

# PREVISÃO DE SAFRAS DE CACAU COM BASE EM MÉTODO DE CONTAGEM DE FRUTOS<sup>1</sup>

MARGARIDA MARIA HOEPPNER ZARONI<sup>2</sup> e TIRZA AIDARI<sup>3</sup>

**RESUMO** - Desenvolveu-se um método de previsão de safras de cacau baseado na relação produção de cacau e observação do fruto. Estudos do perfil da frutificação durante um ano agrícola foram usados para criar o ponto preditivo, quando os perfis da frutificação são classificados em grupos homogêneos, e para auxiliar na verificação da relevância do modelo preditivo, no caso de a predição ser uma interpolação ou uma extrapolação. Previsões com baixos erros de estimação foram obtidas com três a seis meses de antecedência ao final de cada safra.

**Termos para indexação:** *Theobroma cacao*, construção de modelo, interpolação e extrapolação, análise de agrupamentos.

## COCOA CROP FORECAST BASED ON POD COUNTING METHOD

**ABSTRACT** - The development of the method is based on the relationship between cocoa crop production and pod observation. A study of the pod profiles during a crop year is used to create the predictive point when the pod profiles are classified in homogeneous clusters and to verify the predictive model relevance when the prediction is an interpolation or an extrapolation. Predictions with lower estimation errors are obtained three to six months before each crop end.

**Index terms:** *Theobroma cacao*, model building, interpolation, extrapolation, cluster analysis.

## INTRODUÇÃO

No presente trabalho discutir-se-ão os primeiros resultados da previsão de safra de cacau baseada em ensaios para o estabelecimento de um método cuja fonte de informação se origina na contagem de frutos, pela região cacauera do estado da Bahia.

Embora este procedimento esteja amplamente difundido nos países produtores de cacau, encontra-se pouca bibliografia que descreva seu comportamento teórico.

Encontram-se citações sobre a necessidade da experiência de anos para a formulação e ensaio de métodos e sobre a importância do conhecimento das características do cacauero e do hábito de crescimento dos frutos para prever os componentes da produção.

Existem diferentes métodos de previsão de safra de cacau, dentre os quais o de estimativas de safras por amostragem, em que a produção da propriedade cacauera é tomada como unidade amostral, e os que levam em consideração os diversos fatores climáticos que influenciam a produção.

Na prática, os resultados obtidos mostraram-se mais funcionais em relação aos resultados das safras dos anos agrícolas 1981/82 a 1985/86, comparando-os aos de dois outros métodos.

Apesar do custo elevado, a tomada periódica de dados reais sobre a frutificação em pontos dispersos pela região cacauera, além de possibilitar a estimativa da produção das safras com baixo erro de estimação, possibilita a descrição do comportamento médio da frutificação da região em estudo, durante o ano agrícola.

<sup>1</sup> Aceito para publicação em 7 de novembro de 1991.

Apresentado no 2º Simpósio de Estatística Aplicada à Experimentação Agronômica e na 32ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria, Londrina, julho de 1987.

<sup>2</sup> M.Sc. em Estatística, EMBRAPA-Centro Nacional de Pesquisa de Defesa da Agricultura (CNPDA), Km 127, 5, Rodovia Campinas/Mogi-Mirim, Caixa Postal 69, CEP 13820 Jaguariúna, SP.

<sup>3</sup> Bacharel em Estatística, Assessoria de Informática e Estatística, Centro de Pesquisas do Cacau, Caixa Postal 7, CEP 45600 Itabuna, BA.

Formaliza-se o método partindo-se de informações de contagem de frutos em cinco diferentes estágios de desenvolvimento, em períodos igualmente espaçados durante um ano agrícola.

Waters & Hunter (1928) revelaram que medições reais de frutos jovens de cacau em uma fazenda fornecem bases para estimar a produção de cacau em amêndoas secas que dele se espera. Comentam, também, as dificuldades de obtenção de estimativas seguras por esse método, devido à incidência de doenças e ao número necessário para fornecer uma libra de cacau em amêndoas secas. E ainda, sugerem que, se isto fosse feito em um número limitado de fazendas dispersas na Costa do Marfim, África, ou em áreas de cacau representativas, esses dados deveriam fornecer a base para uma estimativa bastante exata da colheita.

Rau (1965) cita que um indicador facilmente mensurável da produção final de uma safra de cacau é o desenvolvimento do fruto na árvore, pois o fruto de cacau é a resposta de todos os fatores que influenciam a produção, com a inclusão de fatores meteorológicos.

Alvim et al. (1965), estudando o ciclo entre a polinização e a maturação, em cacau da variedade Catongo, em Uruçuca, BA, verificou que, entre os frutos colhidos em março-maio (safra do temporão) e outubro-novembro (safra principal), os primeiros têm ciclo médio entre a polinização e a colheita, de aproximadamente 155 dias, ao passo que os últimos, de cerca de 184 dias.

No Informe Técnico da FAO (FAO 1967), vê-se como objetivo a previsão de safra provável, com quatro meses de antecedência ao resultado final da colheita. Toma-se como critério geral a aplicação da fórmula:

$$Y_f = \frac{A_f \cdot N_b}{2240 \cdot R_f}$$

onde,

$Y_f$  = previsão da produção de cacau, em libras, correspondente a uma safra específica;

$N_b$  = número total de cacauzeiros em idade de produzir;

$A_f$  = número médio de frutos que amadurecerão por árvore;

$R_f$  = número de frutos necessários para obter uma libra de cacau seco; e

2240 = constante não definida no texto.

A técnica de amostragem aplicada foi casual estratificada em várias fases. A amostragem casual estratificada em uma única fase estimou o número total de árvores em idade de produzir. Na amostragem casual estratificada em duas fases, tendo como objetivo estimar o número total de frutos fermentáveis, tomaram-se como unidade amostral, na primeira fase, as zonas de cacauais, e, na segunda fase, as árvores em idade de produzir, divididas em duas pequenas parcelas de 18 cacauzeiros. No Informe Técnico da FAO (FAO 1968) menciona-se ser a medida em frutos de cacau uma fonte potencial de informação, com o fim de estabelecer a previsão da produção. Embora se necessite da experiência de anos para a formulação e o ensaio de métodos, é importante conhecer as características do cacauzeiro e o hábito do crescimento dos frutos, pois, assim, prever-se-ão, aproximadamente, os componentes da produção.

Ripailles & Rossion (1968) indicaram que a estimativa da produção feita três meses antes da colheita baseia-se na contagem de frutos em dois tamanhos diferentes.

No Brasil, Pereira (1972) iniciou os trabalhos de amostragem visando a fazer previsões da safra principal e do temporão, de cacau no estado da Bahia, utilizando o método de entrevista a fazendeiros. Posteriormente, Arroyo Vergara et al. (1980a) dimensionaram uma nova amostra de empresas cacauzeiras no estado a Bahia, com a mesma finalidade do trabalho anterior, tomando como base um cadastro de empresas cacauzeiras elaborado por Arroyo Vergara et al. (1980b), com o Departamento de Extensão da CEPLAC (DEPEX).

