exclusivamente, de variáveis numéricas, que desempenham, todas, o mesmo papel, enquanto a análise de correspondência trata de variáveis qualitativas, nas análises canônicas e discriminante as variáveis são repartidas em grupos bem distintos (BOUROCHE & SAPORTA, 1982).

A AF possui, como princípio, cada variável poder ser decomposta em duas partes: uma parte comum e uma parte única. A primeira é a parte da sua variação partilhada com outras variáveis, enquanto a segunda é específica da sua própria variação. Dessa forma, uma diferença entre os dois métodos parte do montante de variância analisada, na qual a ACP considera a variação total presente no conjunto das variáveis originais. Na AF, só é retida a variação comum, partilhada por todas as variáveis (REIS, 1997).

A base fundamental, para a análise de fator comum ACP e AF, é que as variáveis escolhidas poderão ser transformadas em combinações lineares de um conjunto de componentes (fatores) hipotéticos, ou despercebidos. Os fatores poderão ser associados com uma variável individual (fatores únicos), ou, ainda, associados com

duas ou mais das variáveis originais (fatores comuns). As cargas serão responsáveis por relacionarem a associação específica entre os fatores e as variáveis originais. Logo, poder-se-á concluir que o primeiro passo é encontrar as cargas e a solução para os fatores, que aproximarão a relação entre as variáveis originais e fatores encontrados, sendo que as cargas são derivadas dos autovalores, que estão associados às variáveis individuais.

Para ter-se uma melhor visualização das variáveis, que melhor representem cada fator, é realizada uma rotação nos eixos, pois a AF busca colocar os fatores em uma posição mais simples, com respeito às variáveis originais, que ajudam na interpretação de fatores. Essa rotação coloca os fatores em posições em que serão associadas só às variáveis relacionadas distintamente a um fator. Existem várias rotações que poderão ser realizadas para a matriz fatorial, *varimax*, *quartimax* e *equimax*. São todas as rotações ortogonais, enquanto as rotações oblíquas são não-ortogonais. A rotação *varimax rotation* busca minimizar o número de variáveis com altas cargas num fator, ou seja, maximiza a variância da carga e é, também, o mais utilizado. Conforme Pereira (2001), “a rotação da matriz não afeta a inércia (comunalidades) das variáveis nem a percentagem de variações explicadas pelos fatores”.

Antes de aplicar ACP e AF, o pesquisador deverá tomar duas decisões importantes, que são: o método a ser utilizado para a extração dos fatores e o número de fatores para serem extraídos.

Antes de se falar da interpretação da AF, é importante, ter-se claro, dois conceitos: o de ortogonalidade e o de carga fatorial.

O primeiro está relacionado com independência, no qual deve haver dissociação entre variáveis. E isso é conseguido quando se realiza a ACP, onde cada componente é independente da outra. Por isso, a ACP é, geralmente, utilizada como uma técnica para se extrair fatores.

O segundo conceito importante é o de carga fatorial. A matriz de cargas fatoriais é um dos passos finais da análise fatorial. A carga fatorial é um coeficiente: um número decimal, positivo ou negativo, geralmente menor do que um que expressa o quanto um teste, ou variável, observada, está carregado, ou saturado, em um fator.

Entre outras palavras, poder-se-á dizer que: quanto maior for a carga em cima de um fator, mais a variável se identifica com o que quer que seja o fator.

Em resumo, a AF é um método para determinar o número de fatores existente em um conjunto de dados, e serve para determinar quais testes, ou variáveis, pertencem a quais fatores.

A AF, em seus resultados, apresenta alguns conceitos que devem ser entendidos, para que haja uma interpretação correta dos dados. Como neste trabalho utiliza-se o *software statistica*, os resultados são apresentados com conceitos em língua inglesa. Conforme Pereira (2001), conceitos da AF:

- Autovalor: corresponde aos autovalores e à variância total, que pode ser explicada pelo fator. Ou seja, avalia a contribuição do fator ao modelo construído pela análise fatorial: se a explicação da variância pelo fator for alta, existe uma alta explicação desse fator ao modelo, se for baixa, existe uma baixa explicação do fator ao modelo.
- Peso do fator: é a proporção de variação da variável, que é explicada pelo fator, ou, ainda, o quanto cada variável contribui na formação de cada componente.
- Escore do fator: são os autovetores que definem as direções dos eixos da máxima variabilidade. Representam a medida assumida pelos objetos estudados na função derivada da análise.
- Comunalidade: é a medida de quanto da variância, de uma variável, é explicada pelos fatores derivados pela análise fatorial. Avalia a contribuição da variável ao modelo construído pela AF, ou seja, o quanto cada variável participa na formação da outra. Nas *communality*, os valores mais altos são os mais importantes para análise.
- Matriz fatorial: é a matriz de correlação entre as variáveis originais e os fatores encontrados.

Para que se possa nomear os fatores, dever-se-á olhar a pontuação dos mesmos, individualmente, e ver quais variáveis possuem as pontuações mais altas. Dever-se-á olhar, também, a pontuação do fator, para ver se as interpretações iniciais são confirmadas pela mesma.

A ACP adota a premissa de que a relação entre variáveis e fatores é linear. Dessa forma, poder-se-á tentar interpretar um eixo, seja graficamente, por regressão

linear, entre as coordenadas das amostras e os autovetores de cada variável, ou seja, pelo cálculo de um coeficiente de correlação não-paramétrico (Spearman, por exemplo).

Para que se possa resolver a equação característica, em AF, será necessário fazer a inversão de matriz, o que não será possível com uma matriz singular.

A multicolinearidade e singularidade são assuntos derivados de uma matriz de correlação, com alto grau de correlação entre as variáveis. A multicolinearidade acontece quando variáveis são altamente correlacionadas, ou seja, acima de 0.90, o que é muito bom para a AF, e a singularidade acontece quando as variáveis são perfeitamente correlacionadas. Com multicolinearidade, os efeitos são aumentados, as variáveis independentes estão inter-relacionadas. Se a variável é perfeitamente relacionada às outras variáveis, então a singularidade estará presente.

Raramente os resultados da AF serão todos publicados, pois nem todos possuem uma contribuição significativa para a interpretação dos dados e à elaboração de conclusões para o assunto que estiver sendo abordado.

Conforme Valentin (2000), as informações, que devem constar nas publicações, são:

- As dimensões da matriz de dados: número de variáveis e indivíduos;
- A natureza dos dados e as transformações eventuais;
- As figuras dos planos fatoriais;
- A necessidade de análises preliminares para testar a estabilidade e, se for preciso, eliminar certas variáveis ou observações.

2.4 Séries temporais

O estudos de séries temporais teve um grande impulso com o trabalho publicado por Box & Jenkins em 1976, que foi o marco histórico e científico da metodologia que estuda os modelos ARIMA. Desde então, autores como Gujarati (2000), Maddala (1992), Morettin (1987), Morettin (2004) se dedicam-se ao estudo desta metodologia, e nas suas diversas aplicações.

Os modelos de séries temporais possibilitam descrever um processo estocástico, utilizando apenas valores passados e presentes da variável. Assim, dada uma série y_t , os modelos de séries temporais poderão ser descritos como segue:

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}, \quad (2.8)$$

ou, ainda

$$(1 - \phi_1 B - \dots - \phi_p B^p) y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t, \quad (2.9)$$

onde B é o operador de defasagem ($By_t = y_{t-1}$), ϕ e θ são os parâmetros do modelo e ε_t o termo de erro. As variáveis defasadas de y_t representam a parte autoregressiva do modelo (AR), enquanto que as defasagens do termo de erro representam a parte de média móvel (MA). As equações (2.19) e (2.20) representam o modelo ARMA(p,q), onde p representa a ordem de defasagem do termo autoregressivo e q a ordem de defasagem do termo de média móvel. No caso dos modelos puramente sazonais, tem-se que:

$$(1 - \Phi_1 B^S - \dots - \Phi_{PS} B^{PS}) y_t = (1 - \Theta_1 B^S - \dots - \Theta_{QS} B^{QS}) \varepsilon_t, \quad (2.10)$$

onde Φ e Θ são os parâmetros do modelo sazonal, e S é o período sazonal. Como as séries temporais sazonais têm, em geral, componentes não sazonais, o modelo deve ser estimado como misto sazonal - não sazonal. A representação do modelo é feita como segue:

$$(1 - \phi_1 B - \dots - \phi_p B^p) (1 - \Phi_1 B^S - \dots - \Phi_{PS} B^{PS}) y_t = (1 - \theta_1 B - \dots - \theta_q B^q) (1 - \Theta_1 B^S - \dots - \Theta_{QS} B^{QS}) \varepsilon_t, \quad (2.11)$$

Assim, tem-se um modelo misto ARMA(p,q)(P,Q) $_S$. Onde p e q referem-se, respectivamente, às ordens autoregressiva e de média móvel; enquanto que P e Q referem-se, respectivamente, às ordens autoregressiva sazonal (SAR) e de média móvel sazonal (SMA). Para aplicar a metodologia de Box & Jenkins, usando este modelo, a série em estudo tem que ser estacionária. Ou seja, a série tem que ter média, variância e covariância finitas e constantes. O exame de estacionaridade sazonal será realizado observando-se a Função de Autocorrelação (ρ_k), FAC, da série em estudo. A referida função é dada por:

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (2.12)$$

onde \bar{y} é a média do processo e k é o número de defasagem das autocorrelações. A ocorrência de estacionaridade está relacionada a uma queda brusca no valor dos picos sazonais. A FAC também mostrará se há componente sazonal, o que pode ser observado quando esta segue um padrão de picos e depressões ano a ano. Da mesma forma, a FAC mostrará a não estacionaridade sazonal.

Se a série em estudo for não estacionária, com e sem sazonalidade, tem-se um ARIMA $(p,d,q)(P,D,Q)_s$, onde d representará a ordem de diferenciação não-sazonal e D a ordem de diferenciação sazonal. As diferenciações sazonal e não-sazonal podem ser realizadas da seguinte forma:

$$\Delta^d y = y_t - y_{t-1} \text{ (diferença)}$$

$$\Delta^D y = y_t - y_{t-s} \text{ (diferença sazonal).}$$

A diferenciação da série, em termos não-sazonal e sazonal, tem por objetivo tornar a mesma estacionária, o que possibilitará a aplicação da metodologia de Box & Jenkins.

Segundo Morretin & Tolo (1987), a construção dos modelos Box & Jenkins é baseada em um ciclo iterativo, no qual a escolha do modelo é feita com base nos próprios dados. Segundo Box & Jenkins (1976), três são as etapas para construção do modelo:

- Identificação: consiste em descobrir quais, dentre as várias versões dos modelos de Box & Jenkins, sejam sazonais ou não, descreverão o comportamento da série. A identificação do modelo a ser estimado ocorre pelo comportamento das funções de autocorrelações (FAC) e das funções de autocorrelações parciais (FACP).
- Estimação: consiste em estimar os parâmetros ϕ e θ do componente auto-regressivo, os parâmetros Φ e Θ do componente de médias móveis e a variância de ε_t .

- Verificação: consiste em avaliar se o modelo estimado é adequado para descrever o comportamento dos dados.

Pode vir a ocorrer não ser encontrado um modelo significativo. Nesse caso, dever-se-á retornar à fase de identificação. Como, em muitos casos, é possível encontrar vários modelos significativos, ou seja, modelos onde os resíduos encontrem-se dentro do intervalo de confiabilidade da função de autocorrelação residual (FACr) e função de autocorrelação parcial residual (FACPr), será necessário a utilização de outros critérios que auxiliem na tomada de decisão. Neste caso, serão utilizados os critérios AIC e BIC, quanto menor o AIC e o SBC, melhor o ajustamento do modelo. Poderá ocorrer de esses modelos apresentarem empates, então recorrer-se-á ao modelo que apresentar menor número de parâmetros. Ou, ainda, poderão existir modelos nos quais o AIC seja menor em um modelo e o BIC em outro, então a decisão a ser tomada será escolher o menor BIC, pois como $\ln(T)$ é sempre maior que 2, então o BIC será escolhido, sendo este, um modelo mais parcimonioso que o AIC.

Para encontrar o melhor modelo, com base na metodologia Box & Jenkins, dever-se-á ter por base os seguintes passos:

- Identificação

No processo de identificação busca-se determinar a ordem de (p,d,q) e $(P,D,Q)_S$, tomando-se por base o comportamento das Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP) e seus respectivos correlogramas.

Para o caso de séries estacionárias não-sazonais, a Tabela 3 fornece boas sugestões para a escolha de modelos $AR(p)$, $MA(q)$ ou $ARMA(p,q)$, com base nas propriedades das funções de autocorrelação (FAC) e Autocorrelação Parcial (FACP).

onde,

ρ_k = Autocorrelação;

ρ_{kk} = Autocorrelação Parcial;

k = Defasagem das FAC e FACP;

ϕ = Coeficiente do processo $AR(p)$;

- θ = Coeficiente do processo MA(q);
 p = Defasagem do processo AR;
 q = Defasagem do processo MA.

TABELA 3 - Padrões Teóricos das FAC e FACP.

PROCESSO	FAC	FACP
Ruído Branco	$\rho_k = 0, \forall k$	$\rho_{kk} = 0, \forall k$.
AR(1): $\phi_1 > 0$	Queda Exponencial Direta: $\rho_k = \phi_1^k$	$\rho_{11} = \rho_1$; $\rho_{kk} = 0, \forall k \geq 2$.
AR(1): $\phi_1 < 0$	Queda com Oscilação: $\rho_k = \phi_1^k$	$\rho_{11} = \rho_1$; $\rho_{kk} = 0, \forall k \geq 2$.
AR(p)	Queda em Direção a zero. Coeficientes podem oscilar.	Apresenta Pico na Defasagem p; $\rho_{kk} = 0, \forall k > p$.
MA(1): $0 < \theta_1$	Pico Positivo na defasagem 1. $\rho_k = 0, \forall k \geq 2$.	Queda com Oscilação: $\rho_{11} > 0$.
MA(1): $0 > \theta_1$	Pico Negativo na defasagem 1. $\rho_k = 0, \forall k \geq 2$.	Queda sem Oscilação: $\rho_{11} < 0$.
ARMA(1,1): $\phi_1 < 0$	Queda Exponencial, começando na defasagem 1.	Queda Exponencial, começando na defasagem 1.
	Sinal $\rho_1 = \text{Sinal}(\phi_1 + \theta_1)$	$\rho_{11} = \rho_1$; Sinal $\rho_{kk} = \text{Sinal} \rho_{11}$.
ARMA(p,q):	Queda, direta ou com oscilação, começando na defasagem q.	Queda, direta ou com oscilação, começando na defasagem p.

Fonte: MORETTIN (2004, p.157).

Como o objetivo é identificar o melhor modelo, fazer-se-á várias tentativas até encontrar modelos que sejam significativos, ou seja, que estejam representados dentro do intervalo de confiabilidade da FAC e FACP dos modelos encontrados.

- Estimação

Nesta fase da metodologia, cada um dos modelos, sugeridos na fase de identificação, são ajustados, e os vários coeficientes, ϕ , θ , Φ , Θ são examinados, em

termos de significância estatística. Os modelos ajustados serão comparados usando-se os seguintes critérios:

- Parcimônia

A incorporação de coeficientes adicionais melhora o grau de ajustamento do modelo, porém reduz os graus de liberdade do mesmo. Segundo Box & Jenkins (1976), os modelos mais parcimoniosos produzirão melhores previsões do que os modelos sobreparametrizados. Portanto, os modelos com o menor número de parâmetro serão os que melhor explicam a série em estudo.

- Estacionaridade e invertibilidade garantidas, com base nos coeficientes estimados

As FAC e FACP amostrais, e seus respectivos correlogramas, são gerados supondo que a série em estudo seja estacionária. Logo, os coeficientes estimados deverão ser consistentes com a hipótese de estacionaridade. Ou seja, no caso simples de um modelo (processo) AR(1), se ϕ_1 estimado for maior que 1, então a hipótese de estacionaridade será violada.

O modelo também tem que ser invertível, pois assume-se, implicitamente, nas FAC e FACP, que a série em estudo possa ser aproximada por um modelo autoregressivo. Dessa forma, valores estimados de θ , que sugerem não invertibilidade, geralmente estarão associados a modelos mal especificados.

- Diagnóstico do modelo ajustado com base nos resíduos

A análise dos resíduos de modelos ajustados é de extrema importância na escolha final do modelo que melhor explica a dinâmica da série temporal em estudo. Se os resíduos forem autocorrelacionados, então a dinâmica da série em estudo não será completamente explicada pelos coeficientes do modelo ajustado. Deve-se excluir do processo de escolha modelo com esta característica. Uma análise da existência da

autocorrelação serial de resíduos será feita com base nas Funções de Autocorrelação dos Resíduos e Autocorrelação Parcial e seus respectivos correlogramas, nos quais estes deverão apresentar os resíduos todos dentro do seu intervalo de confiabilidade.

Os resíduos também deverão apresentar-se com média zero, variância constante e seguindo a distribuição normal. Por esse motivo o pesquisador conseguirá encontrar diversos modelos que contemplarão estas características, mas que nem sempre oferecerão as melhores previsões. Logo utilizar-se-á os critérios de seleção de modelos ou critérios, penalizadores, que auxiliem na decisão entre quais dos modelos concorrentes encontrados dever-se-á utilizar.

- Critério de seleção de modelos

A identificação dos modelos é mais difícil nas etapas da metodologia de Box & Jenkins. Não são raros os casos em que se consegue identificar um único modelo, e sim vários modelos candidatos a gerador da série em estudo, estes modelos podem ser chamados de modelos concorrentes, isso porque, trabalhando com a FAC e a FACP amostrais, fica difícil, muitas vezes, decidir se elas estão decrescendo ou se são truncadas.

Para escapar desse impasse, muitos pesquisadores preferem utilizar um outro procedimento de identificação, que depende menos do julgamento de quem está analisando a série de tempo. Esse procedimento faz uso de critérios de seleção de modelos construídos com base na variância estimada ε_t , no tamanho da amostra e nos valores de p e q. Os mais utilizados são o critério AIC e o critério BIC.

$$AIC = T \ln (\text{soma do quadrado dos resíduos}) + 2n;$$

$$SBC = T \ln (\text{soma do quadrado dos resíduos}) + n \ln(T);$$

onde,

n = número de parâmetros estimados;

T = número de observações utilizadas.

Ao invés de estabelecer p e q , precisamente, estima-se os modelos correspondentes a vários pares (p, q) , e escolhe-se aquela especificação que apresentar o menor valor para AIC, ou para BIC.

A presença de p e q , nas fórmulas dos critérios AIC e BIC, tem por objetivo “penalizar” os modelos com muitos parâmetros, tendo em vista que modelos mais parcimoniosos deverão ser privilegiados, por apresentarem menor número de parâmetros a ser estimado.

A crítica que se faz, a esses critérios, é que eles podem conduzir a modelos super especificados, ou seja, modelos com valores de p e/ou q maiores do que o correto.

Assim sendo, os critérios AIC e BIC deverão ser usados como procedimento complementar, e não alternativo àquele baseado na FAC e na FACP.

- Previsão

A previsão é uma das principais razões da popularidade da metodologia de Box & Jenkins. Em muitos casos, as previsões, principalmente de curto prazo, obtidas com base em Box & Jenkins, são melhores que as obtidas com base nos modelos econométricos tradicionais.

Tendo em vista que a metodologia de Box & Jenkins tem sua capacidade de previsão comprometida no longo prazo, convém, sempre que possível, atualizar as previsões já realizadas.

Comentários deste capítulo

Nesse capítulo 2, abordou-se a análise de agrupamentos, análise de componentes principais, análise fatorial e séries temporais que servirão de base para o pleno desenvolvimento da aplicação prática.

No capítulo 3, apresentar-se-à metodologia utilizada, bem como as técnicas aplicadas.

3 METODOLOGIA

Neste capítulo abordar-se-á todos os procedimentos utilizados para desenvolver esta pesquisa.

Para a realização deste trabalho foi fundamental uma revisão de literatura sobre o desenvolvimento da bovinocultura nacional, nos últimos anos, pois, essencialmente, desse comportamento é que surgirão as conclusões e interpretações do trabalho. A busca por um conjunto de dados, que refletissem a produção de bovinos, no Brasil, foi fundamental para o desenvolvimento do mesmo. Após a revisão de literatura, foi conduzida uma análise descritiva dos dados, para verificar o comportamento dos mesmos e, logo a seguir, um estudo sobre PDCA, AM e ST foi realizado, o que possibilitou a fundamentação desta pesquisa.

Considerando a bovinocultura brasileira como um sistema produtivo, lança-se mão do ciclo PDCA, que servirá de base para o gerenciamento deste processo, visto que é uma ferramenta do sistema de produção, pois a cada etapa, ou ciclo produtivo, desta atividade é possível cumprir as etapas *Plan*, *Do*, *Check* e *Action*, que significam:

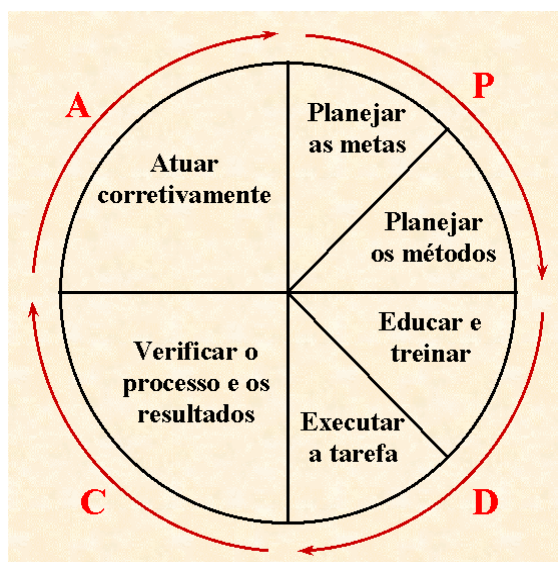


FIGURA 10 - Ciclo PDCA.
Fonte: LOPES (2004, p.31).

O método gerencial, do ciclo PDCA, vem sendo estudado por autores como Ishikawa, K. (1989) e Campos, V. F. (1992, 1994, apud WERKEMA, 1995).

Como se deseja conhecer esta atividade da bovinocultura, lançou-se mão de algumas técnicas, tais como, de AM e de ST, nas quais as etapas metodológicas para a realização deste trabalho são, as que estão inseridas na etapa de planejamento do ciclo PDCA. Com o propósito de gerar subsídios para a tomada de decisão viu-se que o preço do boi gordo é o que impulsiona a cadeia produtiva da criação de gado no Brasil. E, como cada região apresenta características próprias, decidiu-se realizar previsões do preço, para cada região, mas diante do grande número de regiões produtoras, faz-se inicialmente uma triagem entre os expoentes de produção baseado no número de cabeças produzidas, por meio da AM e posteriormente parte-se para a previsão do preço do boi gordo por meio de ST, para as regiões selecionadas.

As etapas metodológicas são:

3.1 Elaboração do banco de dados

Para a elaboração do banco de dados utilizou-se de critérios rígidos, pois buscou-se trabalhar com variáveis que melhor refletissem a bovinocultura nacional. O banco de dados é composto por oitenta variáveis (80), que representam as categorias de bovinos no período de 1997 a 2004, por estados.

O banco de dados inicial foi elaborado no programa Microsoft Excel e transportado, na íntegra, para o *software Statistica, versão 7.0*, o qual possibilitou obter os resultados buscados para esta pesquisa.

3.2 Análise descritiva dos dados

Após a coleta e a elaboração do banco de dados ter sido realizada, uma estatística descritiva foi proferida, para verificar o comportamento das variáveis ao longo do período em estudo.

Para isso, calculou-se a média individual de cada variável, anualmente, o que serviu de suporte para análises subseqüentes, na confirmação de resultados obtidos.

Logo, em seguida, fez-se os gráficos de linhas, que também servirão de suporte para análises subseqüentes. Nesta pesquisa, optou-se em representar, graficamente, aquelas variáveis de maior influência para o estudo.

3.3 Análise multivariada

3.3.1 Análise de agrupamentos

Após a análise descritiva dos dados, uma análise de agrupamentos foi utilizada, para mostrar quais as categorias de bovinos estão associadas, pois esta agrupa variáveis e objetos que possuí características semelhantes. Variáveis, ou objetos, que possuam características diferentes, formarão grupos distintos dos demais. Essa análise foi utilizada, também, para verificar as oscilações ocorridas na bovinocultura, no decorrer período.

Para aplicar as técnicas da AM, utilizou-se variáveis numéricas referentes às categorias de bovinos, no período de 1997 a 2004, sendo essas variáveis abreviadas, para melhor visualização das mesmas, mediante representação gráfica. O nome e as abreviaturas estão descritas conforme segue:

- Touro (TOU);
- Vaca (VAC);
- Novilho de 1 a 2 anos (N 1 a 2);
- Novilho de 2 a 3 anos (N 2 a 3);
- Bezerras (Bas);
- Bezerros (Bos);
- Garrotes de 1 a 2 anos (G 1 a 2);
- Garrotes de 2 a 3 anos (G 2 a 3);
- Bois de 3 a 4 anos (B 3 a 4);
- Bois com mais de 4 anos (B + 4).

Na AA, inicialmente aplicou-se a análise a todas as variáveis e em todo o período em estudo, Se existir uma mudança estrutural na criação de bovinos, esta será

detectada com as técnicas da análise multivariada, caso isso ocorra a análise será realizada em períodos diferentes.

3.3.2 Análise de componentes principais e análise fatorial

Para identificar as variáveis de maior relevância, para a pesquisa, utilizar-se-á a análise de componentes principais, que possibilitará identificar o número adequado de componentes a serem utilizados, que explicarão a amostra em estudo, os quais devem revelar um grau de explicação acima de 70%, atendendo, assim, às pressuposições da ACP.

Para complementar a ACP, utilizou-se a AF, que tem como objetivo identificar novas variáveis, em um número menor que o conjunto inicial, mas sem ocorrer perda significativa de informação. Com esta análise, será possível gerar novas variáveis que expressassem a informação contida no conjunto original de dados. Essas técnicas ACP e AF, auxiliarão na identificação dos fatores de maior relevância para a bovinocultura nacional.

Caso ocorra uma ruptura na estrutura da bovinocultura, no período analisado, sendo esta detectada pela AA, estas duas técnicas da ACP e a AF, também serão realizadas em dois períodos. Através dessas técnicas, será possível identificar os estados que se destacam na produção nacional de bovinos, bem como as categorias que estão associadas aos mesmos.

Com a ACP e AF, buscar-se-á reduzir esta dimensionalidade, fazendo com que se obtenha os melhores resultados, com a menor perda de informação.

Para desenvolver esta pesquisa será utilizada uma amostra composta por oitenta (80) variáveis, as quais representaram as categorias de bovinos e de vinte sete (27) objetos, os quais representaram os estados brasileiros, no período de 1997 a 2004.

3.3.3 Séries temporais

Como se está a analisar um período de oito anos, e durante o mesmo o Brasil sofreu oscilações na economia, isso fez, também, com que todos os segmentos desta

sofressem junto, inclusive a bovinocultura. Devido a isso, monitorou-se a cotação do boi gordo em dólares, sendo que os dados foram coletados junto aos anuários da pecuária brasileira (ANUALPEC), no período de 1997 a 2004, naqueles estados que se destacaram nacionalmente. Utilizou-se, para isso, a metodologia de Box & Jenkins, com o intuito de procurar estimar o modelo adequado da classe geral dos modelos ARIMA, a série e sua estimação mediante previsões para os anos de 2005 e 2006.

Para identificar os modelos apropriados, inicialmente analisou-se o gráfico da série original. A análise desse gráfico indicou a presença de tendência, ou alteração, na variância, o que revela se a série é, ou não, estacionária. Para a utilização da metodologia Box&Jenkins necessita-se que a série seja estacionária. Após a análise visual, da série original, foram analisadas as funções de autocorrelações (FAC) e de autocorrelações parciais (FACP). O comportamento dessas funções indica qual o modelo a ser utilizado, mediante a FACP, bem como auxiliar na confirmação da estacionariedade através da FAC. Caso não se comprove a estacionariedade da série, passa-se a aplicar diferenças, objetivando-se torná-la estacionária.

Logo a seguir, encontrou-se o melhor modelo a ser utilizado para determinar as previsões. Após encontrar vários modelos concorrentes, isto é, aqueles que apresentam ruído branco, optou-se em utilizar aqueles que apresentaram as melhores estatísticas para os critérios do AIC e BIC. A situação ideal é: quanto menor o AIC, e o BIC, melhor o ajustamento do modelo. Porém, faz-se necessário comparar os AICs e os BICs de modelos alternativos, para que se saiba qual o modelo que melhor explique a dinâmica da série temporal em estudo. Para comprovar a validade do modelo escolhido, testou-se a normalidade dos resíduos, mediante gráfico de dispersão. A partir disso, foi possível certificar-se do melhor modelo que represente a série, tornando-se possível, assim, realizar as previsões desejadas.

Decidiu-se monitorar o preço do boi gordo, pois ele afeta, de forma direta, a produção nacional de bovinos, pois se o preço do boi estiver dando um retorno satisfatório para o pecuarista, e órgãos de fomento de incentivo à produção, o pecuarista fará, dessa, sua principal atividade econômica, caso contrário abandonará a mesma, passando a apostar em outra atividade econômica.

3.4 Geração de resultados

Em busca de incentivos fiscais, para que sejam tomadas medidas que venham em auxílio do desenvolvimento das regiões menos favorecidas, não abandonando aquelas que tenham a bovinocultura como uma de suas principais atividades econômicas, esse trabalho será divulgado em congressos nacionais e internacionais, revistas especializadas, bem como será enviado uma cópia ao Instituto Brasileiro de Geografia e Estatística (IBGE), ao Instituto (iFNP), sendo esses órgãos responsáveis pela disseminação dos conhecimentos, aqui gerados, e por possuírem uma ampla circulação nos órgão gestores de política de desenvolvimento.

Comentários deste capítulo

Nesse capítulo apresentou-se a metodologia utilizada, bem como as técnicas aplicadas.

4 APLICAÇÃO DA METODOLOGIA E RESULTADOS

No capítulo 4, item 4.1, apresenta-se uma análise descritiva dos dados, que possibilita verificar o comportamento dos mesmos. No item 4.2 e 4.5, aplica-se a técnica da análise de agrupamentos, que consiste na reunião de elementos semelhantes, ou seja, variáveis (categorias de animais), ou estados, que possuam características em comum serão reunidos em um mesmo grupo. No item 4.3 e 4.6, apresenta-se a aplicação da análise fatorial a qual possibilitará reduzir a dimensão do problema com a menor perda de informação. Nos itens 4.4 e 4.7 aplica-se a análise de componentes principais, a qual possibilita identificar os estados e o tipo de produção que ocorre nos mesmos, bem como aqueles que apresentaram destaque no período de 1997 a 2005. No item 4.8 apresenta-se a previsão do preço do boi gordo, sendo que os valores referentes a esta variável estão em dólar e o peso em arroba, mediante séries temporais, no período analisado, sendo que essa é aplicada apenas nos estados apontados como destaque, pela AM. Fez-se, também, a previsão do preço do boi gordo para o ano de 2005.

4.1 Análise descritiva

Para que fosse possível um melhor entendimento das variáveis, e verificar o comportamento das mesmas, procedeu-se, inicialmente, uma análise estatística descritiva.