Zaroni (1981) estudou a relação entre frutos do primeiro e do quarto estágio de frutificação do cacauzeiro, para previsão de safras. Neste trabalho, previu com boa precisão o número médio de frutos adultos por safra, quando o seu ciclo de bilração se completa. Conseqüentemen-

te, sugeriu a previsão de uma dada safra em duas épocas distintas. A primeira, em função do número estimado de frutos adultos e a segunda em função do número observado de frutos adultos na safra em estudo.

## MATERIAL E MÉTODOS

Para a seleção de amostras, a região cacauera do estado da Bahia foi dividida em duas sub-regiões: uma, ao norte, e outra, ao sul, tendo como centro a cidade de Itabuna. Tomando-se como eixo referencial a direção da rodovia BR-101 que corta a região de norte a sul, foram escolhidas sistematicamente 139 fazendas (conforme mostra o mapa na Fig. 1) e, dentro de cada uma, uma parcela com cinco cacauzeiros, para coleta de dados. A partir de junho de 1976, iniciou-se a contagem de frutos nas árvores mencionadas, em períodos aproximadamente regulares, anotando-se as seguintes informações:

1. número de frutos maduros (quinto estágio de desenvolvimento do fruto do cacauzeiro);
2. número de frutos adultos (quarto estágio de desenvolvimento do fruto do cacauzeiro);
3. número de frutos médios (terceiro estágio de desenvolvimento do fruto do cacauzeiro);
4. número de frutos pequenos (segundo estágio de desenvolvimento do fruto. Frutos entre 3 e 7 cm de comprimento);
5. número de frutos bilros (primeiro estágio de desenvolvimento do fruto. Frutos até 3 cm de comprimento);
6. número de frutos com podridão-parda;
7. número de frutos com outros estragos.

Ainda na área que contém a parcela foram tomadas outras informações, como:

- se houve colheita na área ou não;
- quanto à existência de floração (nenhuma, pouca ou muita);
- quanto à existência de refoliação (nenhuma, pouca ou muita);
- condições de copa (ruim, regular, boa).

Utilizou-se, para elaboração deste trabalho, o total de frutos observados em cada um dos cinco diferentes estágios de desenvolvimento das 139 parcelas com cinco cacauzeiros, referentes aos anos agrícolas de 1977/78 a 1985/86.

### Métodos

Com a finalidade de organizar as observações periódicas obtidas durante oito anos, para cada idade do

fruto, definimos o ano agrícola cacauzeiro e seus períodos de observação.

Assim sendo, o ano agrícola para o cacauzeiro é o tempo necessário para que este, em média, passe os seus frutos do estado de bilros para a maturação e posterior colheita. Como podemos ver no esboço da Fig. 2, inicia-se em outubro de um ano civil, com a bilração de um ano agrícola, que de outubro a início de março corresponde à bilração da safra do temporão, e, de fins de março a 31 de agosto, para a safra principal. Este cacau é comercializado de primeiro de maio a 30 de setembro para a safra do temporão, e, de primeiro de outubro a 30 de abril, para a safra principal.

O ano agrícola cacauzeiro é constituído por 17 meses consecutivos, e está dividido em 24 períodos de observações de 21 dias, conforme descrito na Tabela 1. Um período de observação consiste na contagem de frutos realizada durante dez dias de levantamento de campo pelas 139 fazendas selecionadas. Entre o primeiro dia de uma dada contagem e o primeiro dia da contagem seguinte, há um intervalo de 21 dias, o mesmo ocorrendo entre os segundos dias, e assim por diante. Na descrição do período do ano civil são utilizadas expressões aproximadas das datas dos levantamentos realizados desde o ano agrícola 1977/78 até o ano agrícola 1984/85, ao invés das verdadeiras datas de início e fim das contagens, porque estas datas muitas vezes não coincidem nos anos subsequentes, atrasam ou adiantam um, dois ou três dias. Por questões administrativas o levantamento inicia-se sempre numa segunda-feira, o que prejudica a fidelidade do seu início e término.

Definida a combinação dos cinco estágios de desenvolvimento do fruto que melhor explica a produção da região cacauzeira, em diferentes períodos do ano agrícola, gera-se uma variável aleatória  $S$ , que representa esta frutificação, onde:

$S$  = número médio estimado de sacos de 60 kg de cacau em amêndoas secas, em uma dada safra de um dado período do ano agrícola, a partir da quantidade de frutos observada na amostra.

Estimam-se seus valores do seguinte modo:

$$S = \frac{Z.C.A.}{k}$$

onde:

$Z$  = número médio de frutos por cacauzeiro, que melhor representa a produção de cacau de uma dada safra em um determinado período do ano agrícola;