Na Tabela 6, verifica-se as médias do número de cabeças de bovinos vivos, representadas por cada categoria, no período de 1997 a 2004. Analisando-se a Tabela 6, observar-se-á que estas médias encontram-se entre 17.772 e 2.270.427 cabeças, ou seja, a menor média é encontrada na categoria dos bois com mais de quatro anos, no ano de 2004. A maior média é representada pela categoria das vacas, no ano de 2002.

Com a análise descritiva dos dados foi possível verificar que a menor média encontrada, em todo o período, é representada pela categoria de bois com mais de quatro anos de idade, e que essa categoria está reduzindo com o passar tempo, pois no ano de 1997 possuía-se uma média de 80.151 cabeças, e no ano de 2004 teve-se

uma média de apenas 17.772 cabeças, ou seja, ocorreu uma diminuição significativa em relação a essa categoria de bois. Isso mostra que pode estar havendo o emprego de novas tecnologias no campo, as quais possibilitem a produção de bovino precoce, conforme Tabela 4.

TABELA 4 - Estatística descritiva referente à média do número de cabeças de bovinos vivos, representadas por cada categoria, no período de 1997 a 2004

Variáveis	Ano	Média do n° de cabeças	Variáveis	Ano	Média do n° de cabeças	Variáveis	Ano	Média do n° de cabeças
TOU	1997	80.254,00	Bas	1997	696.604,37	B3a4	1997	178.838,93
TOU	1998	85.258,11	Bas	1998	711.771,70	B3a4	1998	190.326,00
TOU	1999	83.527,48	Bas	1999	760.566,63	B3a4	1999	182.718,48
TOU	2000	83.319,04	Bas	2000	782.076,44	B3a4	2000	152.836,52
TOU	2001	80.434,07	Bas	2001	815.270,15	B3a4	2001	102.406,78
TOU	2002	84.062,00	Bas	2002	825.736,59	B3a4	2002	97.322,56
TOU	2003	83.656,74	Bas	2003	811.340,07	B3a4	2003	98.254,56
TOU	2004	85.897,81	Bas	2004	859.378,93	B3a4	2004	92.425,15
VAC	1997	1.711.120,81	Bos	1997	692.250,00	B+4	1997	80.151,44
VAC	1998	1.774.402,33	Bos	1998	710.097,96	B+4	1998	80.542,96
VAC	1999	1.896.235,07	Bos	1999	763.247,63	B+4	1999	71.367,70
VAC	2000	1.960.262,04	Bos	2000	783.812,93	B+4	2000	59.317,93
VAC	2001	2.202.230,41	Bos	2001	811.753,85	B+4	2001	24.438,44
VAC	2002	2.270.427,00	Bos	2002	820.957,22	B+4	2002	24.007,44
VAC	2003	2.252.818,11	Bos	2003	806.333,67	B+4	2003	19.939,41
VAC	2004	2.221.557,70	Bos	2004	854.881,74	B+4	2004	17.772,26
N 2 a 3	1997	393.708,70	G1a2	1997	565.366,63			
N 2 a 3	1998	416.088,52	G1a2	1998	562.897,78			
N 2 a 3	1999	413.944,41	G1a2	1999	598.268,74			
N 2 a 3	2000	436.981,07	G1a2	2000	597.366,41			
N 2 a 3	2001	455.490,59	G1a2	2001	581.864,30			
N 2 a 3	2002	462.993,48	G1a2	2002	594.242,70			
N 2 a 3	2003	464.806,11	G1a2	2003	608.530,19			
N 2 a 3	2004	495.502,96	G1a2	2004	611.218,85			
N 1 a 2	1997	629.476,70	G2a3	1997	400.271,56			
N 1 a 2	1998	626.325,41	G2a3	1998	408.167,93			
N 1 a 2	1999	647.836,33	G2a3	1999	396.595,48			
N 1 a 2	2000	665.448,96	G2a3	2000	382.211,37			
N 1 a 2	2001	709.977,78	G2a3	2001	304.899,41			
N 1 a 2	2002	709.254,22	G2a3	2002	311.444,85			
N 1 a 2	2003	727.422,33	G2a3	2003	310.918,30			
N 1 a 2	2004	748.491,63	G2a3	2004	314.852,70			

Outra categoria, que teve decréscimo significativo, foi a representada por bois de três e quatro anos de idade, tendo uma média de 178.838,93 cabeças no ano de 1997 e 92.425,15 cabeças, no ano de 2004. Em contrapartida, teve-se um aumento expressivo no número de cabeças de novilhos entre dois e três anos, sendo uma média

de 393.708,70 cabeças no ano de 1997, e 495.502,96 cabeças no ano de 2004, isso mostra o crescimento do abate jovem no país.

Para que seja possível um melhor entendimento dessas categorias abordadas, observe-se as Figuras 11 a 18.

Para que houvesse uma melhor visualização, nestes gráficos, foram eliminados aqueles estados que possuem baixa representatividade, em relação à produção nacional de bovinos, conforme Figura 11:

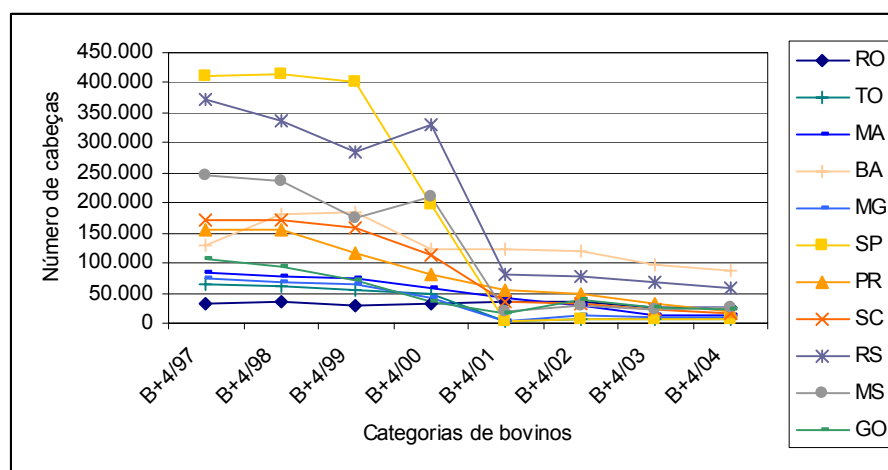


FIGURA 11 – Gráfico referente à categoria dos bois com mais de 4 anos no período de 1997 a 2004.

A Figura 12 mostra que a categoria representada pelas vacas, em onze estados, que se destacam pela quantidade de matrizes. Em primeiro lugar está o MS, que teve um pequeno decréscimo de 1997 a 1998, tendo um aumento significativo entre os anos de 1998 a 2002, tendo, novamente, um decréscimo, em sua produção, nos anos de 2002 e 2003, mantendo-se constante de 2003 a 2004. Em segundo lugar está MG, que teve um pequeno aumento de 1997 a 1999. Manteve-se constante de 1999 a 2000, tendo novamente um aumento entre o ano de 2000 e 2001, manteve-se novamente constante de 2001 a 2003, e teve uma queda entre 2003 e 2004. Em terceiro lugar, o estado do MT, que teve um aumento gradativo entre os anos de 1997 a 2003, tendo uma pequena queda de 2003 a 2004, conforme Figura 12.

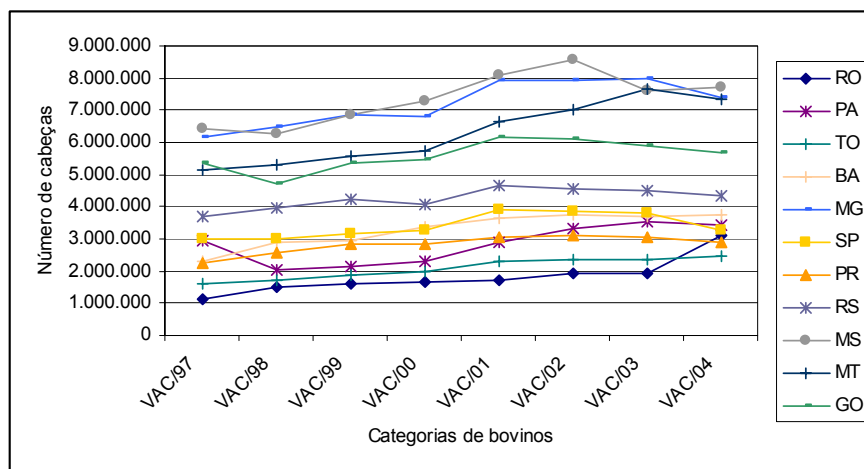


FIGURA 12 – Verificação do número de vacas no período de 1997 a 2004, nos estados brasileiros.

Ainda analisando a Figura 12, é possível verificar que o estado de GO está em quarto lugar, tendo um decréscimo entre 1997 e 1998, tornando a ter um aumento gradativo de 1998 a 2001, decaindo, novamente, de 2001 a 2004. Em quinto lugar está o RS, que teve um pequeno aumento de 1997 e 1999, tendo um pequeno decréscimo de 1999 a 2000. Tornou a crescer de 2000 a 2001, tornando a ter um pequeno decréscimo de 2001 a 2004. Os estados que não se destacaram, de forma significativa, em relação a variável analisada, não serão mencionados na análise descritiva.

A Figura 13 mostra que o estado do MS, de 1999 a 2002, liderou na produção nacional de bezerros, tendo uma queda de 2002 a 2004. Logo, em seguida, está o estado de MG, que sofreu uma queda no período de 2003 a 2004. Oscilando, entre o terceiro e quarto lugares, estiveram os estados de GO e MT, sendo que o estado de MT teve aumentada sua produção, significativamente, no período de 2000 a 2004, sendo que no ano de 2004, possuiu a maior produção do país. Oscilando, entre o quinto e sexto lugar, estão os estados de SP e RS. Oscilando, entre o sétimo, oitavo e nono lugares, estão os estados do PR, PA e a BA, conforme Figura 13.

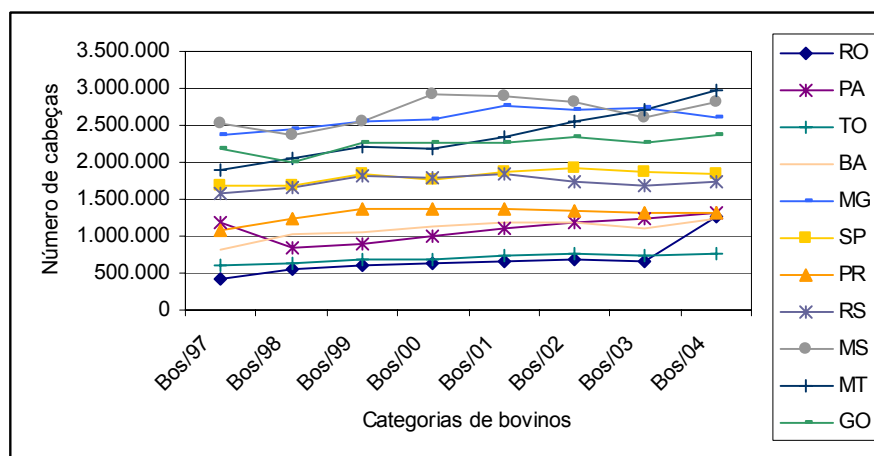


FIGURA 13 – Gráfico referente categoria representada pelos bezerros, no período de 1997 a 2004.

A Figura 14 mostra os estados que melhor representaram a produção nacional de bezerras.

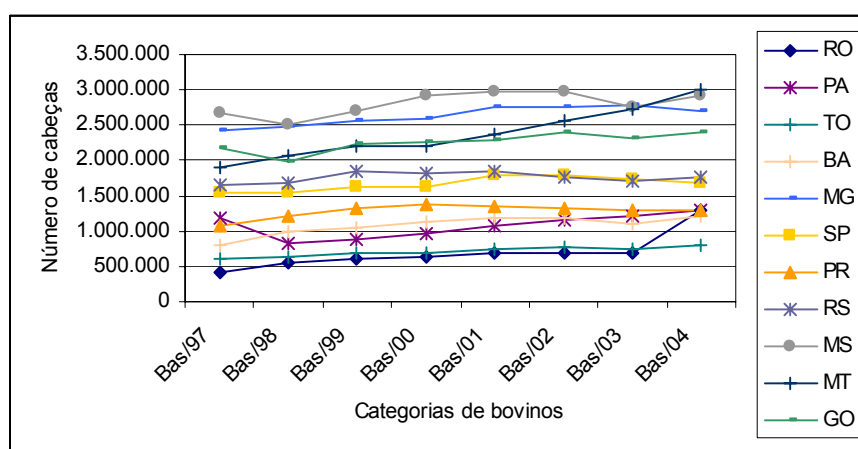


FIGURA 14 – Gráfica referente categoria representada pelas bezerras, no período de 1997 a 2004.

As Figuras 13 e 14 apresentaram que essa produção teve as mesmas características no período de 1997 a 2004, ou seja, os estados que possuíram destaque são os mesmos e as oscilações ocorridas no período também são as mesmas.

A Figura 15 mostra os estados que melhor representaram a produção nacional de novilhos e novilhas de 2 a 3 anos. Como pode-se observar, a produção nacional de

novilhos e novilhas ocorre nos mesmos estados que possuem destaque na produção de bezerros e bezerras.

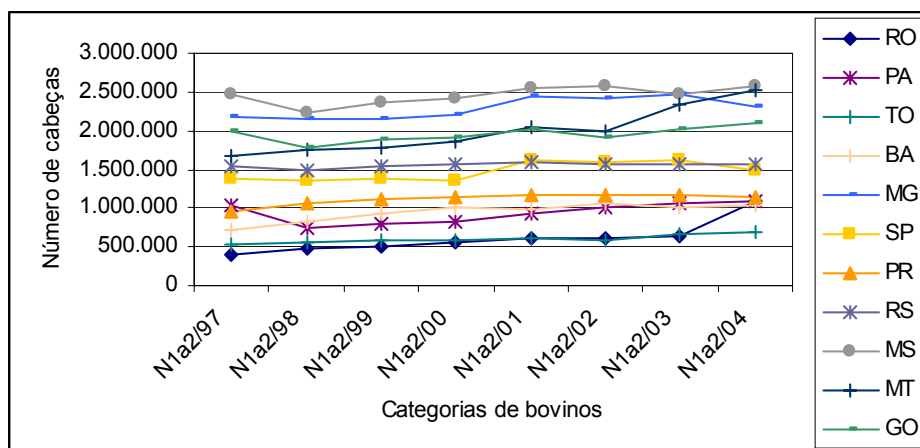


FIGURA 15 – Gráfico referente categoria representada pelos novilhos, no período de 1997 a 2004.

A Figura 16 mostra os estados que melhor representaram a produção nacional de garrotes de 1 a 2 anos.

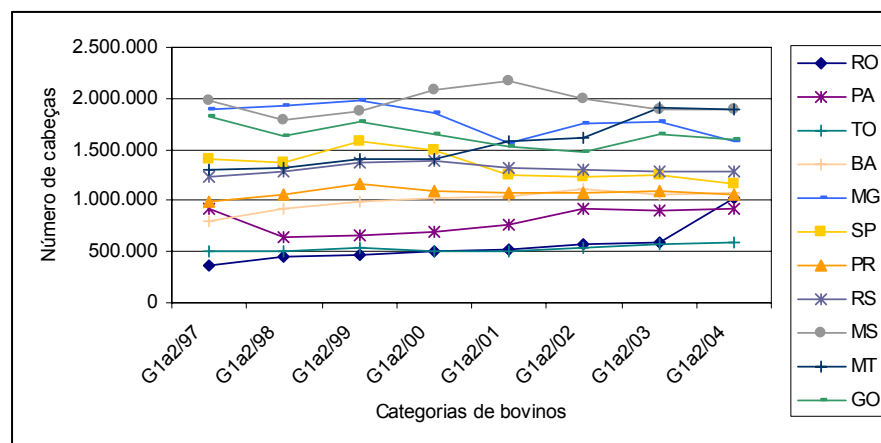


FIGURA 16 – Gráfico referente categoria representada pelos garrotes de 1 a 2 anos, no período de 1997 a 2004.

Pode-se verificar, na Figura 16, que a variável garrotes de 1 a 2 anos sofreu várias oscilações de 1997 a 2004. Desses estados, vale destacar o MT, pois este possuiu um crescimento gradativo ao longo de todo o período, com um aumento

significativo entre 2002 e 2003. Outro estado, que vale destacar, é SP, que no período de 1997 a 2000 representou o quarto lugar na produção, tendo uma queda gradativa de 2000 a 2004, passando a representar o sexto lugar na produção nacional de garrotes de 1 a 2 anos.

A Figura 17 mostra os estados que melhor representaram a produção nacional de garrotes de 2 a 3 anos. Pode-se verificar que os estados de MG, GO, MS e SP, tiveram uma redução significativa nessa produção, principalmente entre 2000 e 2001. Os estados do MT e RO tiveram um aumento gradativo, em sua produção, durante todo o período. Com isso, observa-se que os principais estados, que representam a produção nacional de bovinos, estão diminuindo a produção de garrotes de 2 a 3 anos.

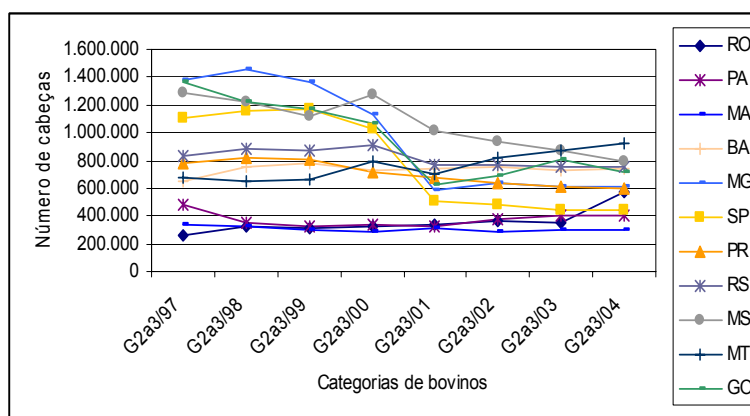


FIGURA 17 – Gráfico referente categoria representada pelos garrotes de 2 a 3 anos, no período de 1997 a 2004.

Pode-se verificar, com essa análise descritiva, que os estados, que se destacam na quantidade de matrizes, lideram, também, a produção de bezerras, bezerras, novilhos e novilhas de 2 a 3, garrotes de 1 a 2 anos e garrotes de 2 a 3 anos, ou seja, são auto-suficientes e responsáveis pela cria, recria e engorda de animais.

A Figura 18 mostra a média de cada estado referente à produção anual de bovinos.

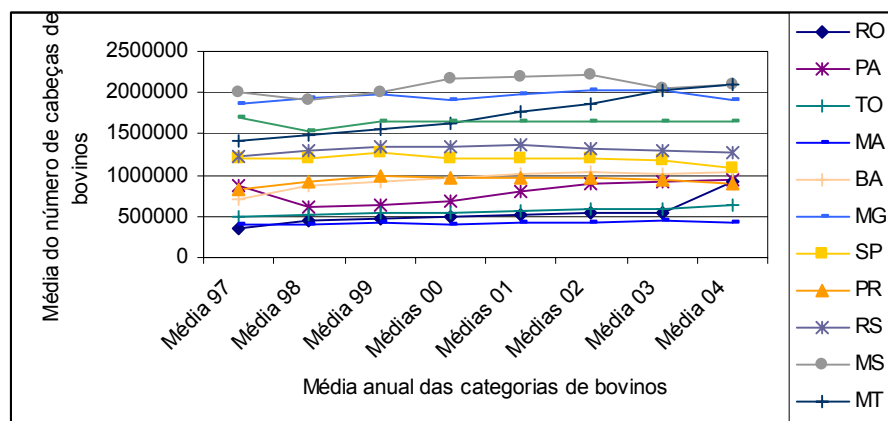


FIGURA 18 – Gráfico referente à média de produção de cada estado, no período de 1997 a 2004.

Analisando-se a Figura 18, pode-se observar que onze (11) estados brasileiros possuíram destaque, no período de 1997 a 2004, na bovinocultura nacional: MS, MG, MT e GO, seguidos do RS, SP, BA, PR, PA, TO e RO, sendo que desses sete estados, vale destacar o estado do MT, que teve um crescimento significativo em todo o período. O PA, tendo um crescimento de 1998 a 2004 e RO, que se destacou de 2003 a 2004, ultrapassando a produção dos estados de TO e PR.

Após a análise descritiva dos dados, passou-se a utilizar a análise multivariada, a qual possibilitou identificar se a estrutura pecuária no Brasil sofreu alterações na produção durante o período analisado, bem como identificar a vocação de cada região, em relação produção nacional de bovinos.

4.2 Análise de agrupamentos de 1997 a 2000

A primeira técnica multivariada a ser utilizada foi à análise de agrupamentos, pois através desta foi possível identificar as variáveis com características comuns, ou seja, que formaram um mesmo grupo. Serviram, também, para auxiliar na caracterização dos fatores que representaram essas variáveis iniciais.

O algoritmo de agrupamento, utilizado para formar esses grupos, é a distância euclidiana, sendo que o método de encadeamento é o *single linkage*, já citado

anteriormente no item 2.2. Os valores utilizados, para constituir a amostra em estudo, não foram padronizados, pois estão na mesma unidade de medida.

A Figura 19 mostra o dendograma, no qual a escala vertical indica a distância entre os grupos, e no eixo horizontal são marcadas as variáveis, na ordem em que são agrupadas. As linhas verticais partem das variáveis e têm altura correspondente ao nível em que as variáveis são consideradas semelhantes.

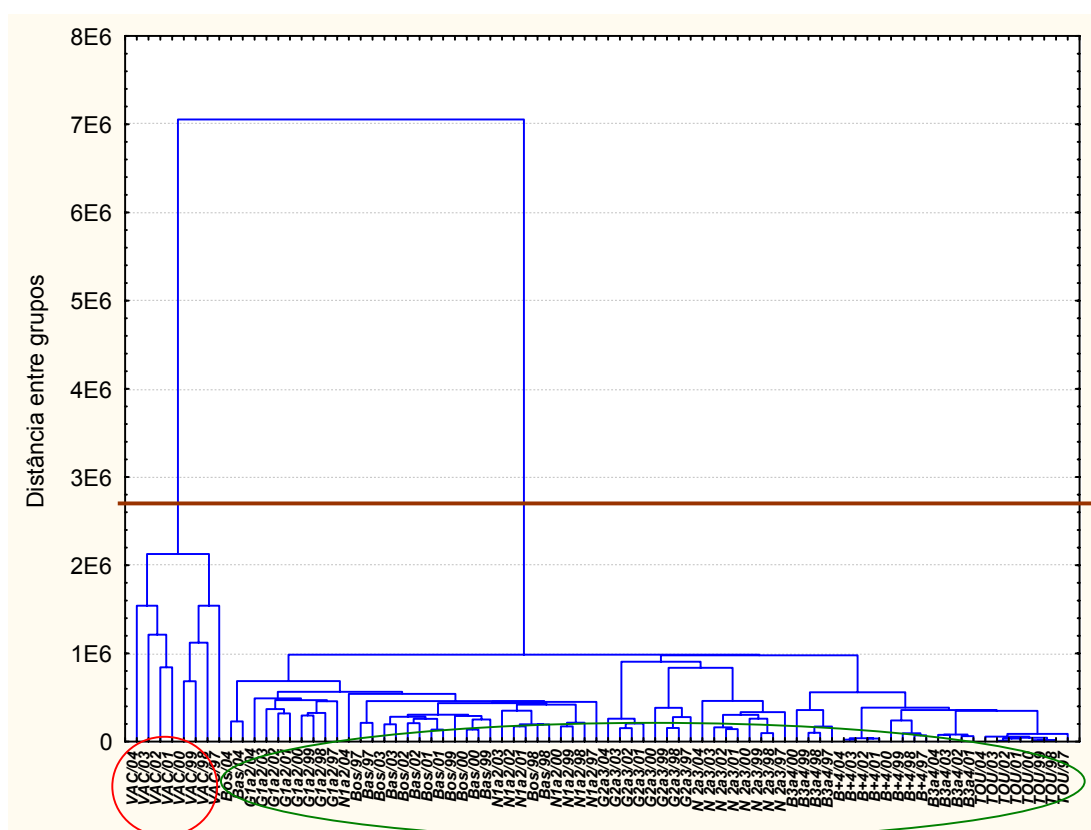


FIGURA 19 - Dendrograma da matriz de distâncias, pelo método de agrupamento por ligação simples, no período de 1997 a 2000.

Analisando a FIGURA 19, verifica-se que o maior salto encontra-se entre a distância 2×10^6 e 7×10^6 , no gráfico referido como 2E6 e 7E6, respectivamente. Fazendo-se um corte no gráfico, entre essas distâncias, ter-se-á dois grupos homogêneos distintos. O primeiro grupo é formado pelas variáveis: touros, bois de 3 a 4 anos, bois + de 4 anos, novilhos de 2 a 3 anos, garrotes de 2 a 3 anos, novilhos de 1 a 2 anos, bezerras, bezerros e garrotes de 1 a 2 anos, representando, estas variáveis, o

maior grupo do dendograma. O segundo grupo é formado pela variável representada pelas vacas, que está contida no círculo, ao lado da elipse. Analisando o grupo, formado pelas vacas, pode-se verificar que no ano de 1997 essa variável representava um alto número de cabeças, tendo uma queda no ano de 1998, e outra queda significativa no ano de 1999, mantendo-se constante no ano de 2000. Só no ano de 2001 passou a recompor o número de matrizes, mantendo-se constante no ano de 2002. Teve aumento significativo no ano de 2003 e outro no ano de 2004, representando, neste ano de 2004, uma população de vacas superior ao ano de 1997.

Observa-se, ainda, na Figura 19, que no período de 1997 a 2004 ocorreu uma quebra na estrutura bovina nacional, ou seja, o período de 1997 a 2000, possuiu uma característica diferente da ocorrida no período de 2001 a 2004, sendo que no primeiro período houve um decréscimo no número de matrizes, já no segundo ocorreu um aumento das mesmas. Essa análise deve ser proferida em dois períodos, em relação ao número de matrizes, pois são estas que geram o rebanho nacional, sendo que através das mesmas torna-se possível aumentar, ou diminuir, a produção de bovinos. Devido a isso, serão realizadas duas análises, a primeira referente ao período de 1997 a 2000 e a segunda referente ao período de 2001 a 2004.

A partir da Figura 20, as análises de AA, ACP e AF passam a ser realizadas em dois períodos, sendo que a Figura 20 mostra o dendograma das categorias de bovinos no período de 1997 a 2000.

Analisando-se a Figura 20, verifica-se que, realizando-se um corte no gráfico entre a distância 1×10^6 e 2×10^6 , no gráfico referido como 1E6 e 2E6, respectivamente, têm-se cinco grupos homogêneos distintos. Como no dendograma ainda não existe uma teoria que defina exatamente em qual distância deva ser feito o corte, fica a critério do pesquisador a decisão de onde este será realizado, pois o mesmo deve ser conhecedor da problemática que está estudando.

O primeiro grupo é formado pelas variáveis: touros, bois de 3 a 4 anos, bois + de 4 anos, representando os reprodutores e os animais para o abate. O segundo grupo, formado por novilhos de 2 a 3 anos, garrotes de 2 a 3 anos, representando o abate jovem de animais. O terceiro grupo é formado por novilhos de 1 a 2 anos, bezerras, bezerros e garrotes de 1 a 2 anos, que está representando a cria e engorda de animais.

O quarto grupo é formado pela variável representada pelo número de vacas, que no ano de 1997 formou um grupo isolado, por ser o maior número de animais encontrados em todo o período. O quinto grupo é também formado pela variável representada pelo número de vacas, no período de 1998 a 2000. Mediante o dendograma, é possível verificar a queda ocorrida no número de matrizes, no período de 1997 a 2000, conforme Figura 20.

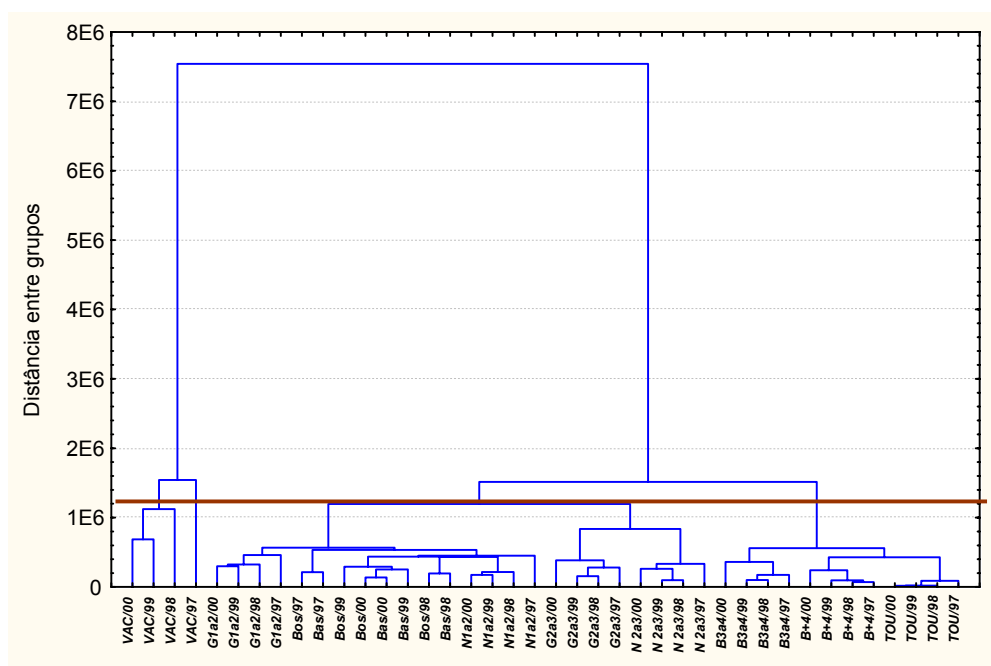


FIGURA 20 – Dendrograma com as variáveis no período de 1997 a 2000, pelo método de agrupamento por ligação simples.

A Figura 21 mostra os estados que se destacaram na produção de bovinos no período de 1997 a 2000.

Fazendo-se um corte no gráfico, da Figura 21, entre as distâncias, 2×10^6 , e $2,5 \times 10^6$ no gráfico referido como 2E6 e 2,5E6, ter-se-á cinco grupos homogêneos distintos. O primeiro, e maior, representado pelos estados RO, MA, TO, AC, AM, RN, AL, SE, RR, AP, DF, PB, PI, ES, PE, RJ, CE, SC e PA, no período de 1997 a 2000.

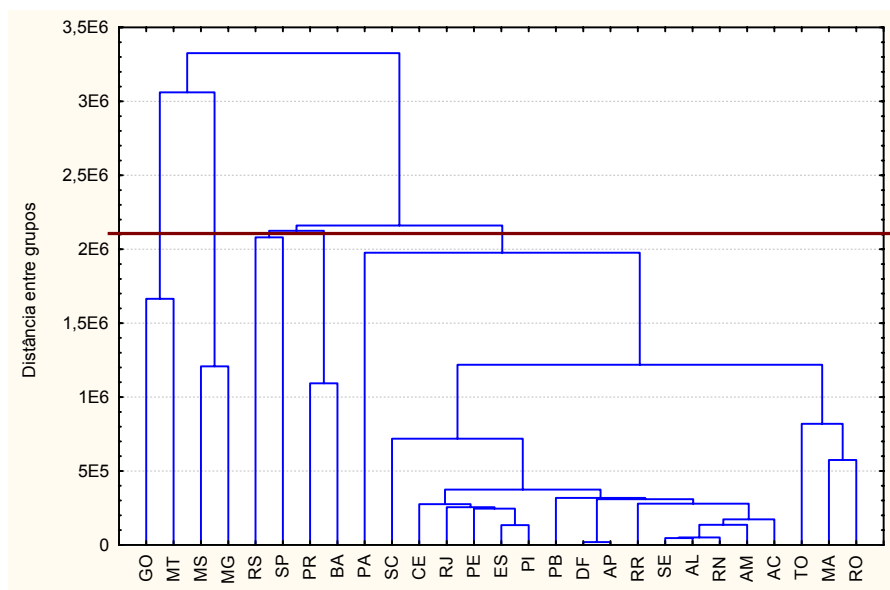


Figura 21 - Dendrograma referente aos estados, utilizando o método de agrupamento de ligação simples, no período de 1997 a 2000.