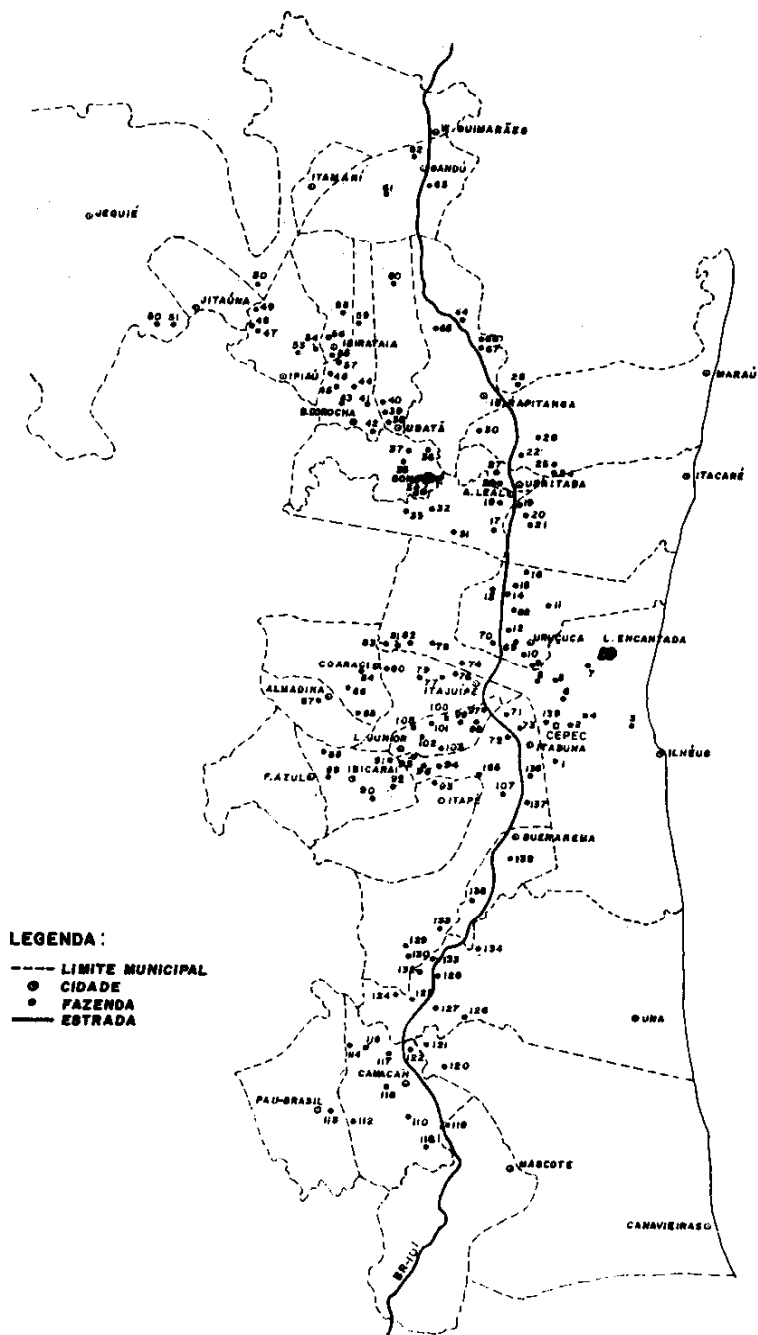


Fig. 1 Localização das fazendas selecionadas.

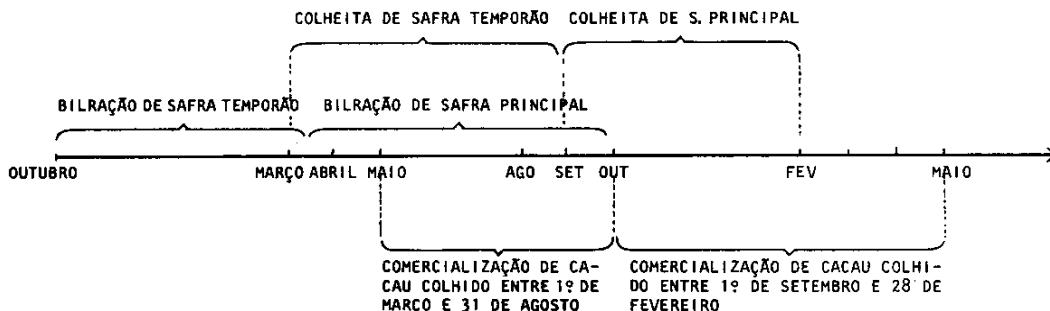


Fig. 2. Ano agrícola cacaueiro e comercialização.

TABELA 1. Definição dos períodos de observação de um ano agrícola cacaueiro.

Período do ano agrícola (J)	Período do ano civil
1	Primeira quinzena de outubro (ano anterior)
2	Segunda quinzena de outubro (ano anterior)
3	Meados de novembro (ano anterior)
4	Primeira quinzena de dezembro (ano anterior)
5	Segunda quinzena de janeiro
6	Primeira quinzena de fevereiro
7	Segunda quinzena de fevereiro
8	Primeira quinzena de março
9	Fins de março/início de abril
10	Meados de abril
11	Primeira quinzena de maio
12	Primeira quinzena de junho
13	Segunda quinzena de junho
14	Meados de julho
15	Fins de julho/início de agosto
16	Fins de agosto/início de setembro
17	Meados de setembro
18	Primeira quinzena de outubro
19	Segunda quinzena de outubro
20	Meados de novembro
21	Primeira quinzena de dezembro
22	Segunda quinzena de janeiro (ano seguinte)
23	Primeira quinzena de fevereiro (ano seguinte)
24	Segunda quinzena de fevereiro (ano seguinte)

C = número médio de cacaueiros por hectare em um dado ano agrícola;

A = área safreira da região cacaueira em um dado ano agrícola;

k = constante que representa o número médio de frutos necessário para compor um saco de 60 kg de cacau em amêndoas secas (1.517).

#### Modelo preditivo

O modelo preditivo da produção de cacau da safra do temporão ou principal da região cacaueira do estado da Bahia, resume-se na relação entre P, S, T e G, expresso pelo modelo de regressão linear múltipla

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \frac{G}{P}$$

onde:

Y = variável aleatória cujos valores são representados pelos valores da variável P; ou alguma transformação de P; ou sua recíproca (1/P); ou o produto de sua recíproca pela variável independente S (S/P);

P = produção de cacau comercializada, em sacos de 60 kg de cacau em amêndoas secas, em uma dada safra, de um dado ano agrícola;

$X_1$  = vetor de ordem  $v_1 \times 1$ , cujos elementos são representados pela variável S ou alguma transformação de S, onde  $v_1 = \epsilon$  o número de variáveis S transformadas ou não;

$X_2$  = vetor coluna de ordem  $v_2 \times 1$ , cujos elementos são representados pela variável T ou alguma transformação de T, onde  $v_2 = \epsilon$  o número de variáveis T transformadas ou não;

$X_3$  = vetor coluna de ordem  $v_3 \times 1$ , cujos elementos são representados pela variável G ou alguma transformação de G, onde  $v_3 = \epsilon$  o número de variáveis G transformadas ou não;

T = ano agrícola a que se refere a observação;

$G$  = agrupamento a que se refere a observação;

$b_0$  = termo independente do modelo;

$b_t$  =  $t$ -ésimo vetor coluna de ordem  $v_t \times 1$ , cujos elementos são os parâmetros lineares do modelo, ( $t = 1, 2, 3$ );  $e$ ,

$\epsilon$  = variável aleatória representando o termo residual, normalmente distribuída, com média zero e variância constante  $\sigma^2$ .

Os valores da variável  $G$  são obtidos a partir do estudo do perfil de desenvolvimento da frutificação, que é reavaliado à medida que se adicionam às observações do conjunto anterior de dados, as observações da frutificação de mais um ano agrícola.

A estimativa dos parâmetros do modelo é obtida pelo método dos mínimos quadrados.

A escolha do melhor modelo preditivo para um dado período do ano agrícola é baseada no exame da adequação dos possíveis modelos de regressão linear múltipla gerados. Este procedimento baseia-se na avaliação do coeficiente de determinação múltiplo, associada à análise gráfica dos resíduos versus regressores; resíduos versus  $\hat{P}$ ; e, resíduos versus ano agrícola.

#### Estudo do perfil do desenvolvimento da frutificação

Este estudo foi motivado pelo aparecimento de efeitos lineares do ano agrícola, nas análises dos resíduos dos modelos preditivos, realizadas em diferentes períodos do ano agrícola.

Empiricamente, parecia possível agrupar perfis com comportamentos semelhantes e gerar uma nova variável  $G$ , representando os agrupamentos, que, além de favorecer a explicação do modelo preditivo, restringiria o intervalo de validade da predição preferencialmente a uma região interpoladora.

Por este fato, agrupam-se perfis de frutificação, em idades diferentes, com comportamentos semelhantes, para gerar um ponto preditor composto pelo valor da variável, correspondente ao número de frutos observado em um dado período do ano agrícola da safra em estudo; pelo valor da variável correspondente ao ano agrícola com que esta safra se parece; e pelo valor da variável correspondente ao grupo que reúne as curvas de perfis semelhantes à curva da safra em estudo.

Assim sendo, a análise de agrupamentos, com a finalidade de agrupar anos agrícolas com curvas de perfil de frutificação semelhante, é realizada através da análise do componente oblíquo (Harman 1976), associada à análise fatorial pelo método de agrupamentos múltiplos, que calcula o primeiro componente principal das variáveis de cada agrupamento, a partir da matriz de correlações ou de covariâncias, e procura ma-

ximizar a soma das variâncias das variáveis originais pelos agrupamentos.

Na análise de agrupamento são considerados três conjuntos de 17 variáveis observadas durante oito anos agrícolas, sendo que para cada conjunto, cada variável recebe o valor de cada uma das 17 observações consecutivas, realizadas por período do ano agrícola cacauzeiro, de um mesmo tipo de frutificação.

Descrevem-se os conjuntos como se segue:

**Primeiro conjunto:** número total de frutos jovens (bilros e pequenos) observado na amostra durante oito anos agrícolas, de 1977/78 a 1984/85, cujas variáveis são representadas por  $bp_{ij}$  ( $i = 1, 2, \dots, 8$ ;  $j = 1, 2, \dots, 17$ );

**Segundo conjunto:** número total de frutos em desenvolvimento (pequenos, médios, adultos e maduros) observado na amostra durante oito anos agrícolas, de 1977/78 a 1985/86, cujas variáveis são representadas por  $pmvm_{ij}$  ( $i = 1, 2, \dots, 8$ ;  $j = 5, 6, \dots, 21$ );

**Terceiro conjunto:** número total de frutos adultos e maduros observado na amostra durante oito anos agrícolas, de 1977/78 a 1985/86, cujas variáveis são representadas por  $vm_{ij}$  ( $i = 1, 2, \dots, 8$ ;  $j = 7, 8, \dots, 23$ ).

O índice  $i$  representa o ano agrícola cacauzeiro, e o índice  $j$  o período do ano agrícola cacauzeiro.

Segundo Mardia et al. (1979), o número de agrupamentos possíveis deve obedecer às seguintes desigualdades:

$$n_k \geq n_f + 1 \text{ e } N \geq g(n_f + 1), \quad k = 1, 2, \dots, g,$$

onde,

$n_k$  é o número de anos agrícolas classificados em um agrupamento;

$n_f$  é o número de frutificações distintas no estudo de agrupamentos;

$N$  é o número de anos agrícolas observados; e,

$g$  é o número de agrupamentos a ser determinado.

#### Previsão de safras

A previsão de uma dada safra ( $r$ ), em um período ( $j$ ) do ano agrícola é o valor estimado  $\hat{P}_{rj} = \hat{b}_{0,rj} + \hat{b}_{1,rj}x_{1,rj} + \hat{b}_{2,rj}x_{2,rj} + \hat{b}_{3,rj}x_{3,rj}$ ,

( $r = 1$ , safra do temporão;  $r = 2$ , safra principal)

onde,

$\hat{b}_{0,rj}$  é o estimador de mínimos quadrados do termo independente do modelo preditivo da safra  $r$  do período  $j$  do ano agrícola;

$\hat{b}_{t,rj}$  são vetores coluna de ordem  $v_t \times 1$ , cujos elementos são estimadores de mínimos quadrados dos parâmetros do modelo preditivo da safra  $r$  do período  $j$  do ano agrícola ( $t = 1, 2, 3$ );

$x_{t,rj}$  são vetores coluna de ordem  $v_t \times 1$ , cujos elementos são os valores observados das variáveis do modelo

preditivo da safra r que se deseja prever no período j do (N+1)-ésimo ano agrícola (t = 1, 2, 3);

Representando-se o ponto preditor pelo vetor de observações

$$X'_{rj} = (1, x'_{1,rj}, x'_{2,rj}, x'_{3,rj}),$$

o intervalo de confiança com probabilidade de 100 (1- $\alpha$ )% para esta observação futura é

$$\hat{p}_{rj} \pm \hat{\sigma} \sqrt{(1+x'_{rj}(X'X^{-1})x_{rj})} < P < \hat{p}_{rj} + \hat{\sigma} \sqrt{(1+x'_{rj}(X'X^{-1})x_{rj})}$$

onde

t é o valor crítico tabelado da distribuição "t" de Student, com N-p graus de liberdade, ao nível  $\alpha$  de probabilidade, onde p é o número de parâmetros do modelo, e  $\hat{\sigma}$  é o desvio-padrão residual estimado do modelo preditivo.

### Intervalo de validade da predição

O intervalo de validade de predição depende dos dados observados na construção amostral do modelo preditivo e é, possivelmente, o menor volume sólido que inclui todos pontos amostrais.

Previsões mais seguras são realizadas nos casos em que o ponto preditor é um ponto de interpolação, pois este ponto comporta-se como um dos pontos da construção amostral e poucas hipóteses devem ser satisfeitas. No caso de o ponto preditivo ser um ponto de extrapolação, deve-se formular uma hipótese de que a função de predição estimada seja válida fora do intervalo estudado e, embora muitas vezes a extrapolação seja menos confiável, pode acontecer que seja o ponto de interesse em um determinado momento.

Segundo Weisberg (1985) e Montgomery & Peck (1982), consideram-se duas aproximações para definir o intervalo de validade da menor figura fechada que inclui todos os pontos amostrais.

Seja

$$h = x'(X'X)^{-1}x$$

x uma constante cujos contornos são elipsóides, onde x é um vetor qualquer de observações, de dimensões (p+1)x1.

Se  $h_{max} = \max [h_i], i = 1, 2, \dots, N$ , na construção amostral, o conjunto de todos os pontos  $x_*$  tais que  $x'_*(X'X)^{-1}x_* \leq h_{max}$  é um elipsóide que inclui toda a construção amostral.

Para uma predição no ponto  $x_*$ , se

$$x'_*(X'X)^{-1}x_* > h_{max}$$

então  $x_*$  não está no elipsóide e a predição pode ser vista como uma extrapolação.

Para prover uma melhor aproximação, obtém-se o menor volume do elipsóide que contenha estes pontos, e que é chamado de elipsóides de cobertura mínima (MCE), conforme um algoritmo em Titterington

(1978). E uma aproximação é chamar qualquer ponto fora deste elipsóide uma extrapolação, pois ainda pode haver regiões dentro da MCE, onde não haja pontos amostrais.