Conforme a Figura 21, esses estados possuíram baixa representatividade em relação à bovinocultura nacional, não representando, esta, a principal atividade econômica para os mesmos. Os grupos formados pelos estados: BA e PR, SP e RS, MG e MS, MT e GO, a bovinocultura representa um importante segmento para a economia dessas regiões, sendo que esses estados destacam-se na bovinocultura nacional, ou seja, são os que melhor representaram esta atividade econômica. Devido a isso, formaram grupos distintos no dendrograma.

4.3 Análise fatorial e de componentes principais de 1997 a 2000

Para aplicar as técnicas da análise multivariada, tais como a AF e ACP, faz-se necessário o cálculo da matriz de correlação entre as variáveis estudadas, para verificar se existe correlação entre as mesmas. Caso não exista correlação entre essas variáveis, torna-se inviável o uso da análise fatorial e de componentes principais.

Com a matriz de correlação, da Tabela 5, é possível observar que a maioria dos valores são superiores a 0,5, o que significa que a correlação entre as variáveis está de moderada a forte. Sendo assim, pode-se concluir que as variáveis estão

interligadas umas às outras, não sendo necessário realizar um teste de significância para as correlações, pois, dessa forma, tem-se a certeza de que estas variáveis estão correlacionadas. Isso mostra que o estudo das variáveis não deve ser feito de forma isolada, mas, sim, de maneira conjunta, com a utilização de uma técnica adequada, neste estudo, a AF e ACP.

A Tabela 5 mostra a correlação de parte das variáveis originais. Como o número dessas é relativamente elevado, optou-se por apresentar, aqui, apenas a correlação das primeiras variáveis, referente ao ano de 1997, sendo que as demais variáveis também possuem correlação alta, ou seja, o resultado é análogo a Tabela 5.

TABELA 5 – Matriz de correlação entre as variáveis, no período de 1997 a 2000

	TOU/97	VAC/97	N 2a3/97	N1a2/97	Bas/97	Bos/97	G1a2/97	G2a3/97	B3a4/97	B+4/97
TOU/97	1,00	1,00	1,00	0,99	0,99	0,99	0,98	0,93	0,80	0,52
VAC/97		1,00	1,00	0,99	0,99	0,99	0,98	0,94	0,80	0,52
N 2a3/97			1,00	0,99	0,99	0,98	0,97	0,94	0,78	0,47
N1a2/97				1,00	1,00	1,00	0,99	0,96	0,85	0,60
Bas/97					1,00	1,00	0,99	0,96	0,85	0,60
Bos/97						1,00	1,00	0,97	0,86	0,61
G1a2/97							1,00	0,99	0,89	0,63
G2a3/97								1,00	0,92	0,67
B3a4/97									1,00	0,88
B+4/97										1,00

Para que seja possível aplicar a AF, deve-se determinar o número de autovalores a serem considerados na análise, bem como o percentual de explicação de cada autovalor. Na Tabela 6, apresenta-se o resultado dos autovalores, bem como a porcentagem de variância explicada por cada componente, e, também, a variância acumulada pelas mesmas.

Numa análise fatorial, considerando-se 40 variáveis, poder-se-ia ter 40 fatores que corresponderiam às variáveis originais. Para que seja possível reduzir a dimensionalidade do problema, ocasionando a menor perda de informação possível, recorreu-se a dois critérios da AF. O primeiro critério inclui, na análise, aquelas componentes que conseguem sintetizar uma variância acumulada em torno de 70%. Como se pode observar, na Tabela 6, os dois primeiros autovalores representam cerca de 98,00% da variância. Portanto, os dados serão resumidos pelos dois primeiros autovalores. Pode-se, também, fazer esta seleção incluindo-se somente aquelas

componentes cujos valores próprios são superiores a 1. Neste caso, são dois autovalores, este critério foi sugerido por Kaiser (1960 apud MARDIA et al., 1979), conforme Tabela 6.

TABELA 6 – Autovalores e percentual da variância explicada de cada componente, no período de 1997 a 2000

Número de componentes	Autovalores Extração dos componentes principais			
	Autovalores	% da variância explicada	Autovalores acumulados	% da variância explicada acumulada
1	35,82	89,56	35,82	89,56
2	3,38	8,44	39,20	98,00
3	0,47	1,17	39,67	99,17
4	0,14	0,35	39,81	99,52
5	0,07	0,18	39,88	99,70
6	0,05	0,13	39,93	99,83
7	0,03	0,08	39,97	99,91
8	0,02	0,04	39,98	99,96
9	0,01	0,02	39,99	99,97
10	0,00	0,01	39,99	99,99
11	0,00	0,01	40,00	99,99
12	0,00	0,00	40,00	100,00

Analisando-se a Tabela 6, pode-se verificar que os dois primeiros autovalores correspondem a 89,56% e 8,44% da variância total explicada pelos autovalores do modelo, ou seja, explicam, juntos, 98,00% das variações das medidas originais. A perda de informação, nesse caso, é de apenas 2,00%. Conforme Pereira (2001), “essa é uma medida de ajuste do modelo à análise de dados: no exemplo, o modelo com dois fatores terá 98,00% de representação real dos dados”.

Após a extração dos autovalores, deve-se decidir pelo número de fatores a serem utilizados na análise. Para isso, recorreu-se ao método gráfico sugerido por Cattell (1966), que consiste em selecionar apenas aqueles fatores que possuam uma

queda no gráfico, ou seja, aqueles fatores que estiverem antes de ocorrer uma estabilização no mesmo.

A Figura 22 mostra a seleção dos fatores através do método gráfico *Scree Plot*, sendo que a porcentagem de variação explicada pela componente está no eixo das ordenadas, e os autovalores estão representados em ordem decrescente no eixo das abcissas.

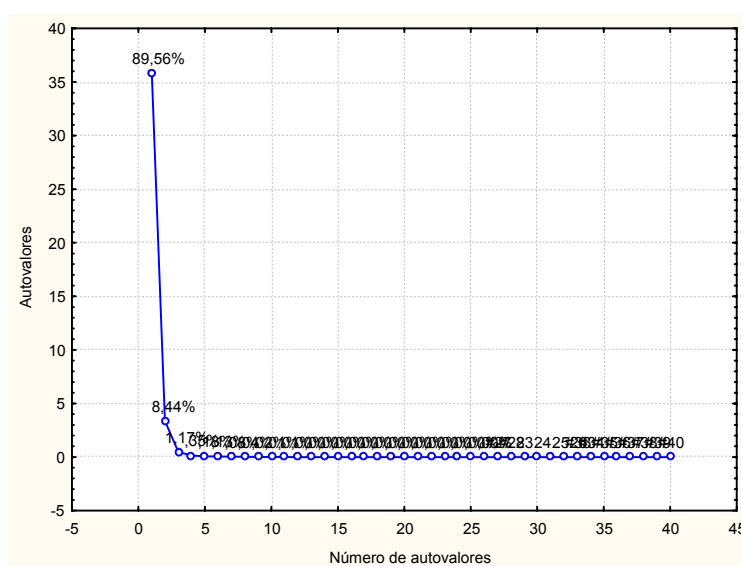


FIGURA 22 - Gráfico de explicação da proporção de variação de cada componente principal, no período de 1997 a 2000.

Como se pode observar, na Figura 22, as duas primeiras componentes explicam 98,00% da variância total, havendo uma estabilização do gráfico após o segundo fator. Sendo assim, considera-se apenas os dois primeiros fatores. Pode-se observar, também, que os outros fatores apresentam uma baixa explicação, não sendo aconselhável incluí-los na análise.

Após ter definido quantos fatores serão utilizados na análise, deve-se verificar a importância de cada variável na composição dos fatores, ou seja, encontrar a matriz de cargas fatoriais, a qual mostra que quanto maior for a carga em cima de um fator, mais a variável se identifica com este.

Conforme Pereira (2001), “o passo final da AF é verificar se os fatores, que são dimensões abstratas, podem ser interpretados de forma coerente com a natureza dos fenômenos estudados”. Para isso, deve-se analisar a matriz de cargas fatoriais, na qual estão os *factor loadings*, e verificar quais as variáveis que melhor se correlacionam com cada fator. A Tabela 7 mostra a matriz de cargas fatoriais, sem realizar a rotação nos eixos.

TABELA 7 – Cargas fatoriais para composição dos fatores, no período de 1997 a 2000.

	Fator 1	Fator 2
TOU/97	-0,97	-0,20
VAC/97	-0,97	-0,21
N 2a3/97	-0,96	-0,25
N1a2/97	-0,99	-0,11
Bas/97	-0,99	-0,12
Bos/97	-0,99	-0,10
G1a2/97	-0,99	-0,06
G2a3/97	-0,98	0,02
B3a4/97	-0,90	0,39
B+4/97	-0,68	0,72
TOU/98	-0,98	-0,18
VAC/98	-0,98	-0,17
N 2a3/98	-0,97	-0,22
N1a2/98	-0,99	-0,09
Bas/98	-0,99	-0,10
Bos/98	-0,99	-0,08
G1a2/98	-1,00	-0,03
G2a3/98	-0,97	0,07
B3a4/98	-0,89	0,40
B+4/98	-0,65	0,75
TOU/99	-0,98	-0,19
VAC/99	-0,98	-0,18
N 2a3/99	-0,97	-0,23
N1a2/99	-1,00	-0,09
Bas/99	-0,99	-0,10
Bos/99	-0,99	-0,07
G1a2/99	-0,99	0,00
G2a3/99	-0,97	0,10
B3a4/99	-0,87	0,43
B+4/99	-0,61	0,78
TOU/00	-0,97	-0,19
VAC/00	-0,98	-0,18
N 2a3/00	-0,96	-0,23
N1a2/00	-0,99	-0,10
Bas/00	-0,99	-0,10
Bos/00	-0,99	-0,08
G1a2/00	-1,00	0,00
G2a3/00	-0,99	0,08
B3a4/00	-0,89	0,41
B+4/00	-0,67	0,61

Na Tabela 7 pode-se visualizar as ponderações de cada variável que irão compor a combinação linear. Observe-se que os valores em destaque são os que possuem uma significância maior do que 0,7.

O ideal é identificar, em cada combinação linear, um conjunto de variáveis que representa este fator e, a partir daí, atribuir-se um nome para o fator. Essa abstração, para o fator, passa a identificá-lo, representando um conjunto de variáveis. Quando essa identificação ficar difícil, por apresentar mais de um grupo de variáveis significativas no mesmo fator, ou em fatores diferentes, recorre-se à realização de rotações, pois, desta forma, mantém-se a mesma inércia no conjunto analisado, mas os eixos são rotacionados, possibilitando uma melhor visualização da disposição dos pontos. A rotação utilizada foi a *Varimax normalizada*, pois esta mantém os eixos perpendiculares entre si, ou seja, ortogonais. A Tabela 8 mostra o resultado dos *Factor Loadings*, após a rotação *varimax normalized*, a qual possibilitou uma melhor visualização dos fatores, nos quais a proporção de variação das variáveis está melhor representada. Observa-se que os valores que possuem uma significância igual, ou superior, a 0,7 estão em destaque em cada fator.

Neste estudo, utilizar-se-á apenas dois fatores que possuem as variáveis explicativas, pois através do método gráfico sugerido por Cattell (1966), esses fatores explicam a maior variância.

Pode-se concluir, ainda, que o fator 1 é o mais importante para o estudo, pois é derivado do maior autovalor e possui uma explicação de 89,56%, sendo que as variáveis, que mais contribuem, neste, são representadas pelas seguintes variáveis: TOU/97, VAC/97, N 2a3/97, N 1a2/97, Bas/97, Bos/97, G 1a2/97, G 2a3/97, TOU/98, VAC/98, N 2a3/98, N 1a2/98, Bas/98, Bos/98, G 1a2/98, G 2a3/98, TOU/99, VAC/99, N 2a3/99, N 1a2/99, Bas/99, Bos/99, G 1a2/99, G 2a3/99, TOU/00, VAC/00, N 2a3/00, N 1a2/00, Bas/00, Bos/00, G 1a2/00, G 2a3/00. Logo, pode-se denominar este fator de cria, cria e engorda de animais. O fator 2 é derivado do segundo autovalor e possui uma explicação de 8,44%, sendo que as variáveis, que mais contribuem na formação desse fator, são as seguintes: B 3 a 4/97, B + 4/97, B 3 a 4/98, B + 4/98, B 3 a 4/99, B + 4/99, B 3 a 4/00 e B + 4/00. O fator 2 pode ser denominado de abate, pois as variáveis que compõe esse fator representam o abate de bovinos, conforme Tabela 8.

Para que haja uma melhor visualização desses fatores, optou-se em utilizar os gráficos de dispersão, ou os planos fatoriais, que examinam a localização das variáveis num sistema de coordenadas criado pelos fatores.

TABELA 8 – Cargas fatoriais para composição dos fatores, no período de 1997 a 2000.

	Fator 1	Fator 2
TOU/97	0,95	0,28
VAC/97	0,96	0,28
N 2a3/97	0,97	0,23
N1a2/97	0,92	0,37
Bas/97	0,92	0,36
Bos/97	0,92	0,37
G1a2/97	0,89	0,41
G2a3/97	0,83	0,46
B3a4/97	0,60	0,75
B+4/97	0,26	0,96
TOU/98	0,95	0,30
VAC/98	0,95	0,31
N 2a3/98	0,96	0,26
N1a2/98	0,92	0,39
Bas/98	0,92	0,38
Bos/98	0,91	0,40
G1a2/98	0,88	0,43
G2a3/98	0,80	0,50
B3a4/98	0,57	0,76
B+4/98	0,22	0,97
TOU/99	0,95	0,30
VAC/99	0,95	0,31
N 2a3/99	0,96	0,26
N1a2/99	0,92	0,39
Bas/99	0,92	0,38
Bos/99	0,90	0,41
G1a2/99	0,86	0,46
G2a3/99	0,79	0,52
B3a4/99	0,55	0,78
B+4/99	0,17	0,98
TOU/00	0,95	0,30
VAC/00	0,95	0,30
N 2a3/00	0,96	0,25
N1a2/00	0,92	0,38
Bas/00	0,92	0,38
Bos/00	0,91	0,39
G1a2/00	0,87	0,47
G2a3/00	0,83	0,53
B3a4/00	0,60	0,78
B+4/00	0,32	0,88

Os fatores a serem relacionados, neste primeiro plano fatorial, são: Fator 1 x Fator 2, conforme Figura 23.

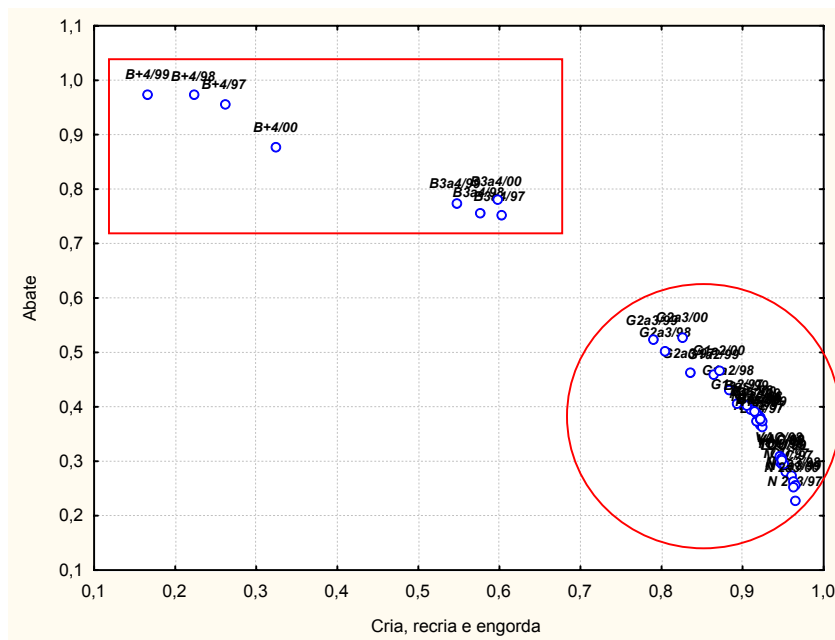


FIGURA 23 - Gráfico representando a relação entre fatores fator 1 x fator 2, no período de 1997 a 2000.

Na Figura 23 pode-se analisar todas as variáveis utilizadas na análise, sendo que, apenas aquelas que melhor explicam o fator 1 e 2, trarão a devida contribuição para o estudo, de forma que se possa identificar quais dessas possuem uma maior representatividade nos planos fatoriais.

É importante salientar que a cria, recria e engorda é composta de trinta e duas variáveis, que possuem variância igual, ou superior a 0,7. Sendo assim, esse é o fator mais importante para análise.

As variáveis que melhor representam a cria, recria e engorda, são as que melhor explicam o fator 1, ou seja, as que estão mais distantes da origem, em relação ao eixo das abscissas, representadas pelo círculo. As variáveis que melhor representam o abate, são B 3 a 4 e B + 4, no período de 1997 a 2000. Essas variáveis estão contidas no retângulo, ou seja, as que estão mais distantes da origem, em relação ao eixo das ordenadas, sendo essas as que melhor explicam o fator 2.

Após a realização da análise fatorial, com a determinação de quantos fatores serão utilizados na análise, e o que estes representam, procedeu-se com a análise de componentes principais.

A interpretação dos componentes principais é, sem dúvida, um dos pontos mais delicados da análise. Aqui, dois aspectos devem ser explorados: o primeiro é a correlação das variáveis originais com as componentes, que fornecerão a contribuição de maior relevância. O outro aspecto é realizar a análise considerando os indivíduos que estão sendo estudados.

A Figura 24 mostra o círculo de correlação unitário, com a nuvem de variáveis, referente ao primeiro plano principal.

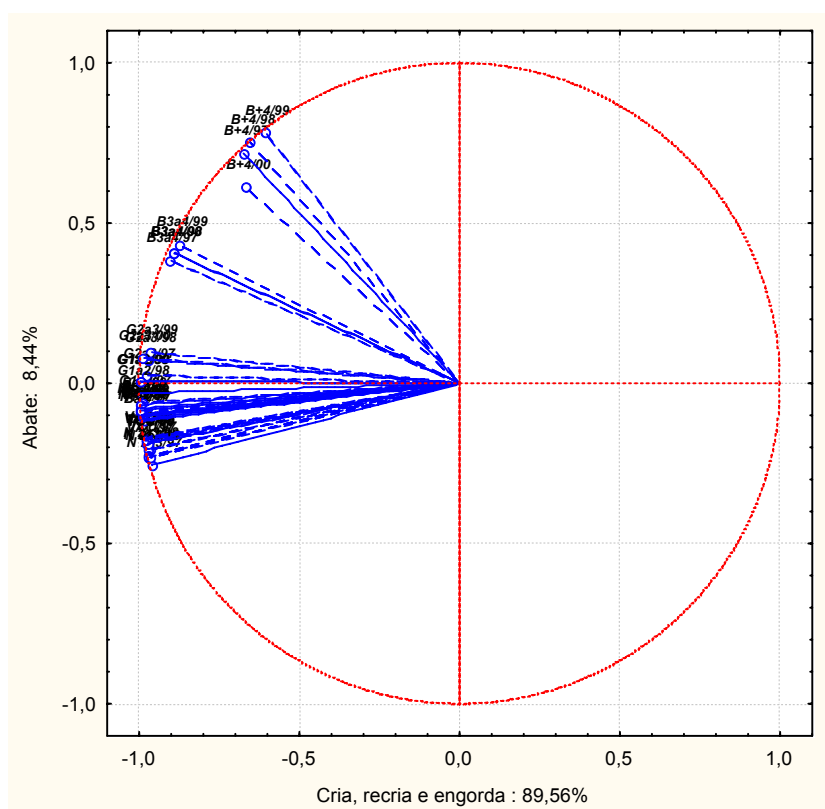


FIGURA 24 - Gráfico da distribuição da nuvem de variáveis, no círculo de correlação, no período de 1997 a 2000.

Observa-se, na Figura 24, que algumas variáveis estão sobrepostas às outras. Isso mostra que essas possuem a mesma representatividade no gráfico. Outro fato importante, é que algumas variáveis estão bem próximas ao círculo unitário, mostrando, assim, que estas possuem uma maior contribuição, em relação às variáveis que estão mais afastadas.

Verifica-se, também, que, o fator 1, que está localizado no eixo das abscissas, é representado pela cria, recria e engorda de animais, e o fator 2, que está localizado no eixo das ordenadas, é representado pelo abate de bovinos. Esse abate engloba o período de 1997 a 2000.

Uma das principais utilizações do círculo unitário é realizar a sobreposição deste sobre o primeiro plano fatorial, desta forma é possível identificar, visualmente, quais variáveis estão relacionadas com os casos em estudo.

A Figura 25 mostra os fatores a serem relacionados para a nuvem de pontos dos indivíduos, a fim de se obter o primeiro plano fatorial. Neste caso, relaciona-se fator 1 x fator 2.

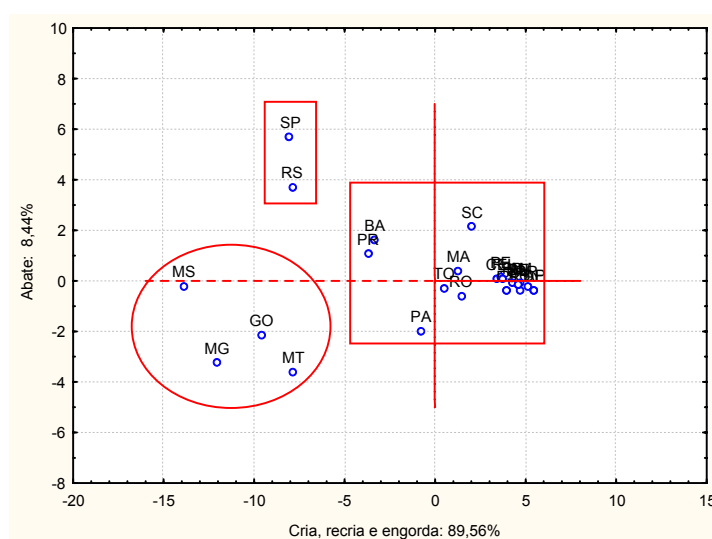


FIGURA 25 - Gráfico da distribuição da nuvem de pontos (os estados), no período de 1997 a 2000.

A Figura 25 contém a distribuição da nuvem de pontos para a totalidade das variáveis do procedimento da ACP. Analisando-se o gráfico da Figura 24, da distribuição da nuvem de variáveis em relação ao gráfico da Figura 25, da distribuição da nuvem de pontos, pode-se verificar a área espacial representada pelo círculo, localizada no segundo e terceiro quadrantes, que representa os estados brasileiros que se destacam na cria, recria e engorda de bovinos no país, que são: MG, MS, MT e GO. A área espacial representada pelo retângulo maior, no segundo quadrante, representa os estados brasileiros que se destacaram no abate: RS e SP. Os demais estados, que

estão dentro do retângulo menor, apresentam pouca representatividade em relação à produção nacional de bovinos, por se encontrarem no centro do eixo coordenado.

Após a realização da análise, do período de 1997 a 2000, procedeu-se a análise do segundo período, de 2001 a 2004.

4.4 Análise de agrupamentos de 2001 a 2004

A Figura 26 mostra o dendrograma com as categorias de bovinos no período de 2001 a 2004.

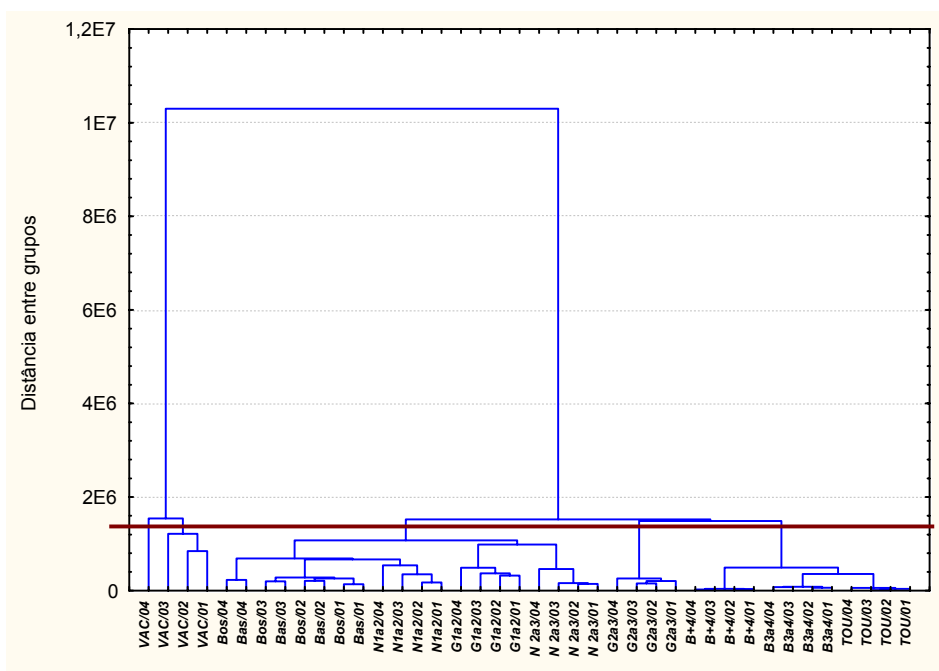


FIGURA 26– Dendrograma com as variáveis, no período de 2001 a 2004, pelo método de agrupamento por ligação simples.

Fazendo-se um corte no gráfico da Figura 26, entre as distâncias 0 e 2×10^6 , no gráfico referido como 0 e 2E6, ter-se-á cinco grupos homogêneos distintos. O primeiro grupo é formado pelas variáveis: touros, bois de 3 a 4 anos, bois + de 4 anos, representando os reprodutores e os animais para o abate. O segundo grupo, formado por garrotes de 2 a 3 anos, representando o abate jovem de animais. O terceiro grupo é formado por novilhos de 2 a 3 anos, garrotes de 1 a 2 anos, novilhos de 1 a 2 anos,

bezerras e bezerros, que representa a cria e engorda de animais. O quarto grupo é formado pela variável representada pelo número de vacas, no período de 2001 a 2003, e o quinto grupo é formado pela variável vacas, apenas no ano de 2004, que formou um grupo isolado, por ser o maior número de animais encontrados em todo o período em análise. O quarto e o quinto grupo representam as matrizes, para a criação de bovinos. Com a representação gráfica do dendograma, pode-se observar que o rebanho bovino brasileiro no período de 2001 a 2004, tornou a crescer.

A Figura 27 mostra os estados que se destacaram na produção de bovinos no período de 2001 a 2004.

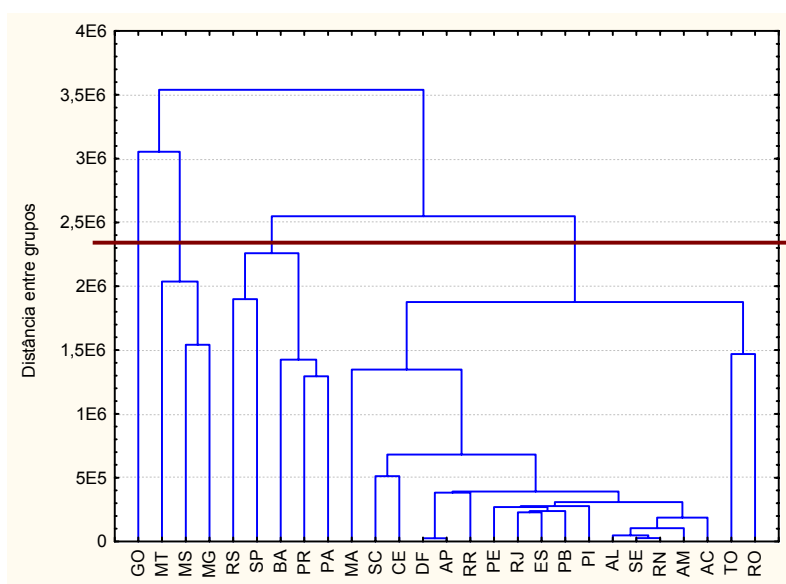


FIGURA 27 - Dendograma referente aos estados, utilizando o método de agrupamento de ligação simples, no período de 2001 a 2004.

Analisando-se a Figura 27, no período de 2001 a 2004, referente aos estados que possuem características semelhantes, chega-se a seguinte conclusão: fazendo um corte no gráfico, entre as distâncias 2×10^6 e $2,5 \times 10^6$, no gráfico referido como 2E6 e 2,5E6, ter-se-á quatro grupos homogêneos distintos. O primeiro, e maior, representado pelos estados RO, TO, AC, AM, RN, SE, AL, PI, PB, ES, RJ, PE, RR, AP, DF, CE, SC e MA. É possível observar, nesse primeiro grupo, que a distribuição dos estados não é a mesma da Figura 21, ou seja, na Figura 21, os três primeiros estados que apareciam

A Tabela 9 mostra a correlação de parte das variáveis originais. Como o número dessas é relativamente elevado, optou-se por apresentar, aqui, apenas a correlação das primeiras variáveis referente ao ano de 2001, sendo que as demais variáveis também possuem correlação alta, ou seja, o resultado é análogo à Tabela 9.

Num segundo momento, determinou-se os autovalores e o percentual de explicação de cada um desses.

Na Tabela 10, apresenta-se o resultado dos autovalores, bem como a porcentagem de variância explicada por cada componente, e, também, a variância acumulada pelas mesmas.

TABELA 10 – Autovalores e percentual da variância explicada de cada componente, no período de 2001 a 2004

Número de componentes	Autovalores Extração dos componentes principais			
	Autovalores	% da variância explicada	Autovalores acumulados	% da variância explicada acumulada
1	33,86	84,64	33,86	84,64
2	5,19	12,97	39,05	97,62
3	0,41	1,01	39,45	98,63
4	0,21	0,53	39,67	99,16
5	0,15	0,38	39,82	99,55
6	0,06	0,16	39,88	99,70
7	0,05	0,12	39,93	99,82
8	0,02	0,06	39,95	99,88
9	0,01	0,04	39,97	99,91
10	0,01	0,03	39,98	99,94
11	0,01	0,03	39,99	99,97
12	0,00	0,01	39,99	99,98
13	0,00	0,01	40,00	99,99
14	0,00	0,01	40,00	99,99

Observando-se a Tabela 10, verifica-se que, novamente, são apenas dois os autovalores que possuem a maior explicação, representando cerca de 97,62% da variância. Portanto, os dados serão resumidos pelos dois primeiros autovalores.

Analisando-se, ainda, a Tabela 10, pode-se verificar que o número de fatores a serem retirados, para a análise, é dois, que correspondem a 84,64% e 12,97% da variância total, explicada pelos autovalores do modelo. Ou seja, explicam, juntos, 97,62% das variações das medidas originais. A perda de informação, nesse caso, é de apenas 2,38%. Esses fatores podem ser visualizados na Figura 28.