## RESULTADOS E DISCUSSÃO

As previsões obtidas para os anos agrícolas 1981/82 a 1985/86 são encontradas nas Tabelas 2a a 2e, respectivamente, comparadas às de dois métodos, através do erro de estimação, relativo à produção real.

O primeiro método comparativo baseia-se na amostragem casual do estrato da produção por microrregião cacauceira, em que a unidade amostral é a informação subjetiva da produção de cacau da propriedade, e o segundo é o utilizado pelo Gill & Duffus Group, cujas previsões

**TABELA 2a. Previsões das safras do temporão e principal do ano agrícola 1981/82, em milhões de sacos de 60 kg de cacau em amêndoas secas, com base no método de contagem de frutos, amostragem e Gill & Duffus, com os respectivos erros de estimação, em % (Erro), relativos ao verdadeiro valor das safras, com N = 4 anos de observações.**

Safras	Período do ano agrícola (J)	Métodos		
		Contagem	Amostragem	Gill & Duffus
		Previsão (erro)	Previsão (erro)	Previsão (erro)
Temporão	6	2,804 (3,03)	2,718 (6,00)	2,900 (0,29)
	9	2,891 (0,02)	2,718 (6,00)	2,900 (0,29)
	12	2,900 (0,28)	2,727 (5,66)	2,555 (11,81)
Principal	13	1,740 (1,32)	2,107 (22,70)	2,000 (16,43)
	16	1,870 (8,84)	2,094 (21,92)	1,950 (13,52)

**TABELA 2b. Previsões das safras do temporão e principal do ano agrícola 1982/83, em milhões de sacos de 60 kg de cacau em amêndoas secas, com base no método de contagem de frutos, amostragem e Gill & Duffus, com os respectivos erros de estimação, em % (Erro), relativos ao verdadeiro valor das safras, com N = 5 anos de observações.**

Safras	Período do ano agrícola (J)	Métodos		
		Contagem	Amostragem	Gill & Duffus
		Previsão (erro)	Previsão (erro)	Previsão (erro)
Temporão	6	3,043 (3,99)	2,560 (19,23)	3,500 (10,42)
	9	3,137 (1,02)	2,560 (19,23)	3,500 (10,42)
	12	3,148 (0,70)	2,647 (16,48)	3,300 (4,11)
Principal	13	1,736 (17,90)	1,831 (13,39)	1,740 (17,71)
	16	1,736 (17,90)	2,103 (0,51)	1,900 (10,14)

**TABELA 2c. Previsões das safras do temporão e principal do ano agrícola 1983/84, em milhões de sacos de 60 kg de cacau em amêndoas secas, com base no método de contagem de frutos, amostragem e Gill & Duffus, com os respectivos erros de estimação, em % (Erro), relativos ao verdadeiro valor das safras, com N = 6 anos de observações.**

Safras	Período do ano agrícola (J)	Métodos		
		Contagem	Amostragem	Gill & Duffus
		Previsão (erro)	Previsão (erro)	Previsão (erro)
Temporão	6	2,720 (11,80)	2,493 (19,18)	2,200 (28,67)
	9	2,621 (15,00)	2,493 (19,18)	2,500 (18,94)
	12	3,031 (1,73)	2,547 (17,41)	2,700 (12,45)
Principal	13	-	2,444 (10,38)	2,400 (10,46)
	16	2,758 (1,14)	3,012 (10,46)	2,550 (6,49)

são encontradas em Gill & Duffus Cocoa Market Report (1981/1986).

Com relação aos resultados obtidos para os anos agrícolas 1981/82 a 1984/85, restringimos a apresentar os valores das previsões e os seus erros de estimação relativos ao verdadeiro valor da safra, comparando os resultados aos dos outros dois métodos.

São exemplificados os passos utilizados no estudo das safras do ano agrícola 1985/86. A construção amostral do modelo preditivo baseia-se em N=8 observações. Os dados usados são encontrados na Tabela 3.

Durante o ano agrícola 1981/82, as previsões obtidas para a safra do temporão pelo método de contagem de frutos têm erros de estimação variando entre 0,28% e 3,03%, enquanto que para o método de amostragem entre 5,66% e 6,00%, e para o método do grupo Gill & Duffus

entre 0,29% e 11,81%. Para a safra principal, a variação do erro de estimação do método de contagem de frutos é de 1,32% a 8,84%, e do grupo Gill & Duffus, de 21,92% a 22,70%.

Para o ano agrícola 1982/83, os erros de estimação para a safra do temporão variam entre 0,70% e 3,99%, 16,48% e 19,23%, e, 4,11% e 10,42%, pelos métodos de contagem de frutos-amostragem e do grupo Gill & Duffus, respectivamente. E para a safra principal, obtém-se um único valor, pelo método de contagem, de 17,90%; pelo segundo método, uma variação entre 0,51% e 13,39%; e pelo terceiro, entre 10,14% e 17,71%.

Para o ano agrícola 1983/84, os erros de estimação para a safra do temporão variam entre 1,73% e 11,80%, 17,41% e 19,18%, e, 12,45% e 28,67%, pelo método de contagem de frutos-amostragem e do grupo Gill & Duffus, respecti-



**TABELA 2d.** Previsões das safras do temporão e principal do ano agrícola 1984/85, em milhões de sacos de 60 kg de cacau em amêndoas secas, com base no método de contagem de frutos, amostragem, e Gill and Duffus, com os respectivos erros de estimação, em % (Erro), relativos ao verdadeiro valor das safras, com N = 7 anos de observações.

Safras	Período do ano agrícola (J)	Métodos		
		Contagem	Amostragem	Gill & Duffus
		Previsão (erro)	Previsão (erro)	Previsão (erro)
Temporão	6	2,092 (17,15)	2,176 (21,84)	2,800 (56,80)
	9	2,062 (15,47)	2,176 (21,84)	1,850 (3,60)
	12	1,772 (0,77)	1,947 (10,56)	1,700 (4,80)
Principal	13	2,855 (11,99)	3,235 (0,29)	2,800 (13,69)
	16	3,178 (2,04)	3,132 (3,44)	3,100 (4,44)

vamente. E para a safra principal, obtém-se um único valor, pelo método de contagem, de 1,14%; pelo segundo método, uma variação entre 10,38% e 10,46%; e pelo terceiro, entre 6,49% e 10,46%.

Para o ano agrícola 1984/85, os erros de estimação para a safra do temporão variam entre 0,77% e 17,15%, 10,56% e 21,84%, e, 3,60% e 56,80%, pelo método de contagem de frutos-amostragem e do grupo Gill & Duffus, respectivamente. E para a safra principal, obtém-se uma variação, pelo método de contagem, entre 2,04% e 11,99%; pelo segundo método, uma variação entre 0,29% e 3,44%; e pelo terceiro, entre 4,44% e 13,69%.