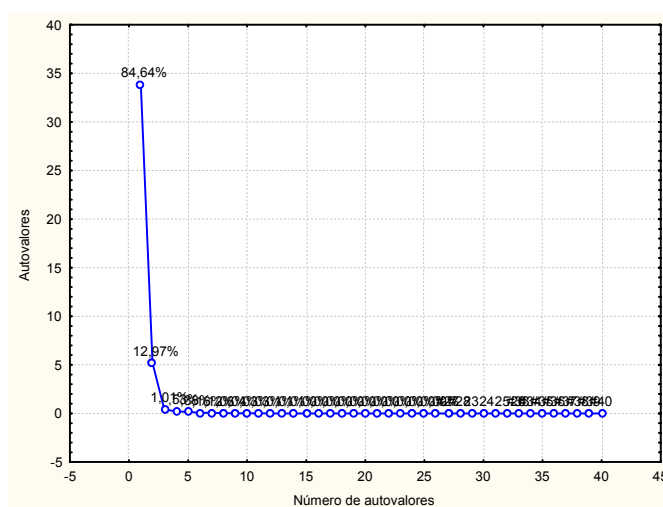


FIGURA 28 - Gráfico de explicação da proporção de variação de cada componente principal, no período de 2001 a 2004.

Pode-se verificar, na Figura 28, dois são os fatores a serem retirados para análise, pois são estes que possuem um decaimento no gráfico. Pode-se observar, também, que os outros fatores apresentam uma baixa explicação, não sendo aconselhável incluí-los na análise.

Após ter definido quantos fatores serão utilizados na análise, deve-se verificar a importância de cada variável na composição dos fatores, ou seja, encontrar a matriz de cargas fatoriais, a qual mostra que quanto maior for a carga em cima de um fator, mais a variável se identifica com este. A Tabela 11 mostra a matriz de cargas fatoriais após a rotação nos eixos, sendo que esta possibilita uma melhor compreensão do resultado. Nesta pode-se verificar que as variáveis que possuem as maiores cargas fatoriais são as que melhor explicam o fator 1 e o fator 2, sendo que aquelas que melhor representam o fator 1, no período de 1997 a 2000, são as mesmas no período de 2001 a 2004. O mesmo ocorre em relação ao fator 2, conforme Tabela 11.

TABELA 11 – Cargas fatoriais para composição dos fatores no período de 2001 a 2004

	Fator 1	Fator 2
TOU/01	0,96	0,21
VAC/01	0,97	0,21
N 2a3/01	0,97	0,22
N1a2/01	0,97	0,19
Bas/01	0,97	0,20
Bos/01	0,97	0,20
G1a2/01	0,94	0,29
G2a3/01	0,81	0,51
B3a4/01	0,49	0,79
B+4/01	-0,06	0,99
TOU/02	0,97	0,20
VAC/02	0,97	0,20
N 2a3/02	0,97	0,20
N1a2/02	0,97	0,20
Bas/02	0,98	0,19
Bos/02	0,98	0,19
G1a2/02	0,95	0,29
G2a3/02	0,84	0,50
B3a4/02	0,49	0,82
B+4/02	0,11	0,99
TOU/03	0,97	0,20
VAC/03	0,97	0,20
N 2a3/03	0,97	0,21
N1a2/03	0,98	0,19
Bas/03	0,98	0,18
Bos/03	0,98	0,17
G1a2/03	0,95	0,27
G2a3/03	0,84	0,48
B3a4/03	0,57	0,76
B+4/03	0,09	0,98
TOU/04	0,96	0,20
VAC/04	0,96	0,22
N 2a3/04	0,96	0,19
N1a2/04	0,97	0,20
Bas/04	0,97	0,20
Bos/04	0,97	0,20
G1a2/04	0,93	0,29
G2a3/04	0,82	0,50
B3a4/04	0,51	0,80
B+4/04	0,12	0,96

Logo, é possível verificar na Tabela 11, que o fator 1 é o mais importante para o estudo, pois é derivado do maior autovalor, sendo este representado pelas seguintes variáveis: TOU/01, VAC/01, N 2a3/01, N 1a2/01, Bas/01, Bos/01, G 1a2/01, G 2a3/01, TOU/02, VAC/02, N 2a3/02, N 1a2/02, Bas/02, Bos/02, G 1a2/02, G 2a3/02, TOU/03, VAC/03, N 2a3/03, N 1a2/03, Bas/03, Bos/03, G 1a2/03, G 2a3/03, TOU/04, VAC/04, N 2a3/04, N 1a2/04, Bas/04, Bos/04, G 1a2/04 e G 2a3/04. Logo, pode-se denominar esse fator de cria, recria e engorda de animais. O fator 2 é derivado do segundo autovalor, sendo é representado pelas seguintes variáveis: B 3 a 4/01, B + 4/01, B 3 a 4/02, B + 4/02, B 3 a 4/03, B + 4/03, B 3 a 4/04 e B + 4/04, denominado de abate.

Para melhor visualizar esses fatores, fez-se os gráficos de dispersão, conforme Figura 29.

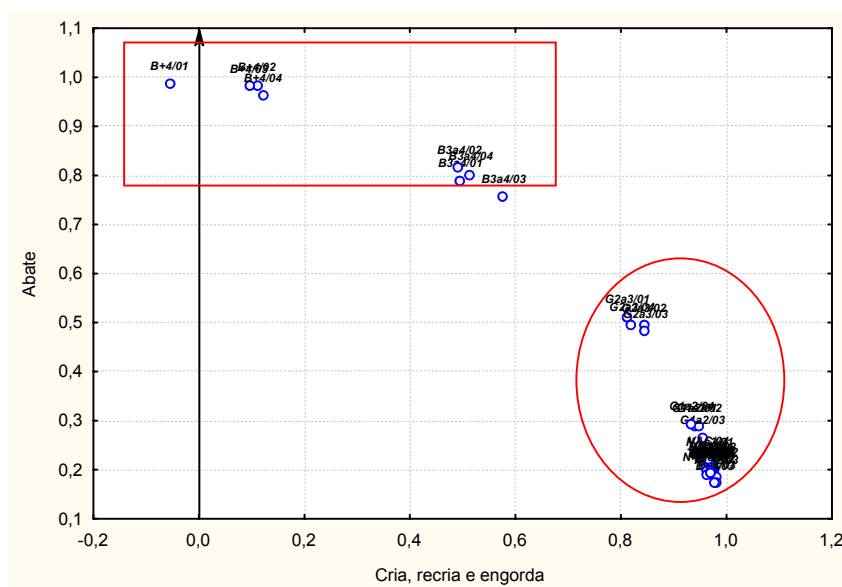


FIGURA 29 - Gráfico representando a relação entre fatores fator 1 x fator 2 no período de 2001 a 2004.

Observando-se o plano fatorial da Figura 29, pode-se verificar que as variáveis que melhor representam a cria, recria e engorda, são as que melhor explicam esse fator, ou seja, as que estão mais distantes da origem em relação ao eixo das abscissas, representadas pelo círculo. As variáveis que melhor representam o abate, são: B 3 a 4 e B + 4, no período de 2001 a 2004. Essas variáveis estão contidas no retângulo, ou seja, são as que estão mais distantes da origem, em relação ao eixo das ordenadas, sendo essas as que melhor explicam o fator 2.

Após a realização da análise fatorial, procedeu-se com a análise de componentes principais.

A Figura 30 mostra o círculo de correlação unitário, com a nuvem de variáveis, referente ao primeiro plano principal, no período de 2001 a 2004.

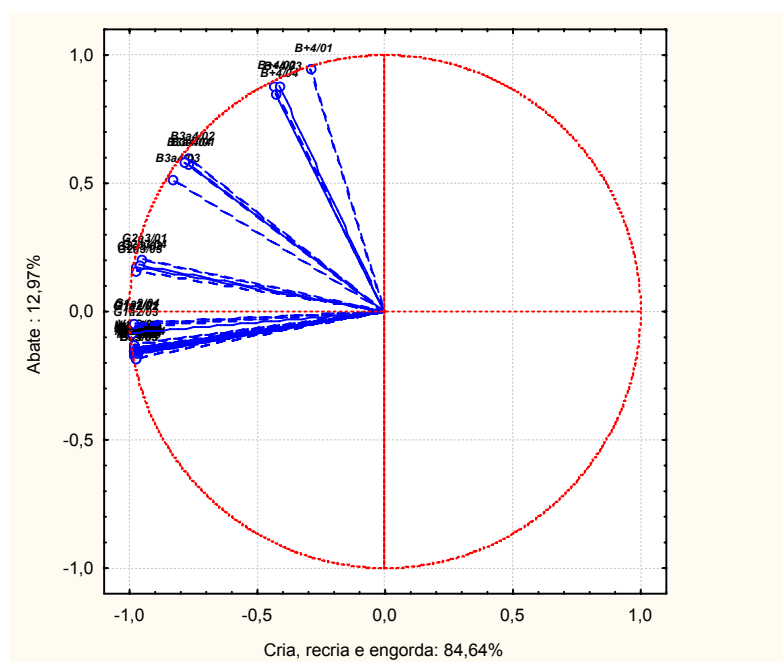


FIGURA 30 - Gráfico da distribuição da nuvem de variáveis, no círculo de correlações, no período de 2001 a 2004.

Na Figura 30, pode-se observar que as variáveis, que melhor explicam o fator 1, estão relacionadas com a cria, recria e engorda de animais, pois essas formam o menor ângulo em relação ao eixo das abscissas. Isso significa que estão mais correlacionadas entre si. Em relação ao abate, as variáveis que melhor explicam são os B + de 4 anos e, em segundo, os B 3 a 4 anos de idade, pois as variáveis que estão em primeiro são as que formam o menor ângulo em relação ao eixo das ordenadas.

A Figura 31 contém a distribuição da nuvem de pontos para a totalidade das variáveis do procedimento da ACP. A área espacial representada pelo círculo, localizada no terceiro quadrante, representa os estados brasileiros que se destacaram na cria, recria e engorda de bovinos, no país, no período de 2001 a 2004. A área espacial representada pela elipse no segundo quadrante, representa os estados brasileiros que se destacaram no abate, e a área representada pelo retângulo apresenta pouca representatividade, pois se encontra no centro do eixo coordenado.

A Figura 31 mostra os fatores a serem relacionados para a nuvem de pontos dos indivíduos. Nesse caso, relaciona-se *factor 1 x factor 2*.

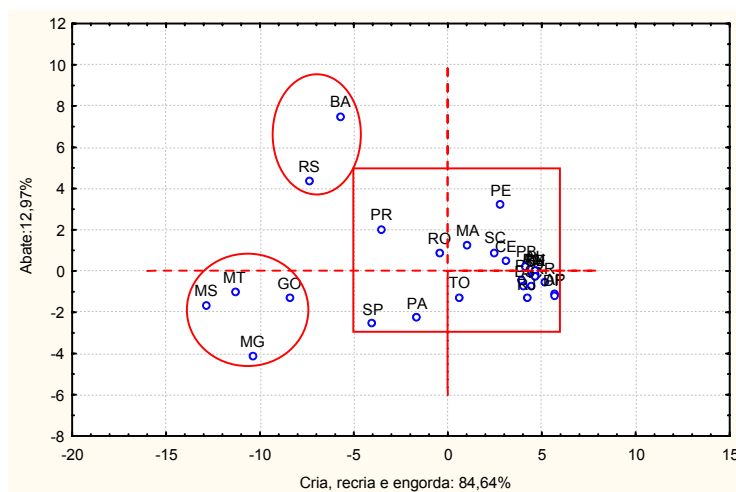


FIGURA 31 - Gráfico da distribuição da nuvem de pontos (os estados), no período de 2001 a 2004.

Realizando-se a sobreposição dos gráficos das Figuras 30 e 31, pode-se concluir que as variáveis que melhor representam a cria, recria e engorda de bovinos são representadas pelos estados MG, MS, MT e GO. As variáveis que melhor representam o abate de bovinos são representadas pelos estados RS e BA. Os demais estados, que estão dentro do retângulo, apresentam pouca representatividade em relação à produção nacional de bovinos, por de encontrarem no centro do eixo coordenado.

Pode-se verificar que, com a análise de 1997 a 2000, e 2001 a 2004, os estados que possuíram destaque na bovinocultura foram MG, MS, MT e GO, em relação à cria, recria e engorda. Os estados do RS, SP e BA possuíram destaque em relação ao abate, sendo que SP esteve como destaque no período de 1997 a 2000, e a BA passou a se destacar no período de 2001 a 2004. Mediante a análise multivariada, e através da análise descritiva, pode-se concluir que o estado de SP teve decréscimos na produção de bovinos desde 1997. Isso mostra que esse estado está deixando de investir nessa atividade econômica. O estado da Bahia teve um crescimento considerável em todo o período analisado, isso é resultado do investimento que esse estado está fazendo na bovinocultura. Conforme Revista Globo Rural nº 231, “o estado da Bahia produz por ano 380 mil toneladas de carne bovina, para o consumo de 315 mil, mostrando-se auto-suficiente e, ainda, abastecendo outros mercados”. Isso tornou-se possível devido aos incentivos fiscais, investidos nesta atividade econômica.

O restante dos estados brasileiros, após as análises realizadas, pode ser classificado com desempenho inferior por possuírem pouca ou nenhuma representatividade em relação à produção nacional de bovinos. Embora alguns desses, pelas análises realizadas, mostra estar em crescimento. São eles: RO, AC, AM, RR, AP, TO, MA, PI, CE, RN PB, PE, AL, SE, ES, RJ, PR, SC, DF, PA.

Como todo o processo produtivo deve ser analisado de forma a se trazer um benefício para o setor, até o presente momento foi possível identificar os estados produtores e o tipo de produção que cada um apresenta. Também foi possível verificar as oscilações ocorridas no período de análise. Ainda, utilizou-se nesta pesquisa a metodologia de séries temporais, a qual possibilitará monitorar a cotação em arroba e em dólar, do preço do boi gordo no mercado nacional. Embora essa variável não esteja diretamente ligada ao número de cabeças de gado produzidas, sabe-se que a elevação do valor da cotação influencia diretamente no aumento do número de cabeças. Isto é, essa variável funciona como um termômetro que estimula a produção.

4.6 Séries temporais

A análise de séries temporais é realizada apenas naqueles estados que foram identificados pela análise multivariada, considerando-os como os vocacionados à produção de bovinos. Dessa forma, tem-se uma redução de estados a serem monitorados.

4.6.1 Rio Grande do Sul

A análise descritiva é sempre reveladora, pois é capaz de mostrar o comportamento de uma situação, ou problema. Neste caso utiliza-se medidas descritivas como a média, o desvio-padrão e o coeficiente de variação das variáveis em estudo. Para tal, essas medidas foram realizadas ano a ano, investigando, dessa maneira, se houve alguma modificação do período analisado. Na Tabela 12, apresenta-se as estatísticas para o preço do boi gordo em arroba, que são auto-explicativas.

TABELA 12 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado do RS, no período de janeiro de 1997 a dezembro de 2004

Estatísticas	1997	1998	1999	2000	2001	2002	2003	2004
Média	22,53	24,38	17,18	20,37	17,98	15,08	16,06	17,09
Desvio padrão	2,09	1,49	1,09	1,28	1,66	2,07	2,07	0,99
Coeficiente de								
Varição	9,27	6,10	6,37	6,31	9,22	13,74	12,88	5,79

Pode-se observar, na Tabela 12, o coeficiente de variação de Pearson, para o período de 1997 a 2004, o qual é muito significativo, pois o mesmo encontra-se bem abaixo de 50 %, o que revela que a média do preço do boi gordo em arroba é muito significativa.

Para identificar o modelo adequado, inicialmente será analisado o gráfico da série original, representado na Figura 32, o qual apresenta uma tendência decrescente e uma mudança de nível. Pode-se dizer, através de uma análise visual, que o preço do boi gordo em arroba teve uma tendência decrescente, não podendo, assim, a série ser considerada estacionária, o que é corroborado no dendograma da Figura 19. Para uma melhor decisão sobre a estacionaridade da série, dispõe-se da FAC, e da FACP, e que servirão para encontrar o modelo desejado.

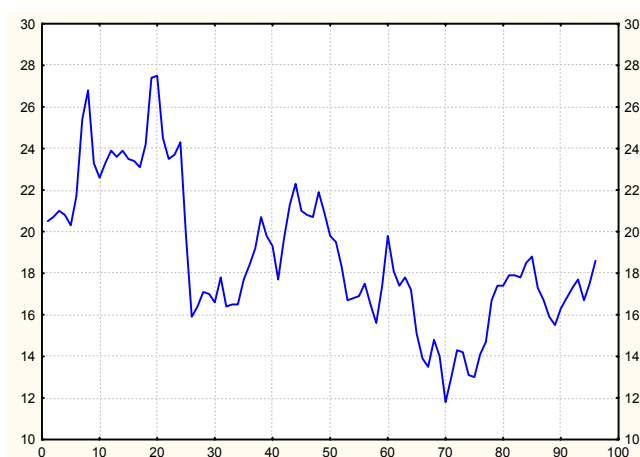


FIGURA 32 – Série original referente ao preço do boi gordo, em arroba, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.

Séries não estacionárias declinam muito lentamente na medida em que os *lag* (retardos ou atrasos) crescem, isto porque séries não estacionárias apresentam fortes

correlações seriais. Observando-se a FAC, na Figura 33, os valores são, inicialmente, altos e não declinam rapidamente para zero, à medida em que os lag crescem, confirmando a não estacionaridade.

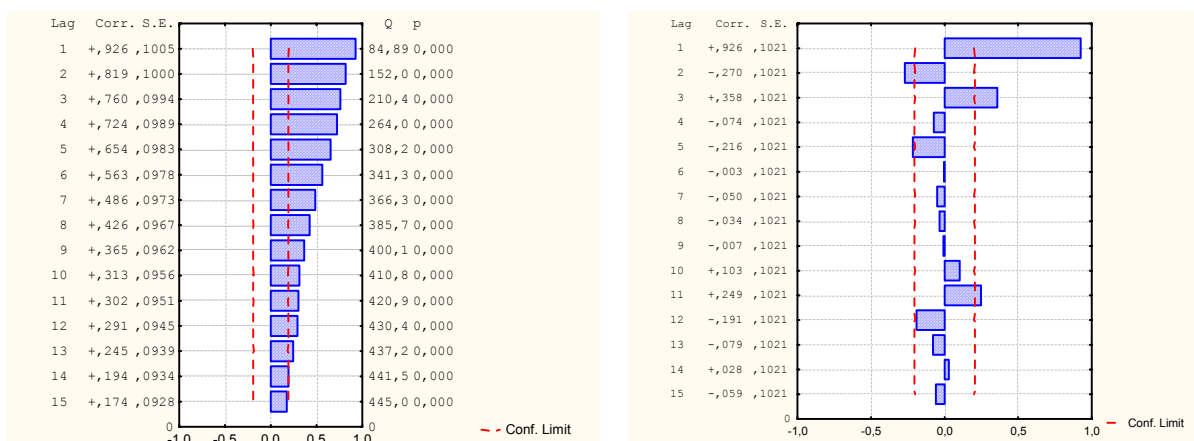


FIGURA 33 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do RS, no período de janeiro de 1997 a dezembro de 2004.

Como a metodologia Box & Jenkins aplica-se a séries estacionárias, ou que se tornam estacionárias após a aplicação de diferenças, o próximo passo é estacionarizar a série através de diferenças. Ainda observando a Figura 33, constata-se sazonalidade, mas é válido lembrar que, ao ser aplicada diferenças na série original, pode-se excluir a sazonalidade da série.

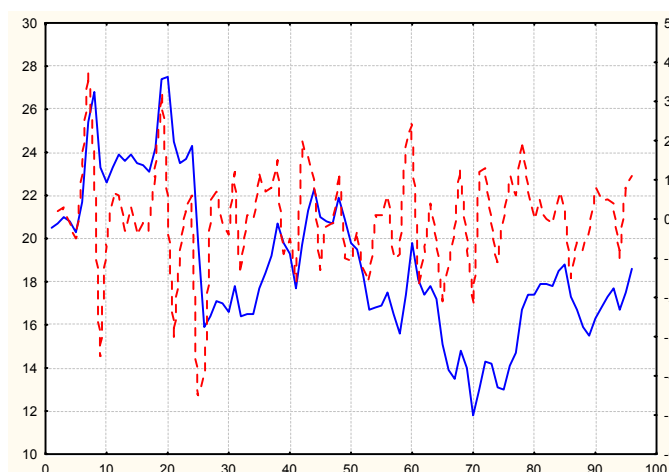


FIGURA 34 – Sazonalidade do preço do boi gordo, em arroba, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.

Ao aplicar-se uma diferença, na série original, é possível observar que ela apresentou um comportamento estacionário, chegando ao objetivo desejado, podendo ser possível utilizar a metodologia para encontrar modelos adequados que representem a série e, assim, ser possível fazer previsões futuras.

O próximo passo é descobrir qual é o melhor modelo que represente a série, os quais estão representados no Quadro 1.

Modelo	Parâmetros	t(calculado)	Nível de Significância	AIC	BIC
SARIMA(2,1,0)(1, 0, 1)₁₂	$\phi(1) = 0,31$ $\phi(2) = -0,32$ $\Phi = 0,88$ $\Theta = 0,72$	3,11 -3,12 4,35 2,91	0,002504 0,002412 0,000035 0,004519	466,21	476,42
SARIMA(0,1,1)(1,0,1) ₁₂	$\theta = -0,37$ $\Phi = 0,88$ $\Theta = 0,68$	-3,80 5,62 2,36	0,000257 0,000000 0,001146	469,64	477,30
SARIMA(1,1,0)(1, 0, 1) ₁₂	$\phi = 0,24$ $\Phi = 0,89$ $\Theta = 0,68$	2,36 6,19 3,51	0,020579 0,000000 0,000701	474,06	481,73
ARIMA(3,2,0) ₁₂	$\phi(1) = -0,43$ $\phi(2) = -0,60$ $\phi(3) = -0,43$	-4,48 -7,02 -4,51	0,000021 0,000000 0,000019	478,89	486,52

QUADRO 1 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.

Pode-se perceber que o melhor modelo encontrado foi um SARIMA. Isso diz que a diferença aplicada não foi suficiente para retirar a sazonalidade de ordem 12 da cotação do preço do boi gordo em arroba, para o RS. Como foram encontrados mais de um modelo significativo, foi necessário a utilização dos critérios AIC e BIC, que, nesse caso, foram os menores encontrados, não sendo necessário utilizar critérios adicionais para escolhê-los. Pode-se dizer que este modelo foi o mais parcimonioso.

Na Figura 35, apresenta-se a FAC e a FACP residual. Os resíduos do modelo estimado são estimativas do ruído branco. Sendo assim, devem estar representados desta forma na FAC e FACP residual.

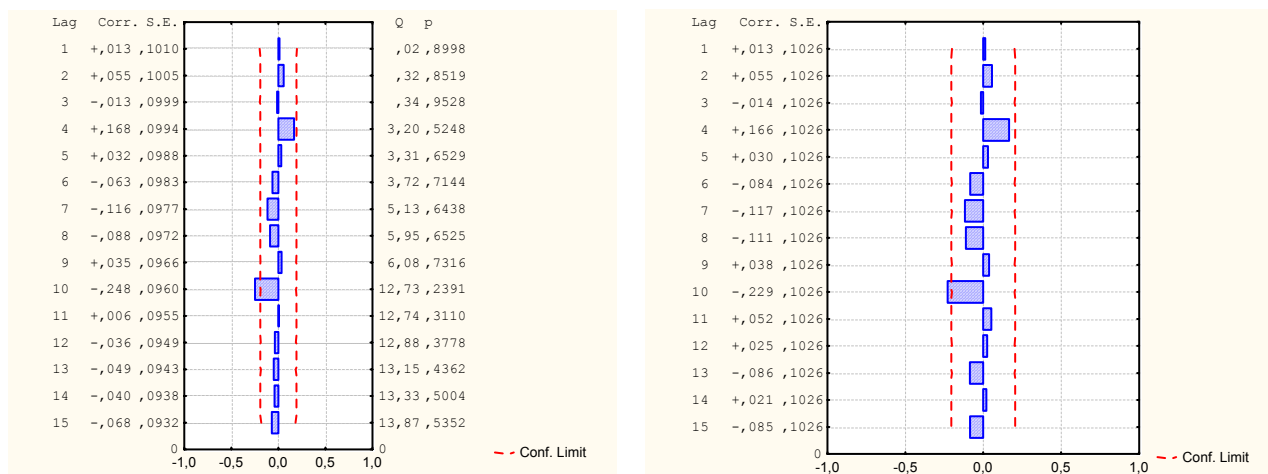


FIGURA 35 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo SARIMA (2,1,0)(1,0,1), no estado do RS, no período de janeiro de 1997 a dezembro de 2004.

A Figura 35 releva que os resíduos apresentam-se quase todos dentro dos limites de confiabilidade, ou seja, são significativos, confirmando a adequação do modelo. Também observa-se que, no instante 26, há uma queda brusca na série original, sugerindo, então, uma intervenção neste instante, a qual será realizada em um próximo estudo, a fim de melhor explicá-la. Como esse modelo foi o que melhor explicou esta série, fez-se as previsões sem levar em conta a intervenção.

Para confirmar a adequação do modelo, dispõe-se ainda, do gráfico da normal dos resíduos, Figura 36, onde mostra que os pontos estão próximos à linha reta, garantindo, assim, a normalidade dos resíduos.

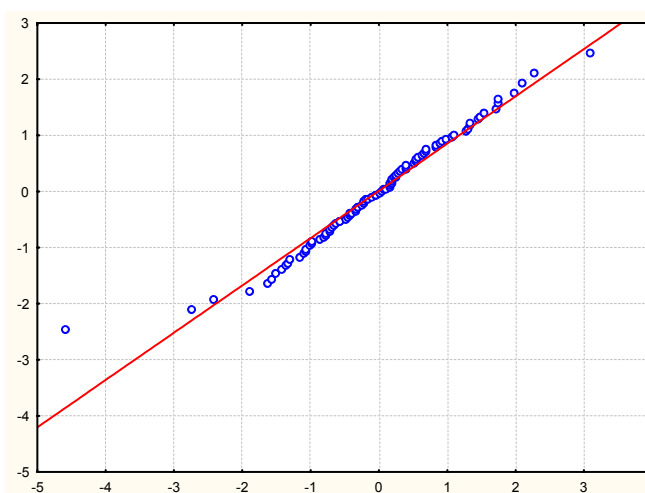


FIGURA 36 - Gráfico da Distribuição Normal, no estado do RS, no período de janeiro de 1997 a dezembro de 2004.

Na Tabela 13, apresentam-se as previsões para os meses de janeiro a dezembro de 2005, para o preço do boi gordo em arroba, para o estado do Rio Grande do Sul, obtidas com base no mês de dezembro de 2004.

TABELA 13 - Previsões para o preço do boi gordo, em arroba, no estado do RS, nos meses de janeiro a dezembro de 2005

Meses	Previsões	Inferior 95%	Superior 95%	Erro Padrão
Jan/05	18,34	16,40	20,28	1,17
Fev/05	17,65	14,45	20,86	1,93
Mar/05	17,48	13,65	21,32	2,31
Abr/05	17,39	13,17	21,62	2,54
Mai/05	17,08	12,48	21,68	2,77
Jun/05	17,45	12,47	22,44	3,00
Jul/05	17,80	12,44	23,16	3,23
Ago/05	18,04	12,35	23,74	3,43
Set/05	17,87	11,86	23,88	3,61
Out/05	17,44	11,14	23,75	3,79
Nov/05	17,81	11,22	24,40	3,97
Dez/05	18,39	11,53	25,26	4,13

Na Figura 37, observa-se o gráfico das previsões do preço do boi gordo em arroba, para os meses de janeiro de 2005 à dezembro de 2005, como representado pela Tabela 13.

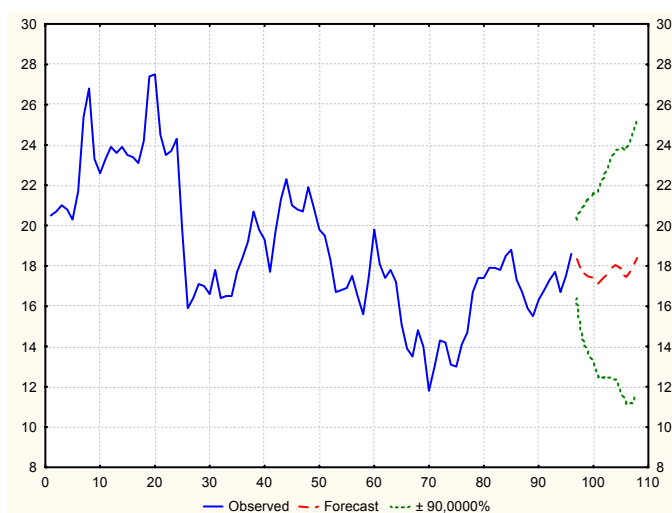


FIGURA 37 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do RS, de janeiro a dezembro de 2005.

Através da metodologia Box & Jenkins, foi possível modelar uma série temporal que melhor explicou o preço do boi gordo em arroba, no estado do RS, a qual possibilitou a realização da previsão para o ano de 2005. O modelo utilizado foi

o melhor, pois foi o que melhor esteve representado nos limites de confiabilidade considerados na FAC e FACP residual.

4.6.2 Minas Gerais

Para verificar se o preço do boi gordo em arroba foi significativo no estado de Minas Gerais, no decorrer do período, utilizou-se medidas descritivas, tais como a média, o desvio-padrão e o coeficiente de variação das variáveis em estudo. Para tal, essas medidas foram realizadas ano a ano, investigando, dessa maneira, se houve alguma modificação do período analisado. Na Tabela 14, apresenta-se as estatísticas para o preço do boi gordo em arroba, que são auto-explicativas.

TABELA 14 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado de MG, no período de janeiro de 1997 a dezembro de 2004

Estatísticas	1997	1998	1999	2000	2001	2002	2003	2004
Média	22,92	22,78	17,65	18,51	16,03	14,55	16,03	17,22
Desvio padrão	0,78	0,62	1,90	1,13	1,15	1,79	2,05	1,53
Coeficiente de Variação	3,39	2,71	10,75	6,12	7,18	12,30	12,78	8,89

Como o coeficiente de variação de Pearson esteve bem inferior a 50% em todo o período, isso mostra que esse é muito significativo, revelando, assim, que a média do preço do boi gordo em arroba é muito significativa.

Para identificar o modelo adequado, inicialmente será analisado o gráfico da série original, representado na Figura 38, o qual apresenta uma tendência decrescente no início e no final certo crescimento. Pode-se dizer, através de uma análise visual, que o preço do boi gordo em arroba teve uma tendência decrescente, não podendo, assim, ser considerada estacionária. Como não é confiável uma análise visual, dispõe-se da FAC para ser verificada a estacionaridade, e da FACP como auxílio para encontrar o modelo desejado.

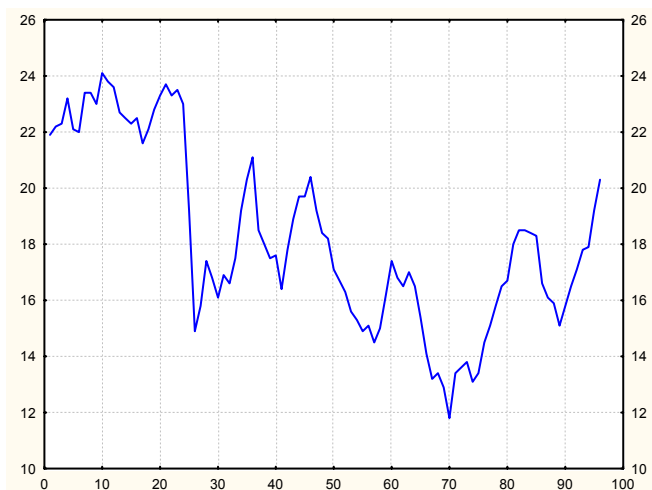


FIGURA 38 - Série original referente ao preço do boi gordo, em arroba, no estado do MG, no período de janeiro de 1997 a dezembro de 2004.

Observa-se, na Figura 39, que a FAC não possui um decaimento exponencial, confirmando a não estacionaridade, mostrando que existem fortes correlações. Logo, será preciso aplicar diferenças para estacionarizá-la e, assim, ser possível aplicar a metodologia Box & Jenkins.

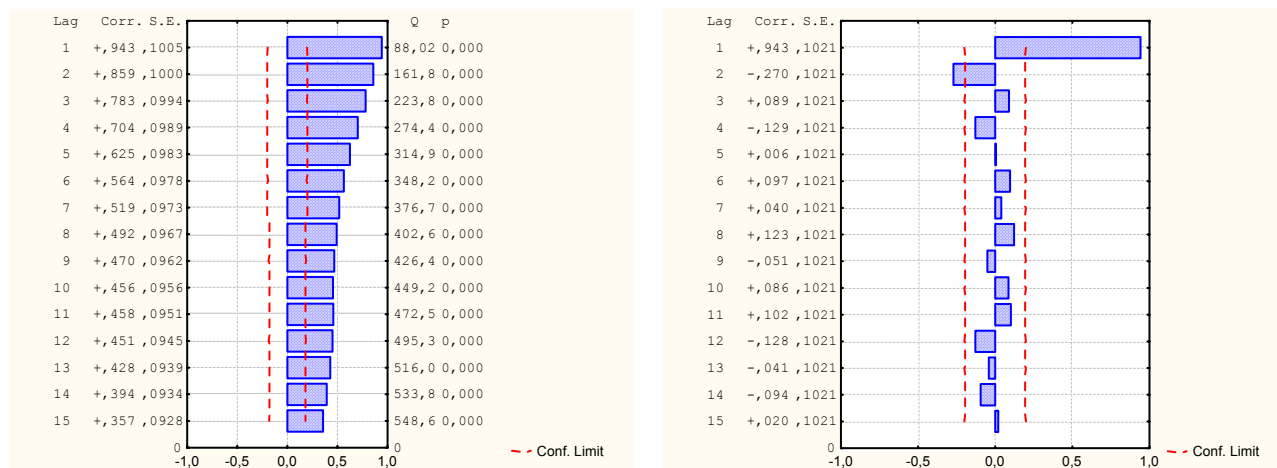


FIGURA 39 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do MG, no período de janeiro de 1997 a dezembro de 2004.

Observando-se a FAC, na Figura 39, os valores são inicialmente altos e não declinam rapidamente para zero, à medida que os *lag* crescem, confirmando a não estacionaridade. O próximo passo é estacionarizar a série através de diferenças. Ainda observando a Figura 39, constata-se sazonalidade, mas é válido lembrar que, ao ser aplicada diferenças na série original, pode-se excluir a sazonalidade da série.

A Figura 40 apresenta a série original e a série com duas diferenças, na qual é possível observar que apresentou um comportamento estacionário, mediante o qual será possível encontrar modelos adequados que representem a série e, assim, fazer previsões futuras.

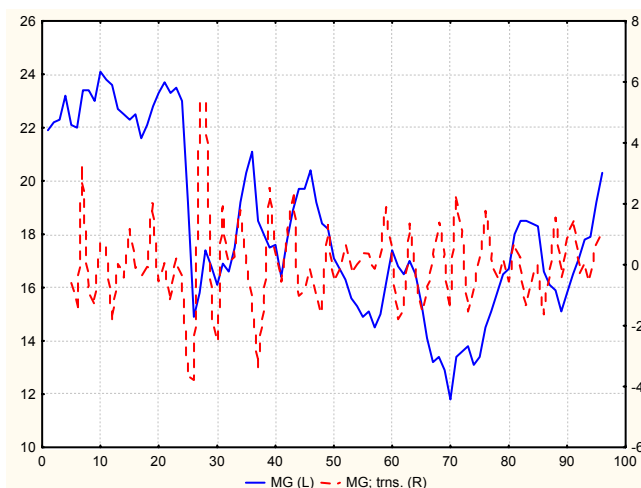


FIGURA 40 – Sazonalidade do preço do boi gordo, em arroba, no estado de MG, no período de janeiro de 1997 a dezembro de 2004.

O próximo passo é descobrir qual o melhor modelo que represente a série. Após várias tentativas, os modelos significativos encontrados estão representados no Quadro 2.

Modelo	Parâmetros	t(calculado)	Nível de Significância	AIC	BIC
SARIMA (1,1,0)(1,0,1)₁₂	$\phi = 0,28$ $\Phi = 0,90$ $\Theta = 0,74$	2,78 3,94 2,58	0,006588 0,000158 0,011577	424,32	431,98
ARIMA (1,2,1) ₁₂	$\phi = 0,32$ $\theta = 0,94$	3,06 49,00	0,002910 0,000000	427,36	432,44
ARIMA (1,1,2) ₁₂	$\phi = 0,84$ $\Theta(1) = 0,54$ $\Theta(2) = 0,36$	5,51 3,45 4,03	0,000000 0,000850 0,000115	426,58	434,24

QUADRO 2 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado de MG, no período de janeiro de 1997 a dezembro de 2004.

Pode-se observar que foi encontrado um SARIMA com uma diferença simples e um diferença sazonal de ordem 12. Isso nos diz que a diferença aplicada não foi suficiente para retirar a sazonalidade da série em estudo. Como foram

encontrados mais de um modelo significativo, foi necessário a utilização dos critérios AIC e BIC.

Na Figura 41, apresenta-se a FAC e a FACP residual. Os resíduos do modelo estimado são estimativas do ruído branco. Sendo assim, devem estar representados desta forma na FAC e FACP residual.

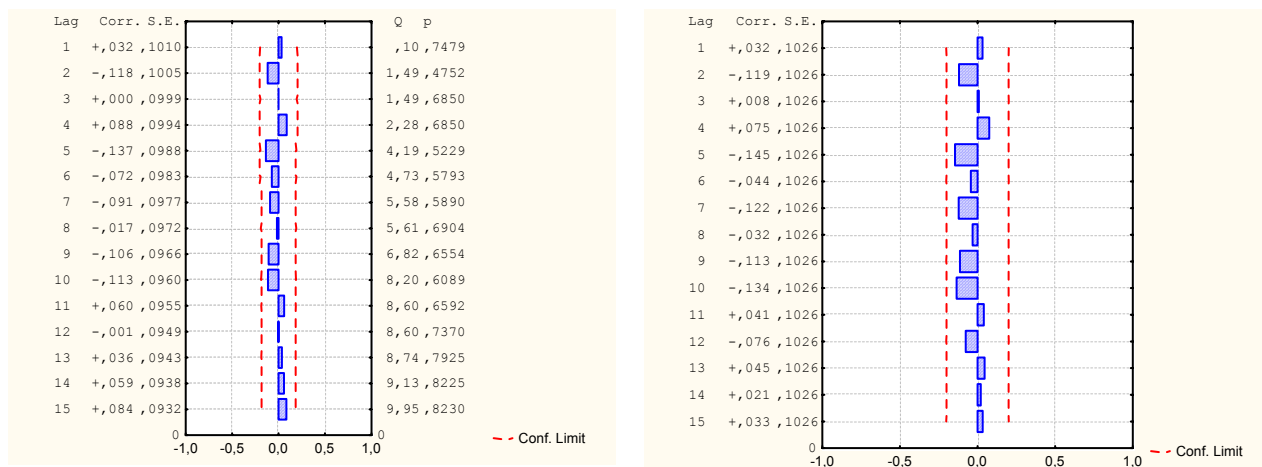


FIGURA 41 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo SARIMA (1,1,0)(1,0,1)₁₂, no estado do MG, no período de janeiro de 1997 a dezembro de 2004.

A Figura 41 relewa que os resíduos apresentam-se dentro dos limites de confiabilidade, ou seja, são significativos, confirmando a adequação do modelo.

Para confirmar a adequação do modelo, dispõe-se do gráfico da normal dos resíduos, Figura 42, onde mostra que os pontos estão próximos à linha reta, garantindo, assim, a normalidade dos mesmos.

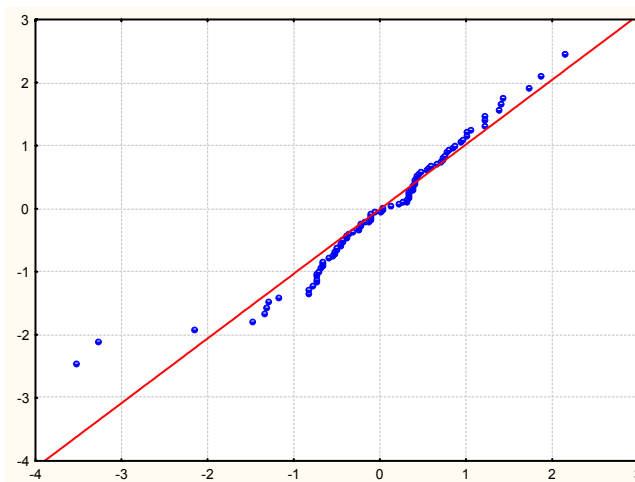


FIGURA 42 - Gráfico da Distribuição Normal, do estado de MG, no período de janeiro de 1997 a dezembro de 2004.

Na Tabela 15, apresentam-se as previsões para os meses de janeiro a dezembro de 2005, para o preço do boi gordo, em arroba, para o estado do Minas Gerais, obtidas com base no mês de dezembro de 2004.

TABELA 15 - Previsões para preço do boi gordo, em arroba, no estado de MG, nos meses de janeiro a dezembro de 2005

Meses	Previsões	Inferior 95%	Superior 95%	Erro Padrão
Jan/05	20,72	19,07	22,37	0,99
Fev/05	20,93	18,18	23,67	1,65
Mar/05	21,06	17,42	24,70	2,19
Abr/05	21,18	16,77	25,58	2,65
Mai/05	21,29	16,20	26,37	3,06
Jun/05	21,39	15,70	27,09	3,42
Jul/05	21,50	15,23	27,77	3,77
Ago/05	21,61	14,80	28,41	4,10
Set/05	21,71	14,39	29,03	4,41
Out/05	21,82	14,00	29,63	4,70
Nov/05	21,92	13,63	30,21	4,99
Dez/05	22,03	13,28	30,78	5,27

Na Figura 43, observa-se o gráfico das previsões do preço do boi gordo em arroba, para os meses de janeiro de 2005 a dezembro de 2005, como representado pela Tabela 15.

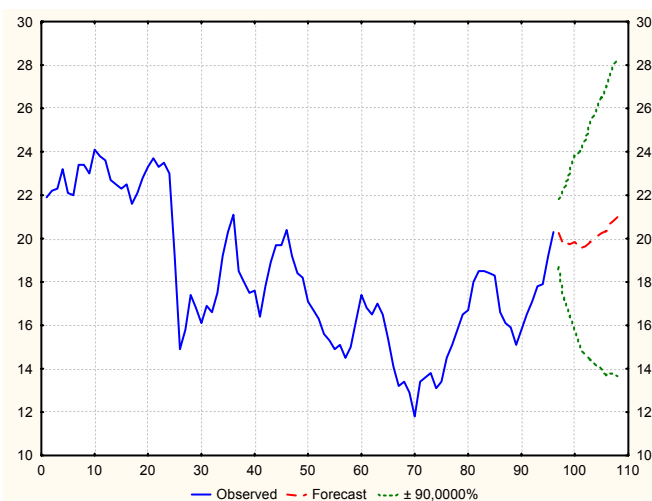


FIGURA 43 - Gráfico de previsões para o estado de MG, de janeiro a dezembro de 2005.

Concluiu-se, para o estado de MG, que o melhor modelo que representou este estado foi um SARIMA (1,1,0)(1,0,1)₁₂, o qual manteve-se dentro dos limites de

confiabilidade, representados pela FAC e FACP, seguindo uma distribuição normal. Esse foi o melhor modelo encontrado para estimar a previsão para o ano de 2005.

4.6.3 Mato Grosso do Sul

Para verificar se o preço do boi gordo em arroba, foi significativo no estado de Mato Grosso do Sul, no decorrer do período, utilizou-se medidas descritivas, tais como a média, o desvio-padrão e o coeficiente de variação das variáveis em estudo. Para tal, essas medidas foram realizadas ano a ano, investigando-se, dessa maneira, se houve alguma modificação do período analisado.

Na Tabela 16, apresenta-se as estatísticas para o preço do boi gordo, em arroba, que são auto-explicativas.

TABELA 16 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado do MS, no período de janeiro de 1997 a dezembro de 2004

Estatísticas	1997	1998	1999	2000	2001	2002	2003	2004
Média	22,88	22,71	17,93	17,98	15,73	14,33	16,33	17,45
Desvio padrão	0,84	0,66	1,68	1,12	1,01	1,62	1,96	1,22
Coeficiente de Variação	3,66	2,89	9,39	6,23	6,42	11,31	12,00	6,98

Como pode-se observar, na Tabela 16, o coeficiente de variação de Pearson, esteve bem inferior a 50% em todo o período. Isso mostra que esse é muito significativo, revelando, assim, que a média do preço do boi gordo, em arroba, é muito significativa para o estado do MS.

Para identificar o modelo adequado, inicialmente será analisado o gráfico da série original, representado na Figura 44, o qual apresenta uma tendência decrescente no início e, no final, certo crescimento. Pode-se dizer, através de uma análise visual, que o preço do boi gordo, em arroba, teve uma tendência decrescente, não podendo, assim, ser considerada estacionária. Como não é confiável uma análise visual, dispõe-se da FAC para ser verificada a estacionaridade, e da FACP como auxílio para encontrar o modelo desejado.

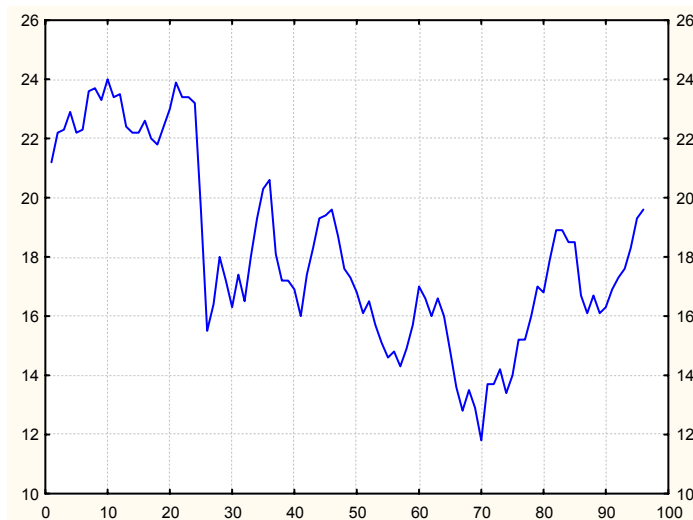


FIGURA 44 - Série original referente ao preço do boi gordo, em arroba, no estado do MS, no período de janeiro de 1997 a dezembro de 2004.

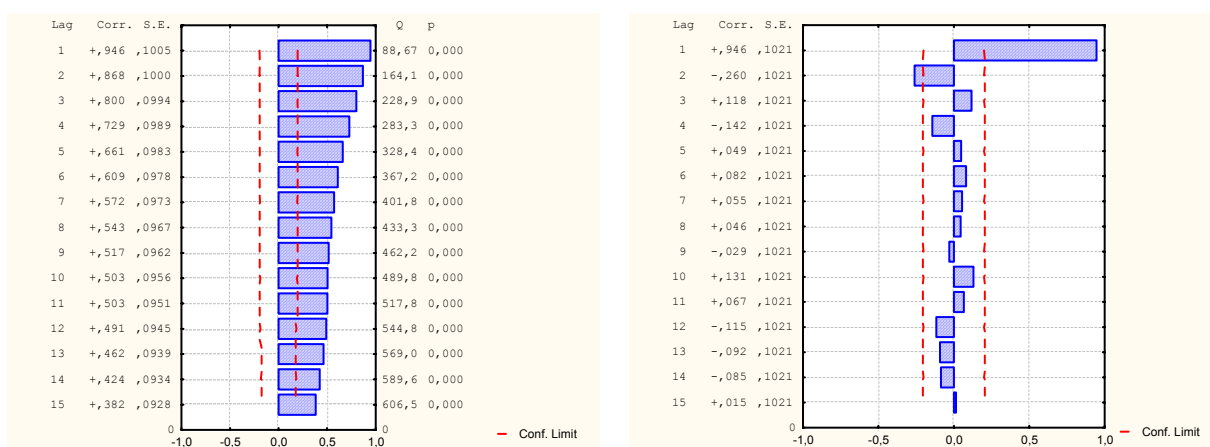


FIGURA 45 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do MS, no período de janeiro de 1997 a dezembro de 2004.

Observa-se, na Figura 45, que a FAC não possui um decaimento exponencial, confirmando a não estacionaridade. Logo, será preciso aplicar diferenças para estacionarizá-la, e assim ser possível aplicar a metodologia Box & Jenkins. Ainda observando a Figura 45, constata-se sazonalidade, sendo importante lembrar que, ao se aplicar diferenças na série original, pode-se excluir a sazonalidade da série.

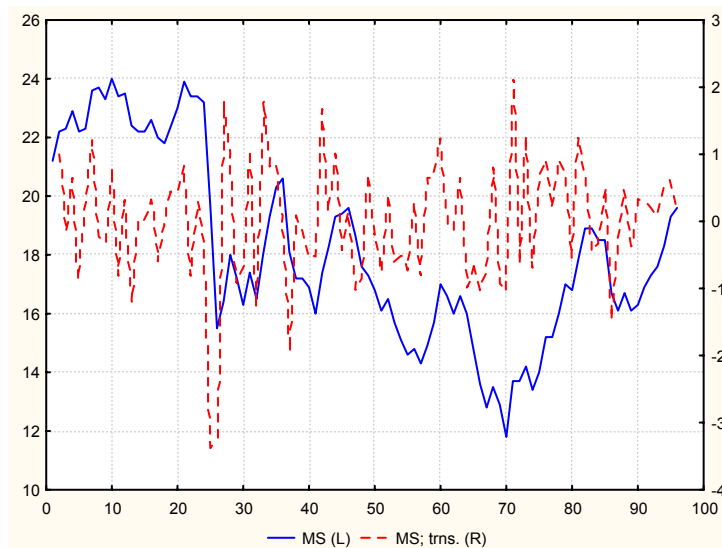


FIGURA 46 – Sazonalidade do preço do boi gordo, em arroba, no estado do MS, no período de janeiro de 1997 a dezembro de 2004.

Ao ser aplicada uma diferença na série original, é possível observar que esta apresentou um comportamento estacionário.

O próximo passo é descobrir qual o melhor modelo que represente a série, que estão dispostos no Quadro 3, qual foi um SARIMA (0,1,1)(1, 0, 1)₁₂. Isto nos diz que a diferença aplicada não foi suficiente para retirar a sazonalidade de ordem 12, da cotação do preço do boi gordo em arroba, para o MS.

Modelo	Parâmetros	t(calculado)	Nível de Significância	AIC	BIC
SARIMA(0,1,1)(1, 0, 1)₁₂	$\phi = -0,27$ $\Phi = 0,93$ $\Theta = 0,80$	-2,40 3,09 2,16	0,018348 0,002619 0,033327	420,37	422,92
SARIMA(0,2,1)(1,0,1) ₁₂	$\theta=0,88$ $\Phi =0,96$ $\Theta = 0,82$	14,36 3,21 2,17	0,000000 0,001826 0,032517	434,57	442,21
SARIMA(1,1,0)(1, 0, 1) ₁₂	$\phi =0,21$ $\Phi = 0,95$ $\Theta = 0,82$	2,06 3,37 2,32	0,041655 0,001086 0,022564	425,52	433,18
SARIMA(1,2,1)(1,0,0) ₁₂	$\phi =-0,24$ $\theta=0,94$ $\Phi =0,22$	2,20 25,56 2,01	0,030671 0,000000 0,047023	431,13	438,76

QUADRO 3 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado do MS, no período de janeiro de 1997 a dezembro de 2004.

Como foram encontrados mais de um modelo significativo, no Quadro 3, foi necessário a utilização dos critérios AIC e BIC, que, neste caso, foram os menores encontrados, no mesmo modelo, não sendo necessário utilizar critérios adicionais para escolhê-los.

Na Figura 47, apresenta-se a FCA e a FACP residual. Os resíduos do modelo estimado são estimativas do ruído branco. Sendo assim, devem estar representados desta forma na FAC e FACP residual dentro dos limites.

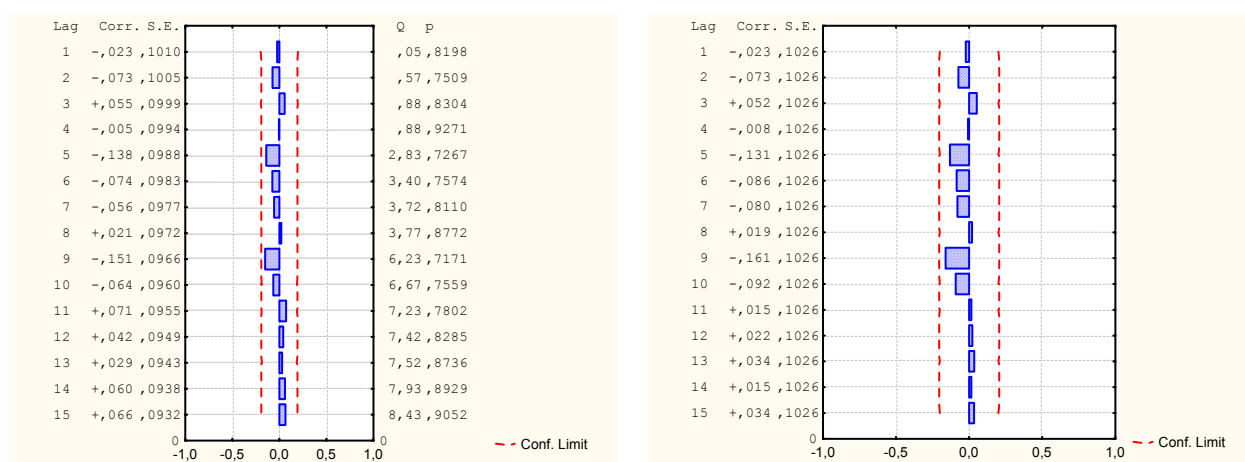


FIGURA 47 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo SARIMA $(0,1,1)(1,0,1)_{12}$, no estado do MS, no período de janeiro de 1997 a dezembro de 2004.

A Figura 47 relewa que os resíduos apresentam-se dentro dos limites de confiabilidade, ou seja, são significativos, confirmando a adequação do modelo.

Para confirmar a adequação do modelo, ainda dispõe-se do gráfico da normal dos resíduos, Figura 48, onde mostra que os pontos estão próximos à linha reta, garantindo, assim, a sua normalidade.

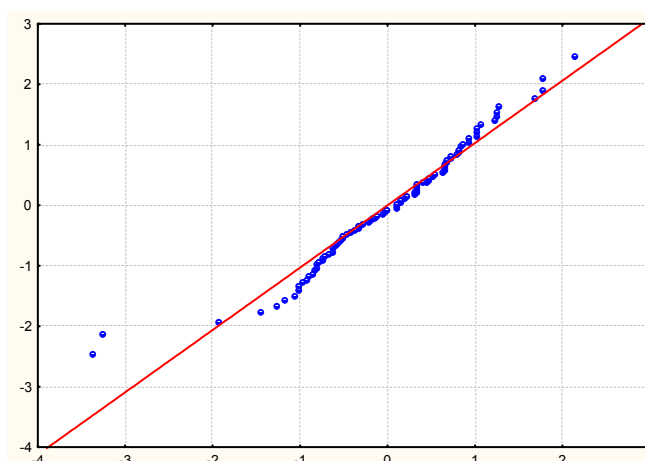


FIGURA 48 - Gráfico da Distribuição Normal, do estado do MS, no período de janeiro de 1997 a dezembro de 2004.

Na Tabela 17, apresentam-se as previsões para os meses de janeiro a dezembro de 2005, para o preço do boi gordo, em arroba, para o estado do Mato Grosso do Sul, obtidas com base no mês de dezembro de 2004.

TABELA 17 - Previsões para o preço do boi gordo, em arroba, no estado do MS, nos meses de janeiro a dezembro de 2005

Meses	Previsões	Inferior 95%	Superior 95%	Erro Padrão
Jan/05	19,30	17,74	20,87	0,94
Fev/05	18,70	16,17	21,23	1,52
Mar/05	18,73	15,51	21,95	1,94
Abr/05	18,99	15,21	22,77	2,28
Mai/05	18,63	14,36	22,91	2,57
Jun/05	18,63	13,92	23,35	2,84
Jul/05	18,87	13,75	23,99	3,08
Ago/05	19,01	13,52	24,50	3,30
Set/05	19,17	13,34	25,01	3,51
Out/05	19,39	13,22	25,55	3,71
Nov/05	19,71	13,23	26,19	3,90
Dez/05	19,74	12,96	26,52	4,08

Na Figura 49, observa-se o gráfico das previsões do preço do boi gordo em arroba, para os meses de janeiro de 2005 a dezembro de 2005, como representado pela Tabela 17.

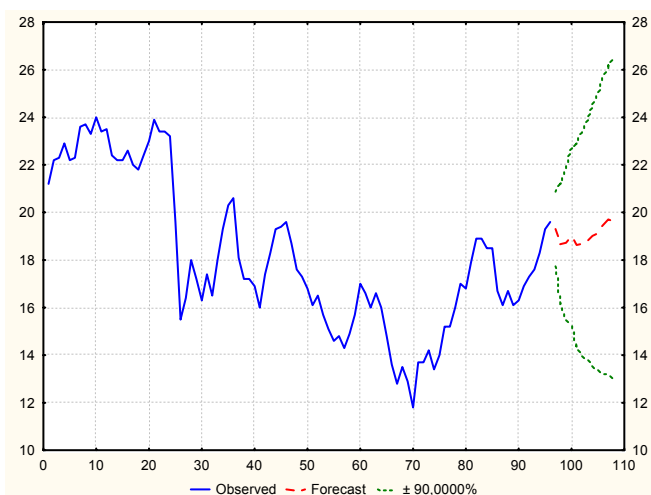


FIGURA 49 - Gráfico de previsões para o preço do boi gordo, em arroba, no estado do MS, de janeiro a dezembro de 2005.

Concluiu-se, para o estado de MS, que o melhor modelo que representou este estado foi um SARIMA (0,1,1)(1,0,1)₁₂, o qual manteve-se dentro dos limites de confiabilidade, representados pela FAC e FACP residual, seguindo uma distribuição

normal. Esse foi o melhor modelo encontrado para estimar a previsão para o ano de 2005.

4.6.4 Mato Grosso

Para verificar se o preço do boi gordo, em arroba, foi significativo no estado de Mato Grosso, no decorrer do período, utilizou-se medidas descritivas, tais como a média, o desvio-padrão e o coeficiente de variação das variáveis em estudo. Para tal, essas medidas foram realizadas ano a ano, investigando, desta maneira, se houve alguma modificação do período analisado. Na Tabela 18, apresenta-se as estatísticas para o preço do boi gordo, em arroba, que são auto-explicativas.

TABELA 18 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado do MT, no período de janeiro de 1997 a dezembro de 2004

Estatísticas	1997	1998	1999	2000	2001	2002	2003	2004
Média	21,65	21,82	16,74	19,48	16,22	15,04	17,07	18,29
Desvio padrão	0,63	0,54	1,78	0,84	1,14	1,57	1,98	1,16
Coeficiente de Variação	2,91	2,49	10,64	4,33	7,03	10,41	11,58	6,32

O coeficiente de variação de Pearson para os anos de 1997 a 2004, é muito significativo, pois é bem abaixo de 50 %, revelando que a média do preço do boi gordo, em arroba, é muito significativa para o estado do MT.

Para identificar o modelo adequado, inicialmente será analisado o gráfico da série original, representado na Figura 50, a série mostra-se não estacionária e com sazonalidade. Pode-se dizer, através de uma análise visual, que o preço do boi gordo, em arroba, teve uma tendência fracamente decrescente, não podendo, assim, ser considerada estacionária. Como não é confiável uma análise visual, dispõe-se da FAC para ser verificada a estacionaridade, e da FACP como auxílio para encontrar o modelo desejado.

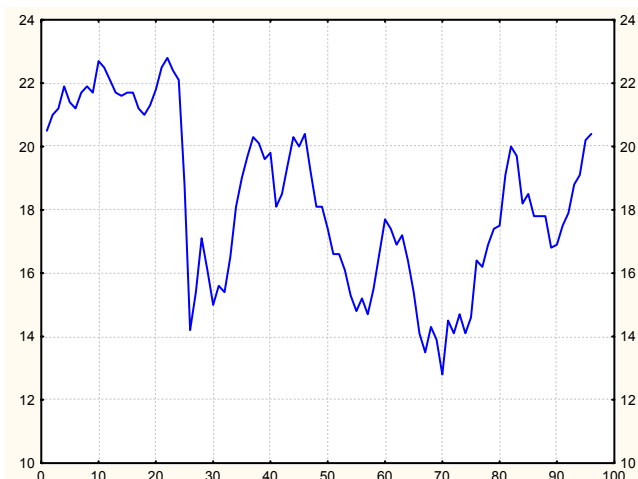


FIGURA 50 - Série original referente ao preço do boi gordo, em arroba, no estado do MT, no período de janeiro de 1997 a dezembro de 2004.

Séries não estacionárias declinam muito lentamente, à medida em que os lag crescem. Isso porque séries não estacionárias apresentam fortes correlações seriais. Observando a FAC, na Figura 51, os valores são inicialmente altos e não declinam exponencialmente à medida que os lag crescem, confirmando a não estacionaridade.

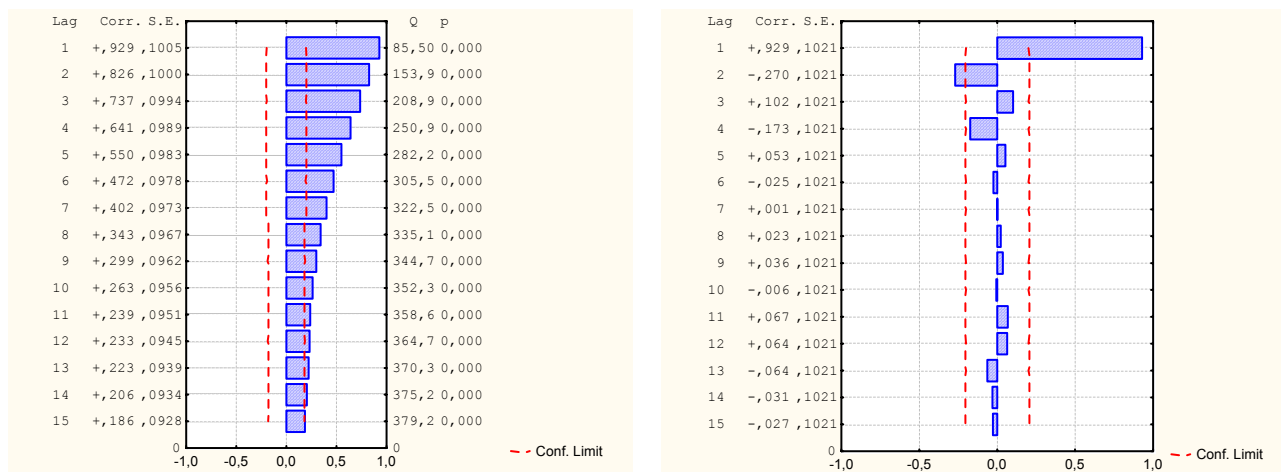


Figura 51 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do MT, no período de janeiro de 1997 a dezembro de 2004.