#### Classificação de perfis de frutificação

O procedimento inicial deste estudo é a classificação dos oito anos agrícolas, para três idades

**TABELA 2e.** Previsões das safras do temporão e principal do ano agrícola 1985/86, em milhões de sacos de 60 kg de cacau em amêndoas secas, com base no método de contagem de frutos, amostragem e Gill & Duffus, com os respectivos erros de estimação, em % (Erro), relativos ao verdadeiro valor das safras, com N = 8 anos de observações.

Safras	Período do ano agrícola (J)	Métodos		
		Contagem	Amostragem	Gill & Duffus
		Previsão (erro)	Previsão (erro)	Previsão (erro)
Temporão	6	3,091 (1,11)	2,717 (13,08)	-
	9	3,051 (2,39)	2,717 (13,08)	3,000 (4,03)
	12	3,165 (1,09)	2,881 (7,83)	2,850 (8,82)
Principal	13	2,994 (3,20)	-	3,250 (12,03)
	16	2,855 (1,58)	-	2,800 (3,48)

des diferentes dos frutos, a fim de agrupar perfis de frutificação com comportamentos semelhantes e gerar a variável G, efetuando a análise do componente oblíquo, conforme o item Estudo do perfil do desenvolvimento da frutificação.

Em uma análise de agrupamentos, pode ocorrer pelo menos um perfil distinto por agrupamento ( $N_f \geq 1$ ). Assim, o número de perfis classificados no k-ésimo grupo ( $n_k$ ) deve ser, no mínimo, 2, ou seja,  $n_k \geq n_f + 1 \geq 2$ .

Observaram-se os perfis de frutificação de oito anos agrícolas ( $N=8$ ). Como  $N \geq g(p+1)$ , temos que  $g \leq 4$ , ou seja, o número de agrupamentos possíveis, é, no máximo, igual a quatro.

Nos resultados da análise de agrupamentos, que se encontram na Tabela 4, os perfis de frutificações são classificados em três grupos ( $g = 3$ ). As classificações obtidas são originadas pela matriz de correlações, embora a análise te-

**TABELA 3.** Produções das safras do temporão (PT) e principal (PP), em milhões de sacos de cacau em amêndoas secas; área safreira em hectares (A); número médio estimado de milhões de sacos de cacau em amêndoas secas (S<sub>j</sub>), nos períodos j = 6, 9, 12, 13 e 16, por ano agrícola cacaueiro.

Ano agrícola	PT <sup>1</sup>	PP <sup>1</sup>	A <sup>2</sup>	S <sub>6</sub>	S <sub>9</sub>	S <sub>12</sub>	S <sub>13</sub>	S <sub>16</sub>
1977/78	2,102	1,887	410182	1,197	3,006	1,933	9,058	5,525
1978/79	2,524	1,805	417182	0,407	3,071	2,517	8,834	3,718
1979/80	3,129	2,257	427671	1,833	5,392	2,900	10,619	4,173
1980/81	2,383	2,658	438113	0,736	2,132	1,692	8,680	6,347
1981/82	2,892	1,718	446139	1,164	3,810	3,022	8,396	3,618
1982/83	3,170	2,114	457223	1,453	4,600	2,951	9,209	3,921
1983/84	3,084	2,727	479442	0,609	2,429	2,649	9,887	6,171
1984/85	1,786	3,244	506263	0,532	2,243	1,209	8,363	8,072
1985/86	3,126	2,901	548796	1,427	4,611	4,492	11,435	7,303

<sup>1</sup> Dados fornecidos pela COMCAUBA - Comércio de Cacau da Bahia.

<sup>2</sup> Dados fornecidos pelo DEPEX - Departamento de Extensão da CEPLAC/Comissão Executiva do Plano da Lavoura Cacaueira.

na sido realizada também para a matriz de covariâncias. Adota-se a classificação baseada na matriz de correlações, pois além de o resultado ter sido o mesmo para os três conjuntos de dados, este estudo avalia comportamentos de perfis de curvas, cujos dados originam-se de observações de variáveis com as mesmas unidades (número de frutos), e a matriz de correlações aproxima melhor os perfis de formas semelhantes. O resultado com a matriz de covariâncias difere para o terceiro conjunto, quando o ano agrícola 1979/80, do grupo 3 é classificado no grupo 2. Adota-se, então, a seguinte classificação em três grupos, dos perfis de frutificação dos oitos anos agrícolas:

Grupo	Ano agrícola
1	1977/78 , 1980/81 e 1984/85
2	1978/79 e 1983/84
3	1979/80 , 1981/82 e 1982/83

Essa classificação está representada na Fig. 3, para as três diferentes idades do fruto.

Mantida a ordenação dos anos agrícolas, codificam-se os três primeiros anos por 1, 2 e 3, respectivamente. Aos anos restantes são desig-

nados os códigos destes três primeiros, segundo a análise de agrupamentos. Gera-se, através destes códigos, a variável G, qualitativa, que representa o agrupamento de anos agrícolas com perfis de frutificação com comportamentos semelhantes.

#### Modelo preditivo

Os modelos preditivos para um dado período j do ano agrícola para as safras do temporão e principal, considerados adequados, foram os seguintes:

$$\hat{P}_{1,j} = \hat{b}_0 + \hat{b}_{1,1} S_j + \hat{b}_{2,1} T^2 + \hat{b}_{2,2} T + \hat{b}_{3,1} G$$

(j=6,9)

$$\hat{P}_{1,j} = \frac{S_j}{\hat{b}_0 + \hat{b}_{1,1} S_j} + \hat{b}_{3,1} G$$

(j=12)

safr

principal

$$\hat{P}_{2,j} = \hat{b}_0 + \hat{b}_{1,1} S_j + \hat{b}_{2,1} T^2 + \hat{b}_{2,2} T + \hat{b}_{3,1} G$$

(j=13,16)

Mostramos, na Tabela 5, as estimativas dos parâmetros do modelo preditivo, seus desvios-padrão e seus respectivos níveis mínimos de significância; o quadrado médio residual, os coeficientes de determinação e determinação

TABELA 4. Resumo de análise de agrupamentos.

Matriz original	Correlação			Covariância			
	variáveis	bp	pmvm	vm	bp	pmvm	vm
	num. de membros	3	3	3	3	3	3
	variação do agrupamento	3,00	3,00	3,00	9459083	64372396	21360846
Grupo 1	variação explicada	2,49	2,88	2,68	8048266	61857920	19023323
	proporção	0,83	0,96	0,89	0,85	0,96	0,89
	segundo valor característico	0,37	0,07	0,23	890233,9	1712832	1840046
	num. de membros	2	2	2	2	2	3
	variação do agrupamento	2,00	2,00	2,00	11235448	117516110	38422209
Grupo 2	variação explicada	1,90	1,90	1,90	10707269	111982017	33451605
	proporção	0,95	0,95	0,95	0,95	0,95	0,87
	segundo valor característico	0,09	0,09	0,10	528178,9	5534093	4849460
	num. de membros	3	3	3	3	3	2
	variação do agrupamento	3,00	3,00	3,00	8840784	104866593	14644231
Grupo 3	variação explicada	2,68	2,81	2,60	7946765	98600539	13568979
	proporção	0,89	0,94	0,87	0,90	0,94	0,93
	segundo valor característico	0,18	0,14	0,26	505220,2	4708327	1075252
	variação total explicada	7,08	7,60	7,18	26702300	272440476	66043907
proporção	0,88	0,95	0,90	0,90	0,95	0,89	
máximo valor característico da matriz de dados original	8	8	8	315,62	100,04	286,56	

ajustado; e, a estatística "F" de Fisher-Snedecor, com seu nível mínimo de significância. Na Tabela 6, são mostrados os valores da constante h de cada ponto da construção amostral do modelo preditivo, de cada período j do ano agrícola

1985/86; e na Tabela 7, os valores das predições, os respectivos limites de confiança para o verdadeiro valor da safra, com 90% de probabilidade, e os respectivos valores da constante h, para cada valor de T e G.

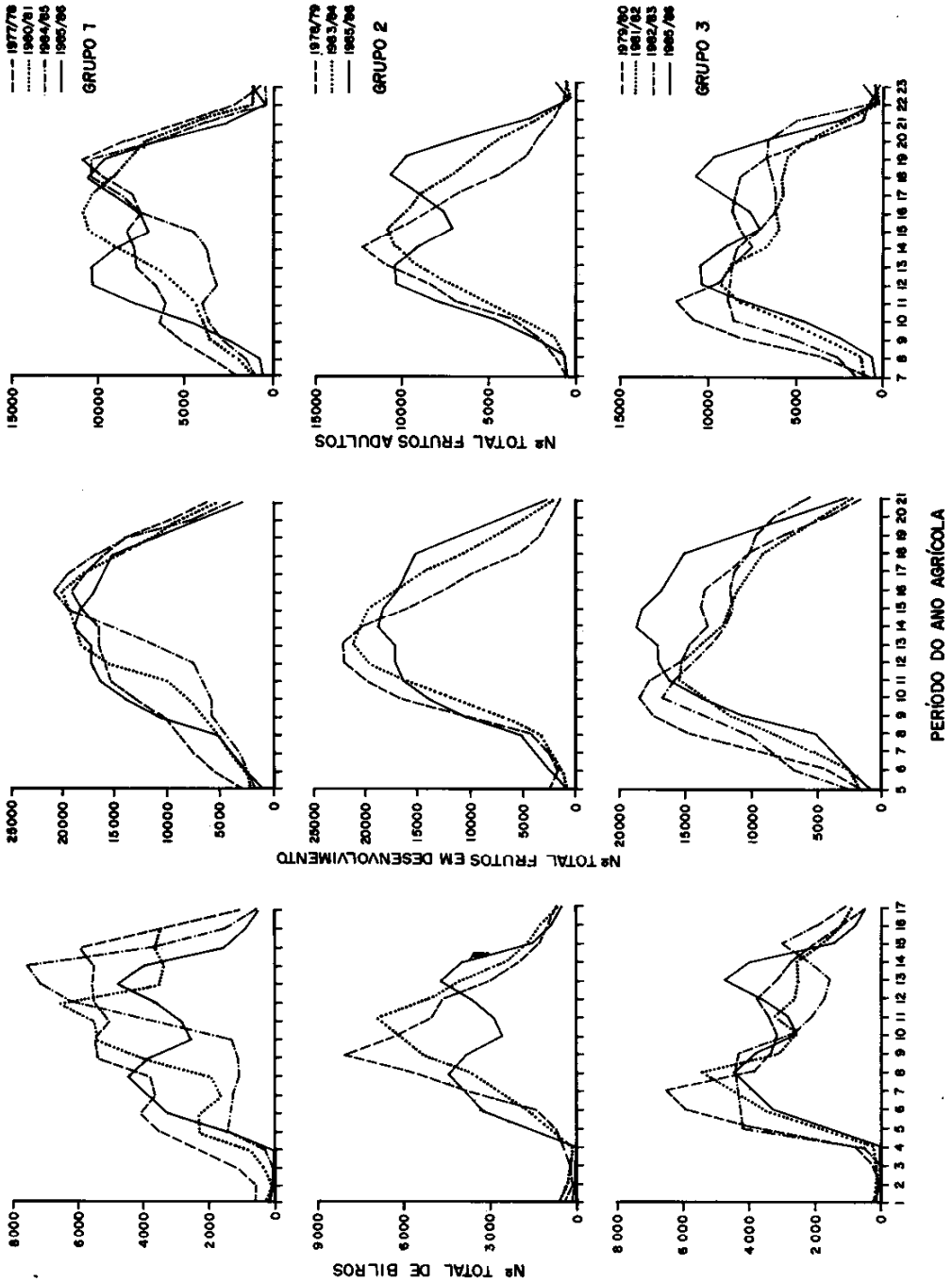


Fig. 3 Perfil da frutificação em três idades diferentes, agrupado em anos com comportamentos semelhantes.

**TABELA 5** - Estimativas para as safras do temporão e principal, em um período  $j$  do ano agrícola, dos parâmetros dos modelos preditivos, seus desvios-padrão e níveis mínimos de significância, entre parênteses; os coeficientes de explicação ( $R^2$ ) e explicação ajustados ( $R^2$ ); o quadrado médio residual estimado ( $\sigma^2$ ), as estatísticas "F" de Fisher-Snedecor e seu nível mínimo de significância, entre parênteses.

j	safra do temporão			safra principal	
	6	9	12	13	16
$\hat{b}_0$	1,7391 0,1516 (0,0014)	1,7976 0,1138 (0,0006)	0,4303 0,1435 (0,0502)	-1,3909 0,4568 (0,0557)	-0,9206 1,1773 (0,4913)
$\hat{b}_{1,1}$	0,0092 0,0060 (0,2227)	$8,05 \times 10^{-6}$ $3,07 \times 10^{-6}$ (0,0788)	0,2510 0,1230 (0,0969)	0,5542 0,0493 (0,0015)	0,4586 0,1707 (0,0747)
$\hat{b}_{2,1}$	$-1,73 \times 10^{-28}$ $3,75 \times 10^{-29}$ (0,0189)	$-1,82 \times 10^{-28}$ $2,80 \times 10^{-28}$ (0,0074)	-	-0,0512 0,0089 (0,0106)	-0,0180 0,0212 (0,4572)
$\hat{b}_{2,2}$	0,1114 0,0327 (0,0422)	0,1155 0,0237 (0,0164)	-	0,6583 0,0829 (0,0042)	0,1621 0,1933 (0,4635)
$\hat{b}_{3,1}$	0,2117 0,0948 (0,1116)	0,1982 0,0666 (0,0587)	-0,0696 0,0890 (0,4697)	-0,1964 0,0521 (0,0326)	0,2858 0,2560 (0,3456)
$R^2$	97,01	98,38	72,87	99,05	93,40
$R^2$	93,03	96,23	62,01	97,78	84,59
$\sigma^2$	0,0187	0,0101	0,0064	0,0078	0,0439
F	24,374 (0,0127)	45,632 (0,0510)	6,714 (0,0384)	78,225 (0,0023)	10,610 (0,0407)

#### Ponto preditor, predição e validade de predição

É necessário, para a criação do ponto preditivo, estabelecer valores para as variáveis T e G, com base na semelhança de perfis. Representam-se as curvas de frutificação, para três idades diferentes, do ano agrícola 1985/86, em relação aos perfis das curvas de cada um dos três agrupamentos, conforme mostrado na Fig. 3.