Observa-se, na Figura 51, que a FAC possui um decaimento exponencial, mas não é possível de se confirmar a estacionaridade. Logo, será preciso aplicar diferenças para estacionarizar a série e verificar se a mesma apresenta um melhor comportamento, e, assim, ser possível aplicar a metodologia Box & Jenkins. Ainda

observando a Figura 51, constata-se a sazonalidade, observando-se sempre que, ao aplicar diferenças na série original, pode-se excluir a sazonalidade da série.

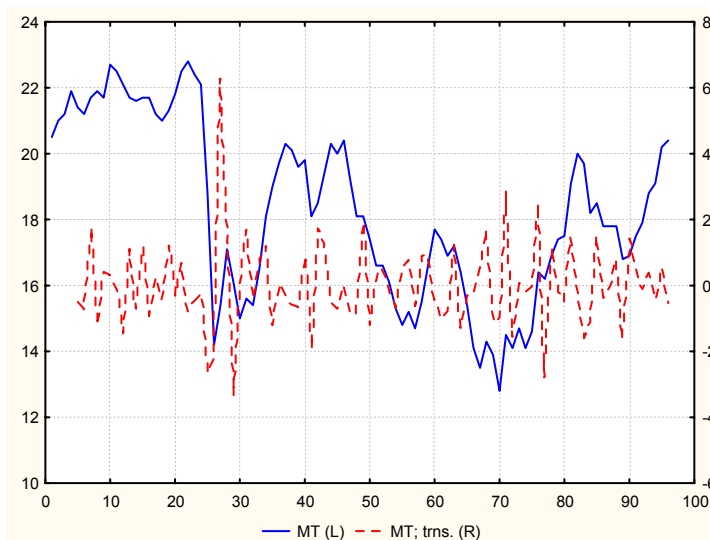


FIGURA 52 – Sazonalidade do preço do boi gordo, em arroba, no estado do MG, no período de janeiro de 1997 a dezembro de 2004.

Ao aplicar-se uma diferença na série original é possível observar que ela apresentou um comportamento estacionário, possibilitando, assim, utilizar a metodologia Box & Jenkins, para poder encontrar modelos adequados que representem a série e, assim, ser possível fazer previsões.

Os modelos mais significativos encontrados estão representados no Quadro 4, que foi encontrado um SARIMA, com uma diferença, a qual não foi suficiente para retirar a sazonalidade da série em estudo. Como foram encontrados mais de um modelo significativo, foi necessário a utilização dos critérios AIC e BIC, conforme Quadro 4.

Modelo	Parâmetros	t(calculado)	Nível de Significância	AIC	BIC
SARIMA(0,1,1)(1,0,1)₁₂	$\phi = -0,34$	-3,16E+00	0,002163	418,49	426,15
	$\Phi = 0,92$	7,69E+18	0,000000		
	$\Theta = 0,90$	1,45E+01	0,000000		
SARIMA(2,2,1)(1, 1, 1) ₁₂	$\phi(1) = -,039$	-3,93	0,000167	439,29	451,96
	$\phi(1) = -,039$	-3,94	0,000161		
	$\Theta = 0,97$	42,86	0,000000		
	$\Phi = 1,00$	4,05	0,000109		
	$\Theta = 0,88$	2,89	0,004791		

QUADRO 4 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado do MT, no período de janeiro de 1997 a dezembro de 2004.

Na Figura 53, apresenta-se a FAC e a FACP residual. Os resíduos do modelo estimado são estimativas do ruído branco. Sendo assim, devem estar representados desta forma na FAC e FACP residual, dentro dos limites de confiabilidade do modelo.

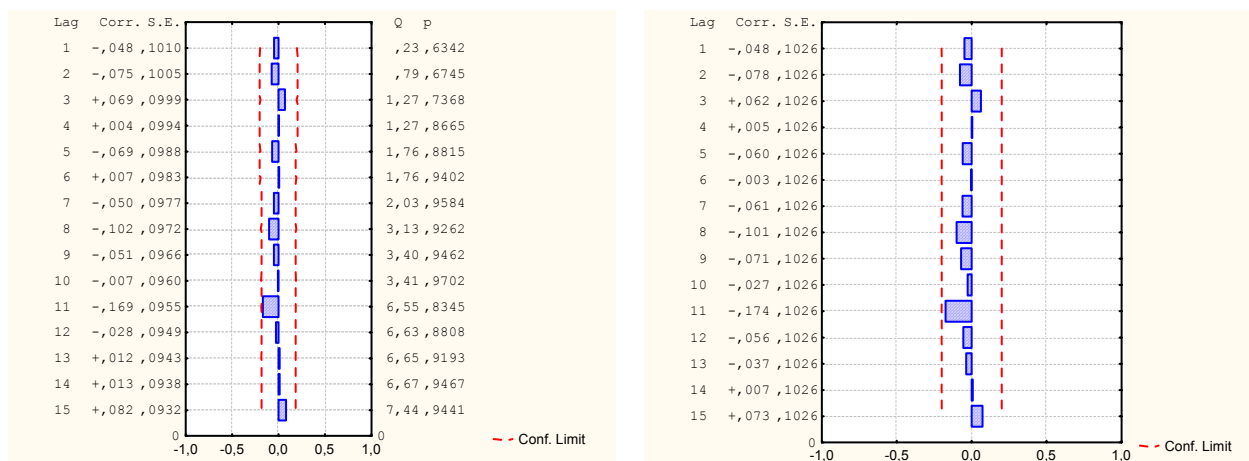


FIGURA 53 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo SARIMA $(0,1,1)(1,0,1)_{12}$, no estado do MT, no período de janeiro de 1997 a dezembro de 2004.

A Figura 53 releva que os resíduos apresentam-se dentro dos limites de confiabilidade, ou seja, são significativos, confirmando a adequação do modelo.

Para confirmar a adequação do modelo, ainda dispõe-se do gráfico da normal dos resíduos, Figura 54, onde mostra que os pontos estão próximos à linha reta, garantindo, assim, a normalidade dos resíduos.

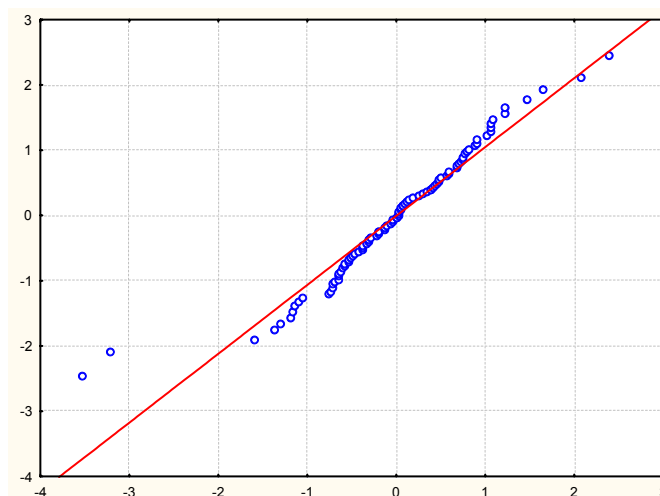


FIGURA 54 - Gráfico da Distribuição Normal, do estado do MT, no período de janeiro de 1997 a dezembro de 2004.

Na Tabela 19, apresentam-se as previsões para os meses de janeiro a dezembro de 2005, para o preço do boi gordo, em arroba, para o estado do Mato Grosso, obtidas com base no mês de dezembro de 2004.

TABELA 19 - Previsões para o preço do boi gordo, em arroba, no estado do MT, nos meses de janeiro a dezembro de 2005

Meses	Previsões	Inferior 95%	Superior 95%	Erro Padrão
Jan/05	20,42	18,83	22,00	0,95
Fev/05	20,53	17,81	23,25	1,64
Mar/05	20,65	17,09	24,20	2,14
Abr/05	20,76	16,49	25,03	2,57
Mai/05	20,88	15,96	25,80	2,96
Jun/05	20,99	15,46	26,52	3,33
Jul/05	21,11	15,00	27,21	3,67
Ago/05	21,22	14,55	27,89	4,01
Set/05	21,34	14,12	28,55	4,34
Out/05	21,45	13,70	29,20	4,66
Nov/05	21,57	13,29	29,84	4,97
Dez/05	21,68	12,89	30,47	5,29

A Figura 55 apresenta o gráfico das previsões para os meses de janeiro a dezembro de 2005, para o preço do boi gordo, em arroba, para o estado do Mato Grosso, obtidas com base no mês de dezembro de 2004, representados na Tabela 19.

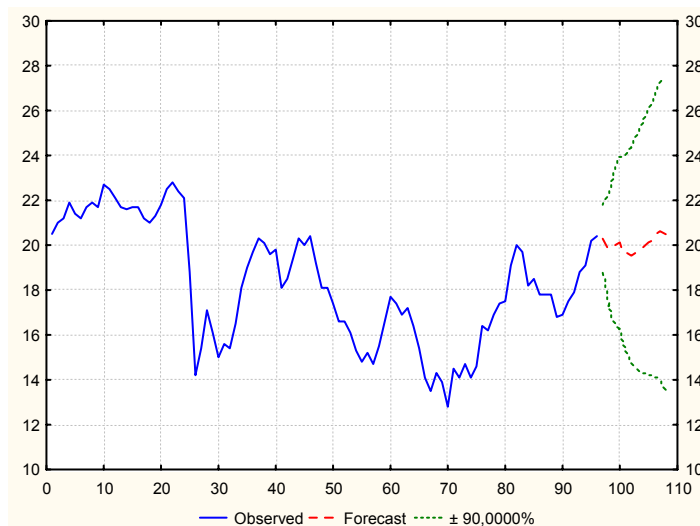


FIGURA 55 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do MG, de janeiro a dezembro de 2005.

Concluiu-se, para o estado de MT, que o melhor modelo que representou este estado foi um SARIMA (0,1,1)(1,0,1)₁₂, o qual manteve-se dentro dos limites de confiabilidade, representados pela FAC e FACP residual, seguindo uma distribuição normal. Esses foi o melhor modelo encontrado para estimar a previsão para o ano de 2005, no qual pode-se constatar haver uma estabilidade entre os meses de janeiro de 2005 a junho de 2005, e de julho de 2005 a dezembro de 2005.

4.6.5 Goiás

Para verificar se o preço do boi gordo, em arroba, foi significativo no estado de Goiás, no decorrer do período, utilizou-se medidas descritivas, tais como a média, o desvio-padrão e o coeficiente de variação das variáveis em estudo. Para tal, essas medidas foram realizadas ano a ano, investigando, desta maneira, se houve alguma modificação do período analisado. Na Tabela 20, apresenta-se as estatísticas para o preço do boi gordo em arroba, que são auto-explicativas.

TABELA 20 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado de GO, no período de janeiro de 1997 a dezembro de 2004

Estatísticas	1997	1998	1999	2000	2001	2002	2003	2004
Média	21,89	21,60	16,54	20,21	17,21	15,66	17,68	19,36
Desvio padrão	0,71	0,85	2,01	0,93	1,26	1,61	2,05	1,30
Coeficiente de								
Variação	3,26	3,96	12,18	4,60	7,30	10,30	11,57	6,73

O coeficiente de variação de Pearson, para os anos de 1997 a 2004, é muito significativo, pois é bem abaixo de 50 %, revelando que a média do preço do boi gordo, em arroba, é muito significativa para o estado do GO.

Para identificar o modelo adequado, inicialmente será analisado o gráfico da série original, representado na Figura 56, o qual não apresenta definida a sua tendência. Como não é confiável uma análise visual, dispõe-se da FAC para ser verificada a estacionaridade, e da FACP como auxílio para encontrar o modelo desejado.

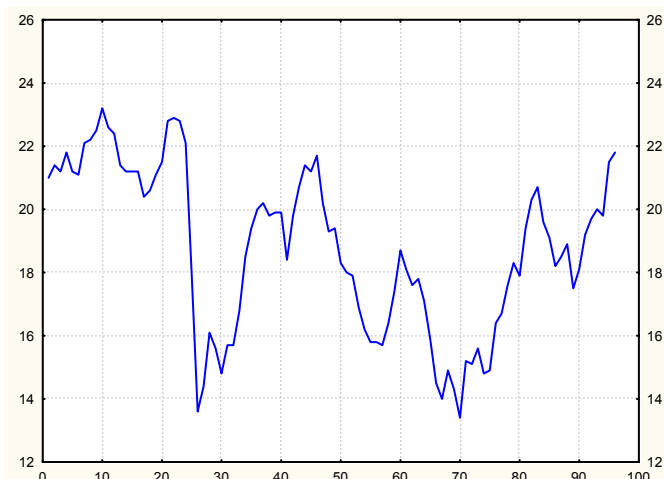


FIGURA 56 - Série original referente ao preço do boi gordo em arroba, no estado do GO, no período de janeiro de 1997 a dezembro de 2004.

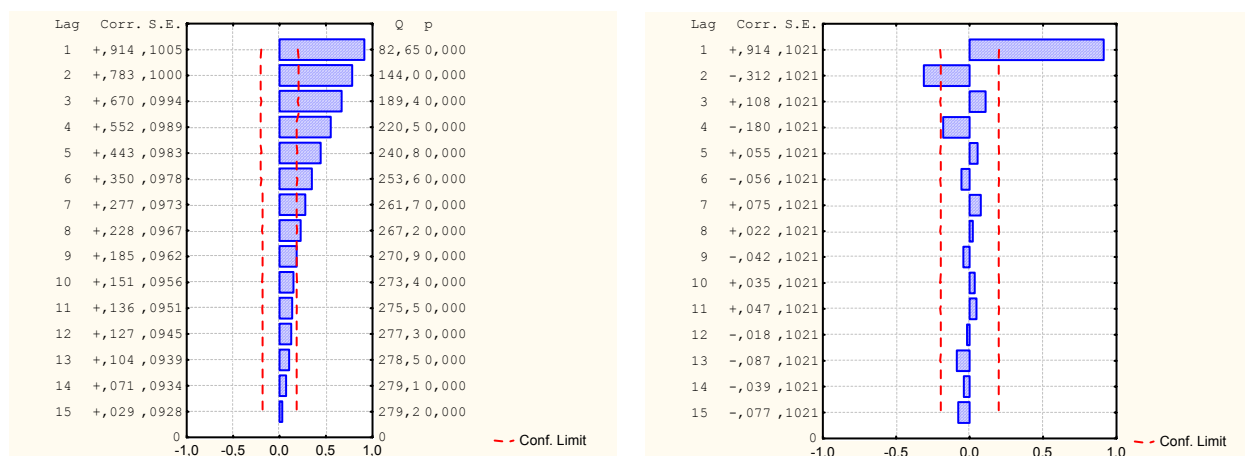


FIGURA 57 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do GO, no período de janeiro de 1997 a dezembro de 2004.

Observa-se, na Figura 57, que a FAC possui um decaimento exponencial, tendendo à estacionaridade. Logo, será aplicado, no mínimo, uma diferença para verificar se a série apresenta um melhor comportamento estacionário e, assim, ser possível aplicar a metodologia Box & Jenkins.

Ao aplicar-se uma diferença, na série original, foi possível observar que ela apresentou um comportamento estacionário, chegando ao objetivo desejado, pois para utilizar a metodologia supõe-se que a série sempre seja estacionária, para poder encontrar modelos adequados que representem a série e, assim, ser possível fazer previsões.

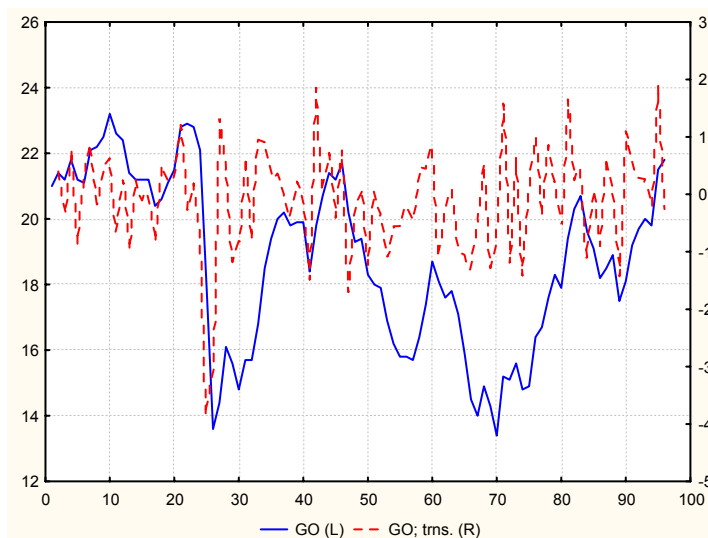


FIGURA 58 – Sazonalidade do preço do boi gordo, em arroba, no estado de GO, no período de janeiro de 1997 a dezembro de 2004.

Os modelos significativos encontrados estão representados no Quadro 5, que foi o modelo encontrado é um ARIMA, com uma diferença, a qual foi suficiente para retirar a sazonalidade da série em estudo. Como foram encontrados mais de um modelo significativo, foi necessário a utilização dos critérios AIC e BIC, que, neste caso, foram os menores encontrados no mesmo modelo, não sendo necessário utilizar critérios adicionais para escolhê-los, conforme Quadro 5.

Modelo	Parâmetros	t(calculado)	Nível de Significância	AIC	BIC
ARIMA(1,1,2)₁₂	$\phi = 0,88$ $\Theta (1) = 0,54$ $\Theta (2) = 0,43$	9,92 4,36 4,05	0,000000 0,000033 0,000107	424,79	432,46
ARIMA(1,2,1) ₁₂	$\phi = 0,32$ $\Theta = 0,98$	3,09 46,16	0,002609 0,000000	428,77	433,86
SARIMA(1,2,1)(1, 0, 1) ₁₂	$\phi = 0,36$ $\Theta = 0,99$ $\Phi = 1,00$ $\Theta = 0,91$	3,54 57,27 3,27 2,69	0,000645 0,000000 0,001522 0,008458	427,08	437,25
SARIMA(2,2,1)(1,1,1) ₁₂	$\phi (1) = -0,31$ $\phi (2) = -0,33$ $\Theta = 0,94$ $\Phi = 0,97$ $\Theta = 0,83$	-2,98 -3,07 27,27 3,93 2,63	0,003666 0,00816 0,000000 0,000169 0,009947	451,77	464,44

QUADRO 5 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado de GO, no período de janeiro de 1997 a dezembro de 2004.

Na Figura 59, apresenta-se a FCA e a FACP residual. Os resíduos do modelo estimado são estimativas do ruído branco. Sendo assim, devem estar representados, desta forma, na FAC e FACP residual dentro dos limites.

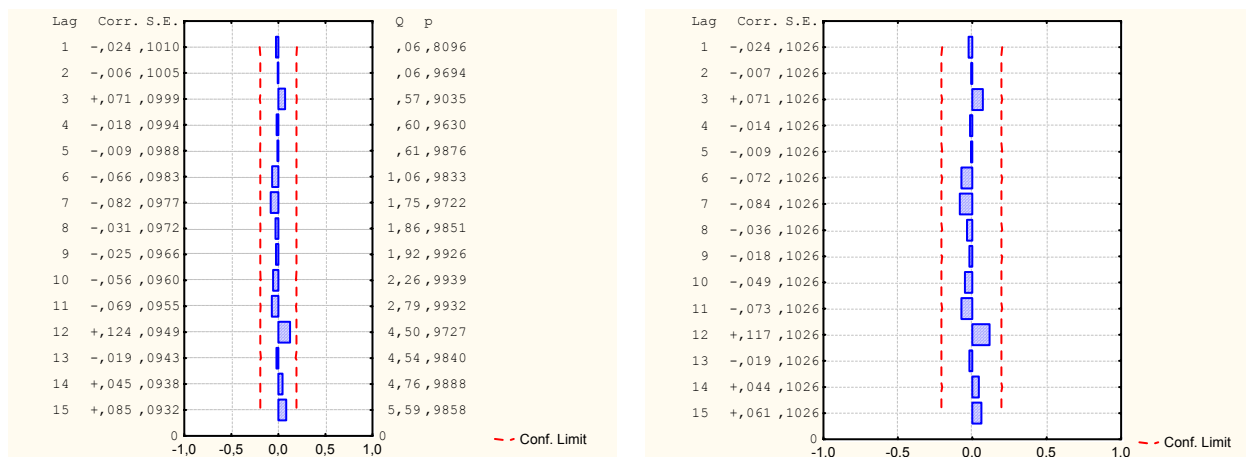


FIGURA 59 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo ARIMA (1,1,2), no estado do RS, no período de janeiro de 1997 a dezembro de 2004.

A Figura 59 relewa que os resíduos apresentam-se dentro dos limites de confiabilidade, ou seja, são significativos, confirmando a adequação do modelo. Para confirmar a adequação do modelo, ainda dispõe-se do gráfico da normal dos resíduos, na Figura 60, a qual mostra que os pontos estão próximos à linha reta, garantindo, assim, a normalidade dos resíduos.

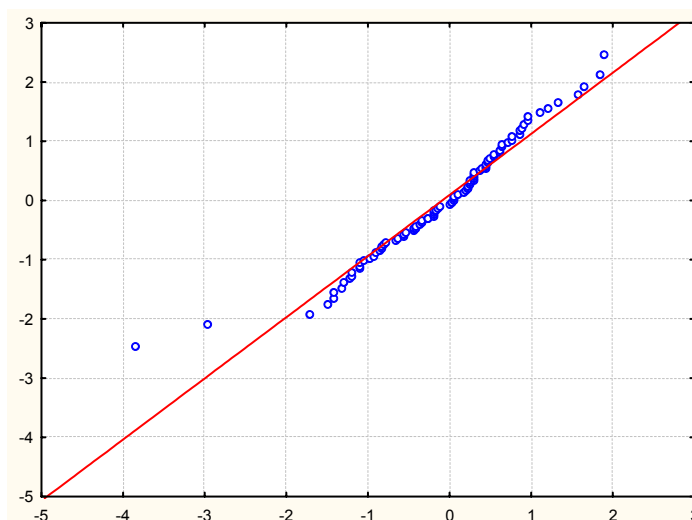


FIGURA 60 - Gráfico da Distribuição Normal, do estado de GO, no período de janeiro de 1997 a dezembro de 2004.

Na Tabela 21, apresentam-se as previsões para os meses de janeiro a dezembro de 2005, para o preço do boi gordo, em arroba, para o estado de Goiás, obtidas com base no mês de dezembro de 2004.

TABELA 21 - Previsões para o preço do boi gordo, em arroba, no estado de GO, nos meses de janeiro a dezembro de 2005

Meses	Previsões	Inferior 95%	Superior 95%	Erro Padrão
Jan/05	21,39	19,82	22,96	0,94
Fev/05	21,14	18,51	23,76	1,58
Mar/05	20,92	17,68	24,15	1,95
Abr/05	20,72	17,06	24,38	20,20
Mai/05	20,55	16,57	24,52	2,39
Jun/05	20,40	16,18	24,61	2,54
Jul/05	20,26	15,86	24,66	2,65
Ago/05	20,14	15,59	24,70	2,74
Set/05	20,04	15,36	24,71	2,82
Out/05	19,94	15,16	24,72	2,88
Nov/05	19,86	15,00	24,73	2,93
Dez/05	19,79	14,86	24,73	2,97

Na Figura 61, observa-se o gráfico das previsões do preço do boi gordo, em arroba, para os meses de janeiro de 2005 a dezembro de 2005, como representado pela Tabela 21.

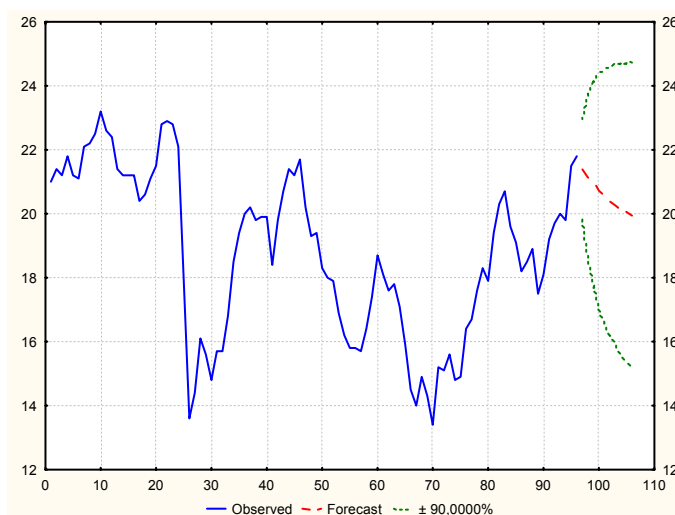


FIGURA 61 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do GO, de janeiro a dezembro de 2005.

Concluiu-se, para o estado de GO, que o melhor modelo que representou este estado foi um ARIMA (1,1,2), o qual manteve-se dentro dos limites de confiabilidade, representados pela FAC e FACP, seguindo uma distribuição normal.

Esse foi o melhor modelo encontrado para estimar a previsão para o ano de 2005, no qual pode-se constatar que ocorreram oscilações durante o período.

4.6.6 Bahia

Para verificar se o preço do boi gordo, em arroba, foi significativo no estado da Bahia, no decorrer do período, utilizou-se medidas descritivas, tais como a média, o desvio-padrão e o coeficiente de variação das variáveis em estudo. Para tal, essas medidas foram realizadas ano a ano, investigando, dessa maneira, se houve alguma modificação do período analisado. Na Tabela 22, apresenta-se as estatísticas para o preço do boi gordo, em arroba, que são auto-explicativas.

TABELA 22 - Estatísticas descritivas, referentes ao preço do boi gordo, em arroba, no estado da BA, no período de janeiro de 1997 a dezembro de 2004

Estatísticas	1997	1998	1999	2000	2001	2002	2003	2004
Média	23,08	23,62	16,91	20,13	17,86	15,88	17,56	18,84
Desvio padrão	0,43	1,13	1,56	1,14	1,87	2,58	1,37	1,52
Coeficiente de Variação	1,88	4,80	9,22	5,65	10,46	16,26	7,82	8,06

Como o coeficiente de variação de Pearson esteve bem inferior a 50%, em todo o período, isso mostra que esse é muito significativo, revelando, assim, que a média do preço do boi gordo, em arroba, é muito significativa, para o estado da BA.

Inicialmente, será analisado o gráfico da série original, representado na Figura 62, o qual apresenta uma série temporal não estacionária. Pode-se dizer, através de uma análise visual, que o preço do boi gordo, em arroba, teve uma tendência decrescente, não podendo, assim, a série ser considerada estacionária. Como não é confiável uma análise visual, dispõe-se da FAC para ser verificada a estacionaridade, e da FACP como auxílio para encontrar o modelo desejado.

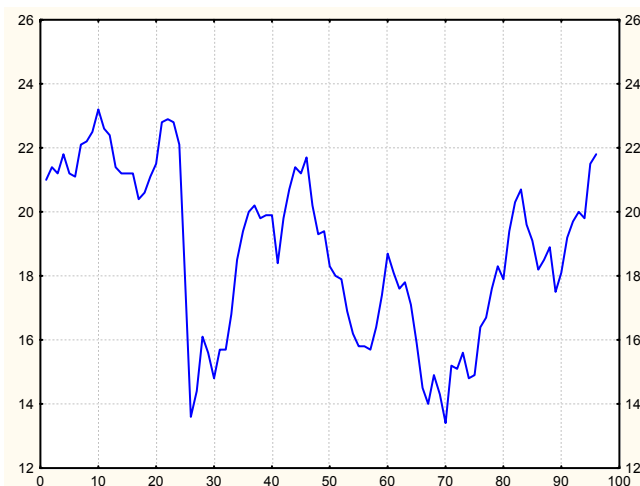


FIGURA 62 - Série original referente ao preço do boi gordo, em arroba, no estado do BA, no período de janeiro de 1997 a dezembro de 2004.

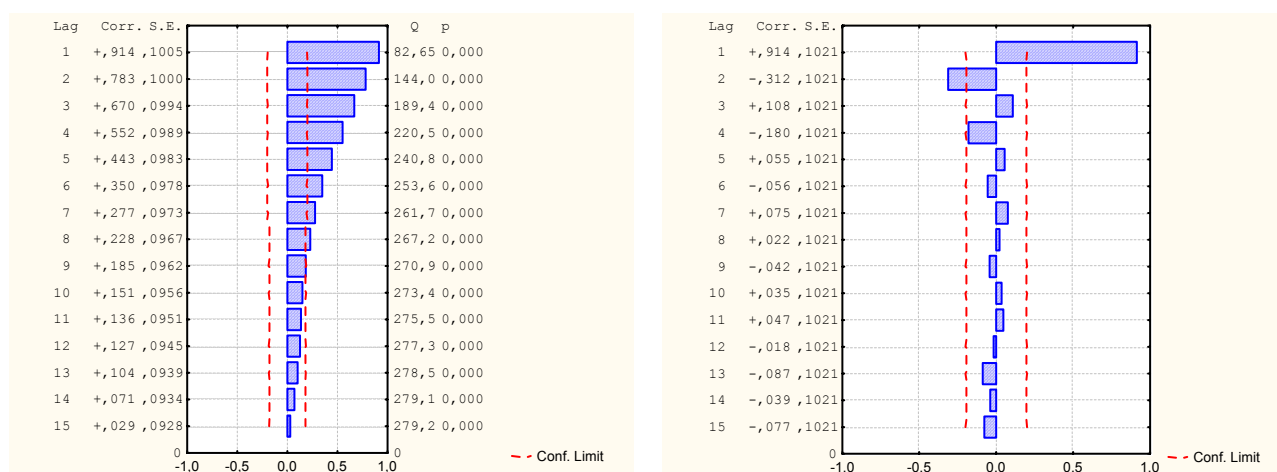


FIGURA 63 - Função de Autocorrelação e Autocorrelação Parcial, da série original, do estado do BA, no período de janeiro de 1997 a dezembro de 2004.

Observa-se, na Figura 63, que o comportamento da FAC não confirma a estacionaridade. Logo, será preciso aplicar diferenças para estacionarizar a série, e, assim, ser possível aplicar a metodologia Box & Jenkins.

A Figura 64 apresenta a série original e a série com duas diferenças, as quais foram necessárias para que a série diferenciada apresentasse um comportamento estacionário, tornando possível encontrar modelos adequados que representem a série e, portanto, fazer previsões futuras.

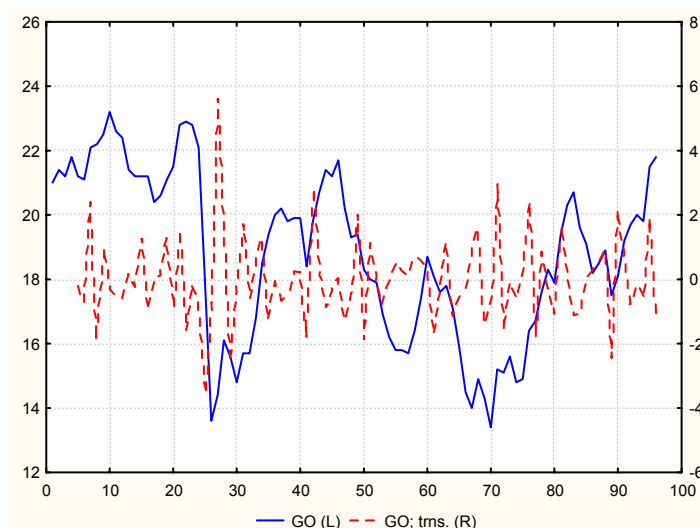


FIGURA 64 – Sazonalidade do preço do boi gordo, em arroba, no estado da BA, no período de janeiro de 1997 a dezembro de 2004.

Os modelos significativos encontrados estão representados no Quadro 6.

Modelo	Parâmetros	t(calculado)	Nível de Significância	AIC	BIC
ARIMA(0,2,2)	$\theta(1) = 0,46$ $\theta(2) = 0,49$	5,29 5,64	0,000001 0,000000	430,90	435,99
ARIMA(2,2,1) ₁₂	$\phi(1) = 0,49$ $\phi(2) = -0,24$ $\theta = 0,96$	4,56 -2,26 28,42	0,000016 0,026317 0,000000	433,37	441,00

QUADRO 6 – Listagem dos possíveis modelos Box & Jenkins, para série com seus parâmetros e suas estatísticas de avaliação, no estado da BA, no período de janeiro de 1997 a dezembro de 2004.

Pode-se perceber que foi encontrado um ARIMA, com duas diferenças, suficientes para retirar a sazonalidade da série em estudo. Como foram encontrados mais de um modelo significativo, foi necessário a utilização dos critérios AIC e BIC, que, neste caso, foram os menores encontrados, em um mesmo modelo, não sendo necessário utilizar critérios adicionais para escolhê-los.

Na Figura 65, apresenta-se a FAC e a FACP residual. Os resíduos do modelo estimado são estimativas do ruído branco. Sendo assim, devem estar representados, confiavelmente, na FAC e FACP residual.

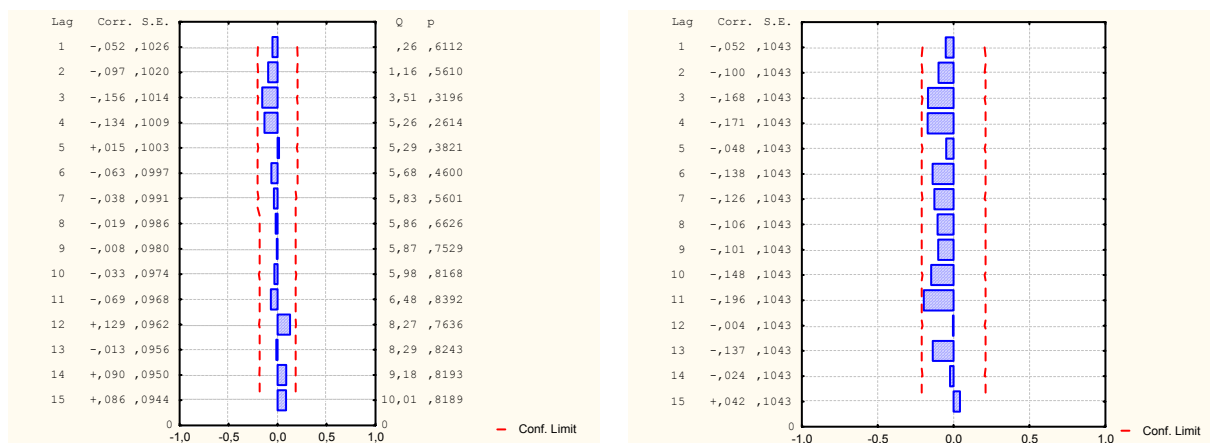


FIGURA 65 - Função de Autocorrelação e Autocorrelação Parcial dos resíduos, modelo ARIMA (0,2,2), no estado do BA, no período de janeiro de 1997 a dezembro de 2004.

A Figura 65 relewa que os resíduos apresentam-se dentro dos limites de confiabilidade, ou seja, são significativos, confirmando a adequação do modelo.

Ainda dispõe-se do gráfico da normal dos resíduos, Figura 66, onde mostra que os pontos estão próximos à linha reta, garantindo, assim, a normalidade dos resíduos.

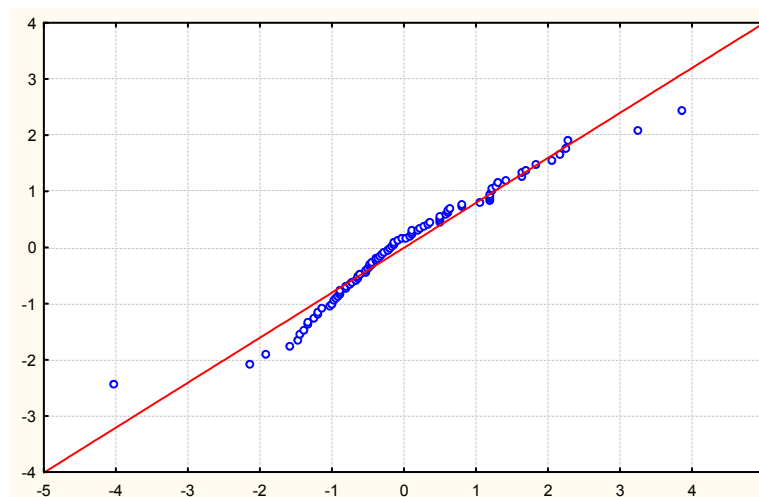


FIGURA 66 - Gráfico da Distribuição Normal, do estado do BA, no período de janeiro de 1997 a dezembro de 2004.

Na Tabela 23 apresentam-se as previsões para os meses de janeiro a dezembro de 2005, para o preço do boi gordo, em arroba, para o estado da Bahia, obtidas com base no mês de dezembro de 2004.

TABELA 23 - Previsões para o preço do boi gordo, em arroba, no estado da BA, nos meses de janeiro a dezembro de 2005

Meses	Previsões	Inferior 95%	Superior 95%	Erro Padrão
Jan/05	23,05	21,37	24,73	1,01
Fev/05	23,21	20,14	26,29	1,85
Mar/05	23,38	19,31	27,44	2,45
Abr/05	23,54	18,64	28,44	2,95
Mai/05	23,70	18,06	29,35	3,40
Jun/05	23,87	17,53	30,21	3,81
Jul/05	24,03	17,03	31,02	4,21
Ago/05	24,19	16,57	31,82	4,59
Set/05	24,36	16,12	32,59	4,96
Out/05	24,52	15,69	33,35	5,31
Nov/05	24,68	15,27	34,10	5,67
Dez/05	24,85	14,86	34,83	6,01

Na Figura 67, observa-se o gráfico das previsões do preço do boi gordo, em arroba, para os meses de janeiro de 2005 a dezembro de 2005, como representado na Tabela 23.

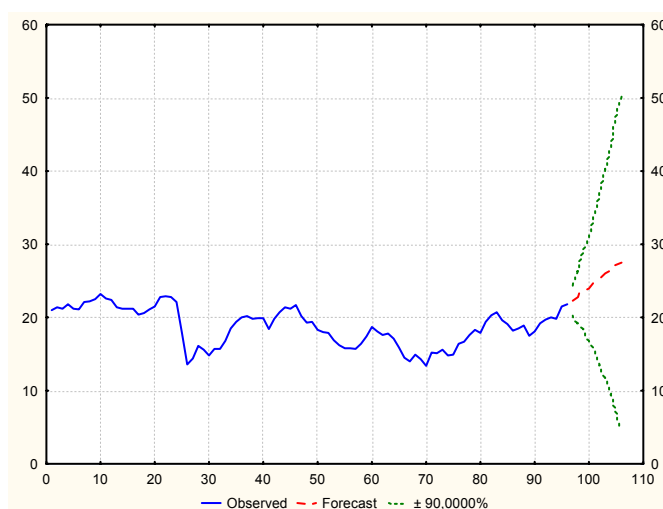


FIGURA 67 - Gráfico de previsões, para o preço do boi gordo, em arroba, no estado do BA, de janeiro a dezembro de 2005.

Concluiu-se, para o estado da BA, que o melhor modelo que representou este estado foi um ARIMA (0,2,2), o qual manteve-se dentro dos limites de confiabilidade, representados pela FAC e FACP, seguindo uma distribuição normal. Esse foi o melhor modelo encontrado para estimar a previsão para o ano de 2005.

- Comentários deste capítulo

Através da metodologia Box & Jenkins, foi possível modelar uma série temporal que melhor explicou o preço do boi gordo, em arroba, nos estados citados nesse capítulo 4, os quais possibilitaram a realização da previsão para o ano de 2005.

5 CONCLUSÕES E RECOMENDAÇÕES

Com a aplicação da estatística descritiva, e a utilização das técnicas da análise de agrupamentos, análise fatorial, análise de componentes principais e o metodologia de séries temporais, foi possível atingir o objetivo proposto neste trabalho, que foi verificar se a estrutura pecuária, no Brasil, sofreu alterações de produção, em termos de número de bovinos, no período de 1997 a 2004, identificar a categoria de bovinos que cada região do país produz, bem como os estados que estão apontados como destaque na produção nacional e, a partir disso, realizar previsões para o futuro e monitorando assim os mesmos.

De acordo com a análise descritiva, foi possível verificar que todos os estados brasileiros estão reduzindo a idade de abate dos animais, de 3 a 4 anos, para 2 a 3 anos, ou seja, a categoria representada pelos bois está deixando de existir, sendo substituída pelo abate jovem, que ocorre com novilhos de 2 a 3 anos. Isso significa um ganho para o país, pois passa a produzir uma carne de melhor qualidade e um retorno maior para o produtor, que recebe, pelo seu produto, num menor espaço de tempo. Outro fato importante é que, hoje, devido às exigências feitas pelo mercado consumidor, cada vez mais busca-se qualidade no produto, redução de custos, preços baixos, fazendo com que haja necessidade de se iniciar o processo de mudança que estabeleça melhorias das condições de criação.

Com a análise de agrupamentos, foi possível identificar as oscilações ocorridas no período, bem como os grupos que as variáveis formaram. Pode-se verificar que a variável representada pelas vacas, manteve-se distinta das demais, devido a sua média ser superior às outras variáveis, embora tivesse um decréscimo no período de 1997 a 2000 tornando a crescer de 2001 a 2004. Isso mostra que algumas regiões do país, no período de 2001 a 2004, tornaram a investir na bovinocultura.

Através da análise fatorial foi possível reduzir o número de variáveis, que passaram a ser explicadas por três fatores, os quais juntos explicam 97,91% dos dados originais.

A análise de componentes principais muitas vezes é criticada, pelo fato de utilizar uma amostragem menor de informações do que aquela da amostra. Corrobora Bégni (1978) afirma que a multicolinearidade pode existir, mesmo na

população, de maneira que as informações descartadas não devem ser consideradas como prejuízo. Entende que, ao utilizar um componente principal, estará substituindo dois ou mais conjuntos de variáveis que contêm informações por uma combinação linear. Sendo assim, a informação original é modificada, e não perdida.

Aplicando-se a análise de componentes principais com a análise fatorial, verifica-se, nas diversas regiões fisiográficas do país, que os processos de cria, recria e engorda concentram-se em pontos mais distanciados dos centros consumidores, e em regiões onde predominam os solos menos férteis, com condições climáticas adversas. Assim, na região Nordeste, essa atividade predomina em áreas de caatinga e cerrados; no Centro Oeste, a atividade é típica do pantanal e, em São Paulo, nas áreas menos produtivas.

O Rio Grande do Sul é uma exceção, pois a atividade pecuária está dividida de duas formas: em uma região do estado está concentrada em pequenas propriedades com rebanhos de até 50 cabeças, e na outra região a criação é realizada em fazendas com rebanhos maiores.

As áreas de cria, recria, e especialmente de engorda, estão concentradas nas regiões com terras mais férteis: a região Oeste de São Paulo e Minas Gerais, Campo Grande, Dourados no Mato Grosso do Sul, Rondonópolis no Mato Grosso e a região do Araguaia em Goiás.

Com as técnicas de análise, da análise multivariada, foi possível identificar o tipo de produção de bovinos, que ocorre nas regiões do Brasil, bem como em quais dessas regiões esta atividade econômica ocorre com uma maior intensidade. A produção de bois de 3 a 4 anos e bois + 4 anos, é representada pela região sul, sudeste e nordeste. A produção de bezerras (as), novilhos (as), garrotes, touros e vacas está concentrada na região centro-oeste do país. Foi possível, identificar, também, aquelas regiões que possuem uma carência maior de recursos, nas quais a bovinocultura ainda é exercida como uma atividade econômica secundária.

Tendo detectado que os estados que possuíram destaque, em todo o período, foram: MG, MT, MS, GO e RS, embora tenham ocorrido várias oscilações na economia durante esse período, refletindo-se, também, na bovinocultura. Aplicou-se, nesses estados, a metodologia de séries temporais, o qual possibilitou monitorar o preço do boi gordo, sendo este o fator que interfere diretamente na produção de

bovinos. Sendo assim, foi possível fazer previsões para o ano de 2005, que servirão de suporte para os pecuaristas que buscarem investir nessa atividade econômica.

Espera-se que os resultados gerados, com esta pesquisa, sirvam para uma tomada de decisão futura a para a implantação de medidas de apoio para o setor, e que a bovinocultura seja tratada como um importante segmento da atividade econômica do país.

Acredita-se que havendo incentivos fiscais, por parte dos órgãos responsáveis, que possibilitem ao criador exercer sua atividade econômica, atingindo uma margem de lucro real, este manterá sua atividade econômica.

Sugere-se que através das inspetorias veterinárias localizadas em cada cidade sejam mantidos banco de dados atualizados e que os mesmos sejam passíveis de utilização, pois não se tem estatísticas referentes aos municípios. Também sugere-se que os estados e o Brasil disponibilizem estes dados via internet, pois sem uma análise estatística atualizada não é possível desenvolver estudos referentes ao setor.

Sugere-se, como pesquisas subseqüentes, que sejam realizados estudos de viabilidade econômica da escala de produção, em relação a um ponto de equilíbrio entre a produção e a rentabilidade da mesma, para pequena, média e grande propriedades.

Também sugere-se um estudo que englobe não só análise referente ao número de cabeças produzidas mas também em relação a sanidade animal, rastreabilidade e genética da bovinocultura brasileira.

6 BIBLIOGRAFIA

- ANUALPEC: Anuário da pecuária brasileira. São Paulo: Instituto iFNP, 1997.
- ANUALPEC: Anuário da pecuária brasileira. São Paulo: Instituto iFNP, 1998.
- ANUALPEC: Anuário da pecuária brasileira. São Paulo: Instituto iFNP, 1999.
- ANUALPEC: Anuário da pecuária brasileira. São Paulo: Instituto iFNP, 2000.
- ANUALPEC: Anuário da pecuária brasileira. São Paulo: Instituto iFNP, 2001.
- ANUALPEC: Anuário da pecuária brasileira. São Paulo: Instituto iFNP, 2002.
- ANUALPEC: Anuário da pecuária brasileira. São Paulo: Instituto iFNP, 2003.
- ANUALPEC: Anuário da pecuária brasileira. São Paulo: Instituto iFNP, 2004.
- ABCZ: a revista brasileira do zebu e seus cruzamentos. Uberaba, v.2, n.12, jan./fev. 2003.
- BARCELLOS, J. O. J.; PRATES, E. R.; SILVA, M. D. Efeitos ambientais sobre a taxa de vacas de corte numa criação comercial no sul do Brasil. In: REUNIÃO ANUAL SOCIEDADE BRASILEIRA DE ZOOTECNIA, 36., 1999, Porto Alegre. Anais... Porto Alegre, 1999. P.119.
- BÉRNI, D. de A. Multicolinearidade e a função de custos: os frigoríficos suínos do Rio Grande do Sul. **Indicadores Econômicos do Rio Grande do Sul**. Porto Alegre, v.6, 1978.
- BOUROCHE, J.; SAPORTA, G. Análise de dados. Rio de Janeiro, 1982.
- BOX, G. E. P. & JENKINS, G.M. **Time séries analysis: forecasting and control**, San Francisco: Holden Day, 1976.
- BUSSAB, W. O.; MIAZAKI, É. S.; ANDRADE, D. F. Introdução à análise de agrupamentos. In: SIMPÓSIO BRASILEIRO DE PROBABILIDADE E ESTATÍSTICA, 9.,1990, São Paulo. **Resumos...**São Paulo, 1990.
- BUSSAB, W. O. & MORETTIN, P. A. **Estatística básica**, 3. ed. São Paulo: Atual, 1986.
- CATTEL, R. B. The scree test for the number of factors. In: ---. **Multivariate behavior research**. v.1, p. 245-276, 1966.
- COSTA NETO, P. L. O. **Estatística**, 2. ed. São Paulo: Edgard Blücher, 2002.
- GLOBO RURAL. **Agropecuária e negócios**. São Paulo, n. 231, jan. 2005.

GUIA AG LEILÕES DO CRIADOR. Porto Alegre: Centauros. n. 4, 2005.

GUJARATI, D. N. **Econometria básica**. 3. ed. São Paulo: Makron Books, 2000.

HAIR, J. F.; *et al.* **Análise multivariada de dados**. 5. ed. Porto Alegre, 2005.

JACKSON, J.E. Principal componets and factor analysis: Part I - principal componets. **Journal of Quality Technology**, v.12, n.4, p.201-213. oct. 1980.

JOHONSON, R.A.; WICHERN, D.W. **Applied multivariate statistical analysis**. 3. ed. New Jersey: Prentice-Hall, 1992.

LOPES, L. F. D. **Análise de componentes principais à confiabilidade de sistemas complexos**. 2001. Tese (Doutorado Engenharia de Produção) – Universidade Federal Santa Catarina, 2001.

LOPES, M. P. D. **Gerenciamento da qualidade no ensino da matemática**. 2004. Dissertação (Mestrado Engenharia de Produção) – Universidade Federal de Santa Maria, 2004.

MADDALA, G. S. **Introduction to econometrics**. 2. ed. New Jersey: Prentice-Hall, 1992.

MAGNUSSON, W. E.; MOURÃO, G. **Estatística sem matemática: a ligação entre as questões e a análise**. Curitiba, 2003.

MALHOTRA, N. K. **Pesquisa de marketing: uma orientação aplicada**. Porto Alegre: Bookman, 2001.

MANLY, B. F. J. **Multivariate statistical methods: a primer**. London: Chapman and Hall, 1986.

MARDIA, K.V.; KENT, J. T. i BIBBY, J. M. **Multivariate analysis**. London: Academic, 1979.

MORETTIN, P. M.; TOLOI, C. M. C. **Análise de séries temporais**. São Paulo: Atual, 1987.

MORETTIN, P. M.; TOLOI, C. M. C. **Análise de séries temporais**. São Paulo: Edgard Blücher, 2004.

NETO, M. M. J. Estatística multivariada. **Revista de Filosofia e Ensino**. 9 maio 2004. Disponível em: http://www.criticanarede.com/cien_estatistica.html. Acesso em: 9 maio 2004.

PLA, L. E. **Analysis multivariado: método de componentes principais**. Washington: Secretaria General de la Organización de Los Estados Americanos, 1986.

PEREIRA, J. C. R. **Análise de dados qualitativos**: estratégias metodológicas para as ciências da saúde, humanas e sociais. São Paulo: Edusp, 2001.

REGAZZI, A. J. **INF 766: análise multivariada**. Viçosa: Universidade Federal de Viçosa, Centro de Ciências Exatas e Tecnológicas, Departamento de Informática, 2001. 166p. Mimeografado.

REIS, E. **Estatística multivariada aplicada**. Lisboa, 1997.

SILVA, J. S. V.; MORAES, A. S. **A evolução da agropecuária no pantanal brasileiro**. Corumbá: EMBRAPA Pantanal, 2001.

SOUZA, A. M. **Monitoração e ajuste de realimentação em processos produtivos multivariados**. 2000. Tese (Doutorado Engenharia de Produção) – Universidade Federal Santa Catarina, 2000.

TOLEDO, G. L. & OVALLE, I. I. **Estatística básica**. 2. ed. São Paulo: Atlas, 1985.

TRIOLA, M. F. **Introdução à estatística** 7. ed. Rio de Janeiro, Livros Técnicos e Científicos, 1999.

VALENTIN, J. L. **Ecologia numérica**: uma introdução à análise multivariada de dados ecológicos. Rio de Janeiro: Interciência, 2000.

VERDINELLI, M. A. **Análise inercial em ecologia**. 1980. Tese (Doutorado) – Universidade de São Paulo, 1980.

WERKEMA, M.C.C. **As ferramentas da qualidade no gerenciamento de processos**. Belo Horizonte: Littera Maciel, 1995. (TQC: Gestão pela qualidade total).

7 ANEXOS

TABELA 24 – Total da produção do rebanho bovino brasileiro, por estado e por categorias, no período de 1997 a 2004

	RO	AC	AM	RR	PA	AP	TO
TOU/97	60389	7385	10264	6876	147595	1544	75988
TOU/98	64152	21023	12298	7368	99780	1176	79680
TOU/99	65802	21416	12764	7556	101765	1194	80870
TOU/00	68247	21180	13381	7825	102380	1234	80680
TOU/01	65247	21439	11628	7927	115152	861	93574
TOU/02	70930	20835	13634	7890	115635	914	89020
TOU/03	67297	21003	14384	8080	119087	1010	87133
TOU/04	90970	21232	15432	8279	124001	1116	86033
VAC/97	1117358	145181	193504	132788	2937224	26626	1619163
VAC/98	1526728	358784	259577	160698	2015569	21591	1726454
VAC/99	1608099	379144	282203	166137	2148187	23070	1883525
VAC/00	1664752	364474	303277	169856	2320712	24568	1983109
VAC/01	1688447	398672	327915	184906	2869647	25045	2292297
VAC/02	1924683	413637	413283	193359	3340622	27789	2381612
VAC/03	1903983	430740	368434	193171	3509423	30224	2363589
VAC/04	3087391	506762	398733	196228	3435324	32009	2484386
N 2a3/97	286884	36076	45813	31623	737527	7588	403616
N 2a3/98	372340	84819	61897	31901	536916	5900	427836
N 2a3/99	365022	86281	61942	32456	568472	6039	434954
N 2a3/00	382097	93385	69990	34604	611894	6919	435357
N 2a3/01	424175	87909	71434	35476	662526	6196	402054
N 2a3/02	417844	90504	99981	36329	712206	6705	429362
N 2a3/03	462124	96894	80433	37207	768615	7037	425392
N 2a3/04	748706	112576	88580	39084	796813	7658	541779
N1a2/97	389312	46277	63004	43026	1046262	9401	539378
N1a2/98	485133	106347	80956	41482	755030	6842	549341
N1a2/99	511367	110643	85438	43341	790150	7143	584285
N1a2/00	547265	116618	93098	47052	836091	7680	580096
N1a2/01	602579	116155	98509	47535	930460	7340	607141
N1a2/02	606731	118869	127392	48318	1004026	7603	594505
N1a2/03	643177	125482	105302	48890	1050481	8341	650870
N1a2/04	1078349	155578	115486	51036	1092913	9067	703410
Bas/97	427012	49444	68720	46292	1183565	9495	600611
Bas/98	557253	117566	88921	46698	837455	6834	625644
Bas/99	603516	128254	97628	49851	887245	7092	686062
Bas/00	628826	129507	104131	52321	959660	7573	702098
Bas/01	684023	134815	112094	54134	1083306	7365	739050
Bas/02	687034	132852	137417	54199	1169194	7791	769361
Bas/03	680865	140347	114312	53841	1219342	8519	755700
Bas/04	1282032	177023	124416	52913	1292091	8769	788186
Bos/97	417394	49922	68185	45902	1193660	9277	599633
Bos/98	544690	115722	88229	46314	840275	6617	627754
Bos/99	596496	126276	96515	49295	895949	6945	682422
Bos/00	624677	127512	102438	51226	994093	7343	694724
Bos/01	645121	135648	112397	54134	1098691	7365	724608
Bos/02	678411	134218	137273	54199	1188611	7545	752793
Bos/03	653615	142215	117231	53841	1236036	8263	738494
Bos/04	1250073	178713	127657	52913	1311067	8498	769512
G1a2/97	366565	45234	63965	42612	923982	8412	497630
G1a2/98	445089	96705	78441	40810	640594	5884	506994
G1a2/99	476133	102397	82669	43152	651519	6046	536091
G1a2/00	509558	107467	88525	45160	702531	6289	509398
G1a2/01	526248	114131	97117	47933	765923	6457	495232
G1a2/02	574900	117061	114106	49168	912690	6657	536828
G1a2/03	581736	123174	106205	48880	900020	7061	578239
G1a2/04	1016835	152607	119545	50266	922325	7684	587043

Continuação ...

	RO	AC	AM	RR	PA	AP	TO
G2a3/97	265641	36714	52598	35474	475878	6595	349201
G2a3/98	324911	77225	64507	33660	349556	4776	358306
G2a3/99	306429	74467	65628	34229	327756	4663	355437
G2a3/00	325113	72071	66481	35150	333982	4669	342569
G2a3/01	341764	81310	77542	37035	326770	5328	238326
G2a3/02	363678	82091	67624	36926	378934	4881	295904
G2a3/03	353467	86490	82030	36997	408726	4965	255561
G2a3/04	566538	104423	98505	37809	408181	5092	270425
B3a4/97	104142	19094	30314	20512	132687	3696	157576
B3a4/98	125925	40686	35611	19824	91457	2778	160824
B3a4/99	104886	36637	36122	19580	75083	2707	154400
B3a4/00	115547	32281	33730	18929	64390	2488	142168
B3a4/01	116170	33599	34812	19290	65645	3055	51604
B3a4/02	115765	34979	23217	20094	53450	2768	63782
B3a4/03	99103	33130	37929	18987	57532	2437	78982
B3a4/04	119462	33870	41657	18953	48117	2327	77964
B+4/97	30960	5623	20359	13473	34613	2368	65867
B+4/98	34815	9963	22119	12613	22271	1701	61205
B+4/99	29114	8409	21263	11625	12082	1582	54104
B+4/00	30806	6319	18754	9819	8017	1337	46984
B+4/01	34396	6046	17306	8937	4250	1786	3580
B+4/02	35555	5482	5485	8993	3916	1545	6896
B+4/03	26925	4712	12925	8064	2921	1287	7299
B+4/04	25421	3165	14548	7191	2120	1019	7634

Continuação ...

	MA	PI	CE	RN	PB	PE	AL
TOU/97	74082	28909	23892	13896	17024	21616	13978
TOU/98	67277	29173	38183	17219	22219	39154	14426
TOU/99	66906	28859	37164	16931	21611	38093	14073
TOU/00	66543	29357	42148	16991	22227	36738	14162
TOU/01	65151	26038	40025	14591	19900	34313	12472
TOU/02	64592	26975	39850	14762	19566	33235	14585
TOU/03	65728	27868	38822	15309	20443	31394	14556
TOU/04	67643	26935	39579	15564	20786	30312	14539
VAC/97	1269807	567905	462591	271863	354170	454101	267005
VAC/98	1361270	605846	763521	330568	458463	669918	305063
VAC/99	1400290	617446	764962	334580	468223	656837	308405
VAC/00	1416113	615828	899024	368025	549256	698774	344381
VAC/01	1502209	655788	981546	383502	573658	574857	353282
VAC/02	1609238	709536	975232	372561	556789	583288	336122
VAC/03	1689418	739286	955342	366937	591599	582894	347411
VAC/04	1544955	621345	958159	393762	599487	618166	357114
N 2a3/97	350380	128071	99975	59065	80030	92933	60737
N 2a3/98	347651	138483	163216	66018	91774	140687	72257
N 2a3/99	351277	145412	162710	70864	96927	135434	75477
N 2a3/00	349859	137256	189116	73091	104633	148255	76124
N 2a3/01	360584	136736	194681	72151	108310	155868	75060
N 2a3/02	348568	140739	184377	75684	101775	171719	67129
N 2a3/03	365734	148031	183421	78416	109401	174379	73827
N 2a3/04	350040	131903	204311	80620	112190	175256	75519
N1a2/97	443699	162194	138546	84852	100405	136064	84585
N1a2/98	445018	186266	230705	101705	133507	201174	104930
N1a2/99	443972	182539	221422	98592	133726	203193	100776
N1a2/00	449856	181029	248660	101855	142541	208928	108749
N1a2/01	442182	170724	246491	101246	138440	225697	100728
N1a2/02	460755	183937	251922	107398	139208	241407	98644
N1a2/03	466239	184642	236785	105117	139966	221763	101297
N1a2/04	434934	164769	255808	103774	146986	216578	99691
Bas/97	462029	189253	153341	93431	112126	153916	94142
Bas/98	476551	205774	257501	114090	154642	240126	117453
Bas/99	491995	199917	247931	110274	149266	237923	115465
Bas/00	476791	192509	276725	118978	167217	240834	125060
Bas/01	487707	194040	289157	119669	171808	237292	115274
Bas/02	485728	197227	278029	113568	161015	235776	107461
Bas/03	491487	196580	251597	108958	155846	216073	107380
Bas/04	457973	173949	278110	109257	165134	220951	109502
Bos/97	444896	180469	151842	90599	111032	153084	93236
Bos/98	461630	196226	256280	111430	153422	238367	116674
Bos/99	472010	191705	245261	107940	148429	236610	114698
Bos/00	457722	182941	272469	117088	168900	239500	124242
Bos/01	478214	190515	290583	119601	171742	242819	115104
Bos/02	472385	195595	279278	113298	160859	240086	107327
Bos/03	485050	194693	252479	108676	153996	220078	107574
Bos/04	449910	171512	279285	108981	160574	227730	110246
G1a2/97	410027	159917	139604	80462	100090	139443	84346
G1a2/98	373034	171684	230127	97167	131158	207373	104803
G1a2/99	385744	167422	224156	96350	134809	211108	101818
G1a2/00	399609	153000	240010	97358	139946	206309	105731
G1a2/01	394436	157863	232490	96564	134821	217042	94792
G1a2/02	405567	170393	234751	95201	134556	236659	91896
G1a2/03	412867	172642	223141	93573	130444	215296	95747
G1a2/04	400769	156682	243403	92335	130790	205506	95772

Continuação...

	MA	PI	CE	RN	PB	PE	AL
G2a3/97	331816	118460	111832	62833	78001	106235	62649
G2a3/98	328546	122121	173916	70508	99045	147858	75497
G2a3/99	298512	115942	173724	72121	96917	145629	75974
G2a3/00	280954	89732	179442	71795	95371	141141	70762
G2a3/01	311623	120516	169712	71830	98085	176048	73501
G2a3/02	287620	115220	156372	66354	91493	180827	65662
G2a3/03	293579	119444	153151	66605	89807	180325	69525
G2a3/04	299140	117519	181764	65970	90488	156249	70772
B3a4/97	178421	55722	61783	34669	40822	59342	31126
B3a4/98	183597	63072	103945	39491	60754	95668	39089
B3a4/99	180448	59935	100784	39437	54221	89789	38464
B3a4/00	138104	39749	107349	40687	51443	74483	38203
B3a4/01	168944	63499	79058	34719	46634	99253	36349
B3a4/02	157697	49606	69113	31457	43023	100861	31478
B3a4/03	150005	46786	65480	31037	40762	95579	30520
B3a4/04	152695	47735	67801	27564	37717	88509	32546
B+4/97	84893	13957	22708	19000	22426	29991	19031
B+4/98	78104	14190	36593	19892	31512	48388	20550
B+4/99	74747	14305	37269	19332	28171	48786	19722
B+4/00	57135	16041	42166	20625	23481	37861	18347
B+4/01	40699	15552	25951	16154	22492	58868	19205
B+4/02	28642	9316	21683	14879	20048	59842	15869
B+4/03	13798	4242	19733	13900	17991	56835	15368
B+4/04	12493	3862	18697	12310	16251	53030	13902

Continuação...

	SE	BA	MG	ES	RJ	SP	PR
TOU/97	15582	94349	274711	22063	19724	134723	103324
TOU/98	17176	137218	328059	27356	21085	153093	105158
TOU/99	16713	134830	313369	26964	20008	148120	101856
TOU/00	16606	142238	300950	26830	21708	144811	99226
TOU/01	14973	146205	303500	21452	18697	120746	90248
TOU/02	15329	151219	332608	21547	22613	127744	97826
TOU/03	15230	148296	329674	21900	22884	128927	95497
TOU/04	15145	152219	326933	21722	23237	131086	92584
VAC/97	299631	2308526	6147124	527364	462124	2981188	2233395
VAC/98	299920	2874728	6470993	537683	430369	2995834	2592929
VAC/99	297562	2961829	6856965	591587	500659	3167768	2861613
VAC/00	337449	3376489	6817740	606295	500905	3243911	2852978
VAC/01	374118	3663283	7933918	552066	611686	3890074	3078112
VAC/02	362369	3750649	7901889	566580	613449	3868520	3127317
VAC/03	357147	3706394	7990610	561718	628800	3790112	3078451
VAC/04	384033	3725382	7380696	504705	582206	3250676	2869208
N 2a3/97	63245	510282	1429070	133249	108218	606447	500292
N 2a3/98	66045	546265	1465653	139083	106332	641723	607564
N 2a3/99	70362	602081	1432153	140130	105445	602936	601241
N 2a3/00	75301	722308	1436390	139684	119312	618330	588769
N 2a3/01	74365	752269	1408513	126701	133061	846884	504007
N 2a3/02	71366	746488	1523127	128200	142538	859059	523640
N 2a3/03	74393	755637	1517508	124413	142702	808442	520253
N 2a3/04	77814	761361	1524845	117016	130774	752238	489365
N1a2/97	87016	719059	2168981	183341	198197	1386304	945626
N1a2/98	97272	831542	2155332	188629	181947	1363169	1071840
N1a2/99	97602	921962	2163591	196015	196570	1389254	1109561
N1a2/00	103680	1011695	2204544	191536	200573	1353196	1131682
N1a2/01	99863	979647	2441643	190503	236905	1629366	1180764
N1a2/02	102361	1054985	2420868	180488	228070	1582988	1173457
N1a2/03	98997	1010275	2465863	191339	231752	1606691	1174379
N1a2/04	99127	1030167	2307266	169181	225575	1480089	1131390
Bas/97	98999	806071	2420219	205024	216807	1552306	1061388
Bas/98	110720	994963	2469700	212196	201440	1538819	1212886
Bas/99	107383	1048426	2566146	228598	226411	1613066	1329081
Bas/00	115913	1126524	2597487	231006	233435	1635289	1366207
Bas/01	115839	1182297	2760322	214928	252719	1789997	1349737
Bas/02	110848	1184784	2767277	220385	247512	1788907	1329441
Bas/03	103569	1104310	2779394	216065	248839	1743621	1308480
Bas/04	110394	1208575	2700783	192419	261499	1686215	1294085
Bos/97	98024	820437	2381397	204194	227642	1679331	1088511
Bos/98	109943	1027990	2449385	210766	211023	1671935	1226605
Bos/99	106792	1053244	2548740	228002	242528	1846962	1365310
Bos/00	115272	1131710	2576034	229865	252699	1756691	1372632
Bos/01	115796	1183654	2752221	213794	251481	1875151	1361728
Bos/02	110744	1186256	2714756	220880	244733	1916244	1346224
Bos/03	103467	1107719	2728809	215013	246563	1866317	1326857
Bos/04	110295	1225477	2610551	198026	263638	1839412	1316403
G1a2/97	86212	792864	1897397	171363	197985	1411099	994328
G1a2/98	97441	928536	1930880	175420	180674	1368449	1061330
G1a2/99	97271	994562	1975585	187726	206738	1581904	1155159
G1a2/00	100954	1021850	1860798	177190	211459	1495436	1099785
G1a2/01	92618	1048413	1557939	201200	194726	1242099	1067764
G1a2/02	93953	1117884	1745277	185374	181727	1235147	1072580
G1a2/03	88800	1063333	1768930	191856	190613	1244216	1096667
G1a2/04	89142	1083516	1578913	166304	187978	1163455	1061300

Continuação...

	SE	BA	MG	ES	RJ	SP	PR
G2a3/97	66412	653616	1376614	123805	181153	1108875	775425
G2a3/98	69108	756486	1451405	125636	161573	1161969	824442
G2a3/99	72414	783003	1367143	126021	165713	1172181	806028
G2a3/00	69078	730692	1125367	117018	170501	1021202	721098
G2a3/01	69070	744794	581470	135804	77246	509585	671365
G2a3/02	64548	753791	643354	131823	87506	477194	635224
G2a3/03	63155	728974	613290	119989	90397	448022	616140
G2a3/04	62026	737333	609765	104850	87596	436726	596288
B3a4/97	35213	331490	391255	42598	106932	706253	399192
B3a4/98	42006	443045	480235	42584	98639	728639	412527
B3a4/99	39876	402963	490253	41773	99955	777611	367024
B3a4/00	35446	329485	242471	34426	84363	503593	282205
B3a4/01	32892	321243	84993	53326	15858	81980	255361
B3a4/02	30221	331623	106570	44789	11842	69937	230039
B3a4/03	29578	312142	102221	40483	15832	79734	198785
B3a4/04	29473	314195	100777	31606	14323	71993	171447
B+4/97	19567	130572	73570	21185	32034	411429	156100
B+4/98	22948	180152	67242	21309	28257	415788	156270
B+4/99	23083	183267	64133	15859	27097	399920	115828
B+4/00	14324	121621	42205	11853	18151	197643	82039
B+4/01	14981	124388	3187	6172	935	4601	56565
B+4/02	13302	119875	12461	5992	1670	5571	48497
B+4/03	12436	97138	10828	4557	1691	5143	33467
B+4/04	11715	86398	11255	3604	1660	5072	19140

Continuação...

	SC	RS	MS	MT	GO	DF
TOU/97	37505	167762	327407	224112	240904	1254
TOU/98	36158	178215	310425	249612	223742	1544
TOU/99	34280	174053	302951	248469	217122	1503
TOU/00	32717	173223	295970	252356	218411	1475
TOU/01	31253	154140	289765	262717	188622	1084
TOU/02	30361	159800	305917	264350	206780	1157
TOU/03	30665	157606	301869	267593	205296	1181
TOU/04	30484	156562	288828	317496	199322	1202
VAC/97	761895	3687215	6446692	5127384	5365414	33024
VAC/98	833569	3979704	6254741	5325175	4725386	23782
VAC/99	826640	4226012	6871709	5588011	5382279	24605
VAC/00	878301	4072467	7269233	5734442	5489720	24996
VAC/01	975141	4672845	8083818	6629369	6160451	23569
VAC/02	997243	4578241	8561263	7016599	6095351	24308
VAC/03	928773	4514861	7603857	7684325	5885616	22974
VAC/04	963890	4345507	7703867	7348485	5670691	18891
N 2a3/97	147134	754727	1528617	1116027	1305137	7372
N 2a3/98	170053	890942	1526404	1313090	1214192	5349
N 2a3/99	153270	868737	1522924	1273423	1205320	5210
N 2a3/00	174518	854503	1711182	1373703	1266106	5803
N 2a3/01	195858	965639	1784454	1454389	1253642	5304
N 2a3/02	200311	938003	1751966	1456231	1271898	5075
N 2a3/03	200194	958822	1672764	1544135	1214733	4858
N 2a3/04	202126	893496	1859851	1799182	1300090	5387
N1a2/97	330978	1531982	2476589	1676656	1989991	14146
N1a2/98	336847	1494080	2222546	1756009	1773446	9691
N1a2/99	334509	1548204	2353769	1771075	1882892	9990
N1a2/00	338541	1553108	2417674	1867281	1913893	10201
N1a2/01	364215	1606065	2539853	2033071	2022097	10181
N1a2/02	368735	1557244	2562228	1997546	1920430	9749
N1a2/03	372269	1555143	2470653	2341179	2023139	10372
N1a2/04	363677	1560654	2587760	2525544	2090440	10025
Bas/97	362898	1667029	2668804	1907882	2181768	15746
Bas/98	386618	1694063	2497709	2058544	1982366	11304
Bas/99	402081	1844075	2710032	2210062	2225722	11797
Bas/00	406765	1819251	2921538	2206197	2262275	11947
Bas/01	422477	1853334	2971213	2360301	2298603	10793
Bas/02	415067	1756577	2975879	2558078	2391863	11618
Bas/03	410482	1712784	2744474	2715415	2306791	11111
Bas/04	404447	1756069	2934311	3000531	2403552	10045
Bos/97	351279	1591950	2524590	1902785	2195820	15659
Bos/98	375059	1654270	2371620	2049186	1999978	11255
Bos/99	394814	1819297	2546670	2216645	2256340	11791
Bos/00	399669	1794282	2917738	2178483	2261059	11940
Bos/01	434153	1829733	2882835	2347893	2270882	11491
Bos/02	421792	1732118	2828828	2558186	2350628	12578
Bos/03	417580	1690177	2613352	2703884	2267064	11966
Bos/04	412727	1724278	2807167	2986729	2369564	10869
G1a2/97	299577	1234089	1982890	1301920	1819927	12959
G1a2/98	300968	1277282	1785445	1326644	1626128	9180
G1a2/99	314933	1371258	1866401	1403204	1769588	9513
G1a2/00	309903	1392083	2079097	1404304	1655955	9188
G1a2/01	319586	1315708	2165640	1578273	1532924	12397
G1a2/02	327145	1309350	1990024	1613496	1480524	11639
G1a2/03	326391	1291208	1896831	1914290	1655015	13140
G1a2/04	313096	1284242	1889211	1890503	1601050	12637

Continuação ...

	SC	RS	MS	MT	GO	DF
G2a3/97	244241	838216	1288463	676399	1370886	9300
G2a3/98	250648	882020	1220324	654969	1224768	6754
G2a3/99	239831	877913	1114034	657170	1172708	6491
G2a3/00	224903	904359	1268956	789066	1061890	6345
G2a3/01	200412	765133	1018755	699840	620554	8866
G2a3/02	201758	762328	930577	822611	695734	8977
G2a3/03	198739	751828	877348	866292	809635	10313
G2a3/04	189141	751523	793466	929161	718951	11322
B3a4/97	163870	465544	632769	183657	435619	4353
B3a4/98	170418	477972	605175	161538	409991	3312
B3a4/99	162666	466974	568917	146366	373493	3035
B3a4/00	133208	489709	550292	266844	272158	2835
B3a4/01	96232	318237	266428	235235	145849	4718
B3a4/02	92508	313840	229329	227470	138571	3680
B3a4/03	87062	296646	247169	273706	176174	5072
B3a4/04	77742	296045	199842	246657	139781	4681
B+4/97	171350	373454	245479	36145	106171	1764
B+4/98	171127	337484	235928	30629	92271	1339
B+4/99	159199	285529	175460	25211	70686	1145
B+4/00	112585	330309	209019	87698	35518	927
B+4/01	35661	81785	18366	22398	15273	304
B+4/02	30895	78411	28454	25367	39110	445
B+4/03	23454	68428	22168	26356	26310	388
B+4/04	16366	59888	24678	23894	24159	379

Tabela 25 - Número de bois com mais de 4 anos de idade, nos estados brasileiros, no período de 1997 a 2004

	B+4/97	B+4/98	B+4/99	B+4/00	B+4/01	B+4/02	B+4/03	B+4/04	B+4/97
RO	30.960	34.815	29.114	30.806	34.396	35.555	26.925	25.421	30.960
AC	5.623	9.963	8.409	6.319	6.046	5.482	4.712	3.165	5.623
AM	20.359	22.119	21.263	18.754	17.306	5.485	12.925	14.548	20.359
RR	13.473	12.613	11.625	9.819	8.937	8.993	8.064	7.191	13.473
PA	34.613	22.271	12.082	8.017	4.250	3.916	2.921	2.120	34.613
AP	2.368	1.701	1.582	1.337	1.786	1.545	1.287	1.019	2.368
TO	65.867	61.205	54.104	46.984	3.580	6.896	7.299	7.634	65.867
MA	84.893	78.104	74.747	57.135	40.699	28.642	13.798	12.493	84.893
PI	13.957	14.190	14.305	16.041	15.552	9.316	4.242	3.862	13.957
CE	22.708	36.593	37.269	42.166	25.951	21.683	19.733	18.697	22.708
RN	19.000	19.892	19.332	20.625	16.154	14.879	13.900	12.310	19.000
PB	22.426	31.512	28.171	23.481	22.492	20.048	17.991	16.251	22.426
PE	29.991	48.388	48.786	37.861	58.868	59.842	56.835	53.030	29.991
AL	19.031	20.550	19.722	18.347	19.205	15.869	15.368	13.902	19.031
SE	19.567	22.948	23.083	14.324	14.981	13.302	12.436	11.715	19.567
BA	130.572	180.152	183.267	121.621	124.388	119.875	97.138	86.398	130.572
MG	73.570	67.242	64.133	42.205	3.187	12.461	10.828	11.255	73.570
ES	21.185	21.309	15.859	11.853	6.172	5.992	4.557	3.604	21.185
RJ	32.034	28.257	27.097	18.151	935	1.670	1.691	1.660	32.034
SP	411.429	415.788	399.920	197.643	4.601	5.571	5.143	5.072	411.429
PR	156.100	156.270	115.828	82.039	56.565	48.497	33.467	19.140	156.100
SC	171.350	171.127	159.199	112.585	35.661	30.895	23.454	16.366	171.350
RS	373.454	337.484	285.529	330.309	81.785	78.411	68.428	59.888	373.454
MS	245.479	235.928	175.460	209.019	18.366	28.454	22.168	24.678	245.479
MT	36.145	30.629	25.211	87.698	22.398	25.367	26.356	23.894	36.145
GO	106.171	92.271	70.686	35.518	15.273	39.110	26.310	24.159	106.171
DF	1.764	1.339	1.145	927	304	445	388	379	1.764

TABELA 26 - Número de vacas, nos estados brasileiros, no período de 1997 a 2004

	VAC/97	VAC/98	VAC/99	VAC/00	VAC/01	VAC/02	VAC/03	VAC/04
RO	1.117.358	1.526.728	1.608.099	1.664.752	1.688.447	1.924.683	1.903.983	3.087.391
AC	145.181	358.784	379.144	364.474	398.672	413.637	430.740	506.762
AM	193.504	259.577	282.203	303.277	327.915	413.283	368.434	398.733
RR	132.788	160.698	166.137	169.856	184.906	193.359	193.171	196.228
PA	2.937.224	2.015.569	2.148.187	2.320.712	2.869.647	3.340.622	3.509.423	3.435.324
AP	26.626	21.591	23.070	24.568	25.045	27.789	30.224	32.009
TO	1.619.163	1.726.454	1.883.525	1.983.109	2.292.297	2.381.612	2.363.589	2.484.386
MA	1.269.807	1.361.270	1.400.290	1.416.113	1.502.209	1.609.238	1.689.418	1.544.955
PI	567.905	605.846	617.446	615.828	655.788	709.536	739.286	621.345
CE	462.591	763.521	764.962	899.024	981.546	975.232	955.342	958.159
RN	271.863	330.568	334.580	368.025	383.502	372.561	366.937	393.762
PB	354.170	458.463	468.223	549.256	573.658	556.789	591.599	599.487
PE	454.101	669.918	656.837	698.774	574.857	583.288	582.894	618.166
AL	267.005	305.063	308.405	344.381	353.282	336.122	347.411	357.114
SE	299.631	299.920	297.562	337.449	374.118	362.369	357.147	384.033
BA	2.308.526	2.874.728	2.961.829	3.376.489	3.663.283	3.750.649	3.706.394	3.725.382
MG	6.147.124	6.470.993	6.856.965	6.817.740	7.933.918	7.901.889	7.990.610	7.380.696
ES	527.364	537.683	591.587	606.295	552.066	566.580	561.718	504.705
RJ	462.124	430.369	500.659	500.905	611.686	613.449	628.800	582.206
SP	2.981.188	2.995.834	3.167.768	3.243.911	3.890.074	3.868.520	3.790.112	3.250.676
PR	2.233.395	2.592.929	2.861.613	2.852.978	3.078.112	3.127.317	3.078.451	2.869.208
SC	761.895	833.569	826.640	878.301	975.141	997.243	928.773	963.890
RS	3.687.215	3.979.704	4.226.012	4.072.467	4.672.845	4.578.241	4.514.861	4.345.507
MS	6.446.692	6.254.741	6.871.709	7.269.233	8.083.818	8.561.263	7.603.857	7.703.867
MT	5.127.384	5.325.175	5.588.011	5.734.442	6.629.369	7.016.599	7.684.325	7.348.485
GO	5.365.414	4.725.386	5.382.279	5.489.720	6.160.451	6.095.351	5.885.616	5.670.691
DF	33.024	23.782	24.605	24.996	23.569	24.308	22.974	18.891

TABELA 27 - Número de bezerros, nos estados brasileiros, no período de 1997 a 2004

	Bos/97	Bos/98	Bos/99	Bos/00	Bos/01	Bos/02	Bos/03	Bos/04
RO	417.394	544.690	596.496	624.677	645.121	678.411	653.615	1.250.073
AC	49.922	115.722	126.276	127.512	135.648	134.218	142.215	178.713
AM	68.185	88.229	96.515	102.438	112.397	137.273	117.231	127.657
RR	45.902	46.314	49.295	51.226	54.134	54.199	53.841	52.913
PA	1.193.660	840.275	895.949	994.093	1.098.691	1.188.611	1.236.036	1.311.067
AP	9.277	6.617	6.945	7.343	7.365	7.545	8.263	8.498
TO	599.633	627.754	682.422	694.724	724.608	752.793	738.494	769.512
MA	444.896	461.630	472.010	457.722	478.214	472.385	485.050	449.910
PI	180.469	196.226	191.705	182.941	190.515	195.595	194.693	171.512
CE	151.842	256.280	245.261	272.469	290.583	279.278	252.479	279.285
RN	90.599	111.430	107.940	117.088	119.601	113.298	108.676	108.981
PB	111.032	153.422	148.429	168.900	171.742	160.859	153.996	160.574
PE	153.084	238.367	236.610	239.500	242.819	240.086	220.078	227.730
AL	93.236	116.674	114.698	124.242	115.104	107.327	107.574	110.246
SE	98.024	109.943	106.792	115.272	115.796	110.744	103.467	110.295
BA	820.437	1.027.990	1.053.244	1.131.710	1.183.654	1.186.256	1.107.719	1.225.477
MG	2.381.397	2.449.385	2.548.740	2.576.034	2.752.221	2.714.756	2.728.809	2.610.551
ES	204.194	210.766	228.002	229.865	213.794	220.880	215.013	198.026
RJ	227.642	211.023	242.528	252.699	251.481	244.733	246.563	263.638
SP	1.679.331	1.671.935	1.846.962	1.756.691	1.875.151	1.916.244	1.866.317	1.839.412
PR	1.088.511	1.226.605	1.365.310	1.372.632	1.361.728	1.346.224	1.326.857	1.316.403
SC	351.279	375.059	394.814	399.669	434.153	421.792	417.580	412.727
RS	1.591.950	1.654.270	1.819.297	1.794.282	1.829.733	1.732.118	1.690.177	1.724.278
MS	2.524.590	2.371.620	2.546.670	2.917.738	2.882.835	2.828.828	2.613.352	2.807.167
MT	1.902.785	2.049.186	2.216.645	2.178.483	2.347.893	2.558.186	2.703.884	2.986.729
GO	2.195.820	1.999.978	2.256.340	2.261.059	2.270.882	2.350.628	2.267.064	2.369.564
DF	15.659	11.255	11.791	11.940	11.491	12.578	11.966	10.869

Tabela 28 - Número de bezerras, nos estados brasileiros, no período de 1997 a 2004

	Bas/97	Bas/98	Bas/99	Bas/00	Bas/01	Bas/02	Bas/03	Bas/04
RO	427.012	557.253	603.516	628.826	684.023	687.034	680.865	1.282.032
AC	49.444	117.566	128.254	129.507	134.815	132.852	140.347	177.023
AM	68.720	88.921	97.628	104.131	112.094	137.417	114.312	124.416
RR	46.292	46.698	49.851	52.321	54.134	54.199	53.841	52.913
PA	1.183.565	837.455	887.245	959.660	1.083.306	1.169.194	1.219.342	1.292.091
AP	9.495	6.834	7.092	7.573	7.365	7.791	8.519	8.769
TO	600.611	625.644	686.062	702.098	739.050	769.361	755.700	788.186
MA	462.029	476.551	491.995	476.791	487.707	485.728	491.487	457.973
PI	189.253	205.774	199.917	192.509	194.040	197.227	196.580	173.949
CE	153.341	257.501	247.931	276.725	289.157	278.029	251.597	278.110
RN	93.431	114.090	110.274	118.978	119.669	113.568	108.958	109.257
PB	112.126	154.642	149.266	167.217	171.808	161.015	155.846	165.134
PE	153.916	240.126	237.923	240.834	237.292	235.776	216.073	220.951
AL	94.142	117.453	115.465	125.060	115.274	107.461	107.380	109.502
SE	98.999	110.720	107.383	115.913	115.839	110.848	103.569	110.394
BA	806.071	994.963	1.048.426	1.126.524	1.182.297	1.184.784	1.104.310	1.208.575
MG	2.420.219	2.469.700	2.566.146	2.597.487	2.760.322	2.767.277	2.779.394	2.700.783
ES	205.024	212.196	228.598	231.006	214.928	220.385	216.065	192.419
RJ	216.807	201.440	226.411	233.435	252.719	247.512	248.839	261.499
SP	1.552.306	1.538.819	1.613.066	1.635.289	1.789.997	1.788.907	1.743.621	1.686.215
PR	1.061.388	1.212.886	1.329.081	1.366.207	1.349.737	1.329.441	1.308.480	1.294.085
SC	362.898	386.618	402.081	406.765	422.477	415.067	410.482	404.447
RS	1.667.029	1.694.063	1.844.075	1.819.251	1.853.334	1.756.577	1.712.784	1.756.069
MS	2.668.804	2.497.709	2.710.032	2.921.538	2.971.213	2.975.879	2.744.474	2.934.311
MT	1.907.882	2.058.544	2.210.062	2.206.197	2.360.301	2.558.078	2.715.415	3.000.531
GO	2.181.768	1.982.366	2.225.722	2.262.275	2.298.603	2.391.863	2.306.791	2.403.552
DF	15.746	11.304	11.797	11.947	10.793	11.618	11.111	10.045

TABELA 29 - Número de novilhos de 2 a 3 anos, nos estados brasileiros, no período de 1997 a 2004

	N 2ª/97	N 2a/98	N 2a3/99	N 2a3/00	N 2a3/01	N 2a3/02	N 2a3/03	N 2a3/04
RO	286.884	372.340	365.022	382.097	424.175	417.844	462.124	748.706
AC	36.076	84.819	86.281	93.385	87.909	90.504	96.894	112.576
AM	45.813	61.897	61.942	69.990	71.434	99.981	80.433	88.580
RR	31.623	31.901	32.456	34.604	35.476	36.329	37.207	39.084
PA	737.527	536.916	568.472	611.894	662.526	712.206	768.615	796.813
AP	7.588	5.900	6.039	6.919	6.196	6.705	7.037	7.658
TO	403.616	427.836	434.954	435.357	402.054	429.362	425.392	541.779
MA	350.380	347.651	351.277	349.859	360.584	348.568	365.734	350.040
PI	128.071	138.483	145.412	137.256	136.736	140.739	148.031	131.903
CE	99.975	163.216	162.710	189.116	194.681	184.377	183.421	204.311
RN	59.065	66.018	70.864	73.091	72.151	75.684	78.416	80.620
PB	80.030	91.774	96.927	104.633	108.310	101.775	109.401	112.190
PE	92.933	140.687	135.434	148.255	155.868	171.719	174.379	175.256
AL	60.737	72.257	75.477	76.124	75.060	67.129	73.827	75.519
SE	63.245	66.045	70.362	75.301	74.365	71.366	74.393	77.814
BA	510.282	546.265	602.081	722.308	752.269	746.488	755.637	761.361
MG	1.429.070	1.465.653	1.432.153	1.436.390	1.408.513	1.523.127	1.517.508	1.524.845
ES	133.249	139.083	140.130	139.684	126.701	128.200	124.413	117.016
RJ	108.218	106.332	105.445	119.312	133.061	142.538	142.702	130.774
SP	606.447	641.723	602.936	618.330	846.884	859.059	808.442	752.238
PR	500.292	607.564	601.241	588.769	504.007	523.640	520.253	489.365
SC	147.134	170.053	153.270	174.518	195.858	200.311	200.194	202.126
RS	754.727	890.942	868.737	854.503	965.639	938.003	958.822	893.496
MS	1.528.617	1.526.404	1.522.924	1.711.182	1.784.454	1.751.966	1.672.764	1.859.851
MT	1.116.027	1.313.090	1.273.423	1.373.703	1.454.389	1.456.231	1.544.135	1.799.182
GO	1.305.137	1.214.192	1.205.320	1.266.106	1.253.642	1.271.898	1.214.733	1.300.090
DF	7.372	5.349	5.210	5.803	5.304	5.075	4.858	5.387

TABELA 30 - Número de garrotes de 1 a 2 anos, nos estados brasileiros, no período de 1997 a 2004

	G1a2/97	G1a2/98	G1a2/99	G1a2/00	G1a2/01	G1a2/02	G1a2/03	G1a2/04
RO	366.565	445.089	476.133	509.558	526.248	574.900	581.736	1.016.835
AC	45.234	96.705	102.397	107.467	114.131	117.061	123.174	152.607
AM	63.965	78.441	82.669	88.525	97.117	114.106	106.205	119.545
RR	42.612	40.810	43.152	45.160	47.933	49.168	48.880	50.266
PA	923.982	640.594	651.519	702.531	765.923	912.690	900.020	922.325
AP	8.412	5.884	6.046	6.289	6.457	6.657	7.061	7.684
TO	497.630	506.994	536.091	509.398	495.232	536.828	578.239	587.043
MA	410.027	373.034	385.744	399.609	394.436	405.567	412.867	400.769
PI	159.917	171.684	167.422	153.000	157.863	170.393	172.642	156.682
CE	139.604	230.127	224.156	240.010	232.490	234.751	223.141	243.403
RN	80.462	97.167	96.350	97.358	96.564	95.201	93.573	92.335
PB	100.090	131.158	134.809	139.946	134.821	134.556	130.444	130.790
PE	139.443	207.373	211.108	206.309	217.042	236.659	215.296	205.506
AL	84.346	104.803	101.818	105.731	94.792	91.896	95.747	95.772
SE	86.212	97.441	97.271	100.954	92.618	93.953	88.800	89.142
BA	792.864	928.536	994.562	1.021.850	1.048.413	1.117.884	1.063.333	1.083.516
MG	1.897.397	1.930.880	1.975.585	1.860.798	1.557.939	1.745.277	1.768.930	1.578.913
ES	171.363	175.420	187.726	177.190	201.200	185.374	191.856	166.304
RJ	197.985	180.674	206.738	211.459	194.726	181.727	190.613	187.978
SP	1.411.099	1.368.449	1.581.904	1.495.436	1.242.099	1.235.147	1.244.216	1.163.455
PR	994.328	1.061.330	1.155.159	1.099.785	1.067.764	1.072.580	1.096.667	1.061.300
SC	299.577	300.968	314.933	309.903	319.586	327.145	326.391	313.096
RS	1.234.089	1.277.282	1.371.258	1.392.083	1.315.708	1.309.350	1.291.208	1.284.242
MS	1.982.890	1.785.445	1.866.401	2.079.097	2.165.640	1.990.024	1.896.831	1.889.211
MT	1.301.920	1.326.644	1.403.204	1.404.304	1.578.273	1.613.496	1.914.290	1.890.503
GO	1.819.927	1.626.128	1.769.588	1.655.955	1.532.924	1.480.524	1.655.015	1.601.050
DF	12.959	9.180	9.513	9.188	12.397	11.639	13.140	12.637

TABELA 31 - Número de garrotes de 2 a 3 anos, nos estados brasileiros, no período de 1997 a 2004

	G2a3/97	G2a3/98	G2a3/99	G2a3/00	G2a3/01	G2a3/02	G2a3/03	G2a3/04
RO	265.641	324.911	306.429	325.113	341.764	363.678	353.467	566.538
AC	36.714	77.225	74.467	72.071	81.310	82.091	86.490	104.423
AM	52.598	64.507	65.628	66.481	77.542	67.624	82.030	98.505
RR	35.474	33.660	34.229	35.150	37.035	36.926	36.997	37.809
PA	475.878	349.556	327.756	333.982	326.770	378.934	408.726	408.181
AP	6.595	4.776	4.663	4.669	5.328	4.881	4.965	5.092
TO	349.201	358.306	355.437	342.569	238.326	295.904	255.561	270.425
MA	331.816	328.546	298.512	280.954	311.623	287.620	293.579	299.140
PI	118.460	122.121	115.942	89.732	120.516	115.220	119.444	117.519
CE	111.832	173.916	173.724	179.442	169.712	156.372	153.151	181.764
RN	62.833	70.508	72.121	71.795	71.830	66.354	66.605	65.970
PB	78.001	99.045	96.917	95.371	98.085	91.493	89.807	90.488
PE	106.235	147.858	145.629	141.141	176.048	180.827	180.325	156.249
AL	62.649	75.497	75.974	70.762	73.501	65.662	69.525	70.772
SE	66.412	69.108	72.414	69.078	69.070	64.548	63.155	62.026
BA	653.616	756.486	783.003	730.692	744.794	753.791	728.974	737.333
MG	1.376.614	1.451.405	1.367.143	1.125.367	581.470	643.354	613.290	609.765
ES	123.805	125.636	126.021	117.018	135.804	131.823	119.989	104.850
RJ	181.153	161.573	165.713	170.501	77.246	87.506	90.397	87.596
SP	1.108.875	1.161.969	1.172.181	1.021.202	509.585	477.194	448.022	436.726
PR	775.425	824.442	806.028	721.098	671.365	635.224	616.140	596.288
SC	244.241	250.648	239.831	224.903	200.412	201.758	198.739	189.141
RS	838.216	882.020	877.913	904.359	765.133	762.328	751.828	751.523
MS	1.288.463	1.220.324	1.114.034	1.268.956	1.018.755	930.577	877.348	793.466
MT	676.399	654.969	657.170	789.066	699.840	822.611	866.292	929.161
GO	1.370.886	1.224.768	1.172.708	1.061.890	620.554	695.734	809.635	718.951
DF	9.300	6.754	6.491	6.345	8.866	8.977	10.313	11.322