As previsões da safra do temporão são obtidas para os períodos  $j = 6, 9$  e  $12$ .

No período  $j = 6$  (primeira quinzena de fevereiro), observa-se que o perfil de frutificação jovem do ano agrícola 1985/86 assemelha-se, até este período, aos perfis dos anos agrícolas pertencentes ao terceiro grupo, e, particularmente, ao perfil do ano agrícola 1981/82, conforme mostra a Fig. 3. Assim, teremos os valores 5 e 3 para as variáveis T e G, respectivamente.

A constante  $h$  máxima (item Intervalo de validade da predição) da construção amostral do

**TABELA 6** Valores da constante h para cada ponto da construção amostral do modelo preditivo.

Ano agrícola	Período do ano agrícola (j)				
	6	9	12	13	16
1977/78	0,572	0,564	0,437	0,711	0,711
1978/79	0,681	0,462	0,184	0,359	0,549
1979/80	0,852	0,979	0,333	0,684	0,630
1980/81	0,505	0,524	0,292	0,983	0,983
1981/82	0,454	0,512	0,292	0,643	0,375
1982/83	0,353	0,363	0,307	0,348	0,427
1983/84	0,580	0,595	0,325	0,518	0,285
1984/85	1,000	1,000	0,831	0,754	0,739

modelo preditivo com N = 8 anos agrícolas, tem valor 1,000, enquanto que a do ponto preditivo,  $h = 0,302$ , garante a interpolação da predição, ou seja, que o ponto preditor se comporte como um ponto da construção amostral. A previsão obtida neste período é de 3,091 milhões de sacos de cacau em amêndoas secas, variando entre 2,914 e 3,268 sacos, cujo erro de estimação é de 1,11%; relativamente ao verdadeiro valor da safra, o método de amostragem apresenta erro de estimação de 13,08%.

No período  $j = 9$  (fins de março, início de abril), observa-se que é mantida a semelhança do novo perfil da curva de frutificação, agora da

**TABELA 7** - Predições ( $\hat{P}_{T,j}$ ) do ano agrícola 1985/86, em milhões de sacos de 60 kg de cacau em amêndoas secas, os respectivos valores da constante h, e os limites superior (LS) e inferior (LI) do intervalo de confiança para o verdadeiro valor da safra, com 90% de probabilidade, para cada valor de T e G.

T	1	2	3	4	5	6	7	8
G	1	2	3	1	3	3	2	1
$\hat{P}_{1,6}$	2,222	2,545	2,868	2,556	3,091	3,202	3,102	1,919
LI	1,949	2,369	2,629	2,228	2,914	3,012	2,786	1,538
LS	2,495	2,721	3,107	2,884	3,268	3,393	3,418	2,300
h	0,721	0,299	0,551	1,038	0,302	0,350	0,967	1,402
$\hat{P}_{1,9}$	2,193	2,506	2,820	2,539	3,051	3,167	3,084	1,866
LI	2,009	2,375	2,631	2,333	2,913	3,024	2,877	1,617
LS	2,376	2,638	3,009	2,745	3,190	3,309	3,291	2,114
h	0,602	0,310	0,638	0,757	0,343	0,361	0,762	1,093
$\hat{P}_{1,12}$	3,018	3,167	3,330	3,018	3,330	3,330	3,167	3,018
LI	1,997	2,238	2,526	1,997	2,526	2,526	2,238	1,997
LS	6,181	5,410	4,884	6,181	4,884	4,884	5,410	6,181
h	18,957	10,909	5,343	18,957	5,343	5,343	10,909	18,957
$\hat{P}_{2,13}$	5,358	5,666	5,783	6,565	6,371	6,466	6,656	6,743
LI	5,034	5,396	5,625	6,172	6,108	6,203	6,354	6,377
LS	5,682	5,937	6,120	6,959	6,633	6,730	6,958	7,110
h	2,411	1,682	1,406	3,561	1,585	1,596	2,096	3,091
$\hat{P}_{2,16}$	2,858	3,252	3,610	3,074	3,645	3,608	3,250	2,855
LI	2,019	2,034	1,953	2,430	2,248	2,366	2,696	2,366
LS	3,698	4,470	5,266	3,718	5,041	4,851	3,803	3,345
h	2,903	6,113	11,301	1,709	8,032	6,356	1,261	0,987

frutificação em desenvolvimento, à do perfil do ano agrícola 1981/82 dentro do grupo 3, mantendo os valores 5 e 3, para as variáveis T e G, respectivamente.

A constante  $h$  máxima da construção do modelo com  $N = 8$  observações, tem valor 1,000, enquanto que o ponto preditor,  $h = 0,343$ , o que garante que a previsão seja uma interpolação. A previsão obtida neste período é de 3,051 milhões de sacos de 60 kg de cacau em amêndoas secas, variando entre 2,913 e 3,190 sacos, cujo erro de estimação, relativamente ao verdadeiro valor da safra é de 2,39%, os erros do método de amostragem e do grupo Gill & Duffus são 13,08% e 4,03%, respectivamente.

No período  $j = 12$  (primeira quinzena de junho), quando é realizada a última previsão da safra do temporão, a construção amostral do modelo baseia-se na frutificação adulta. Nesta fase conta-se com doze observações da curva de perfil da frutificação jovem, oito observações da curva de frutos em desenvolvimento e seis observações de frutos adultos. Neste período, o perfil da curva de frutos adultos corresponde ao perfil da curva de frutos jovens até o período  $j = 6$  e ao perfil da curva em desenvolvimento até o período  $j = 9$ . Os frutos adultos são frutos por amadurecer e portanto com maior probabilidade de sobrevivência, fato que justifica a aproximação deste perfil ao perfil da produção do ano agrícola em estudo, correspondente à produção da safra do temporão. Esta idade do fruto é a resposta de todos os fatores climáticos que influenciam a produção final desta safra, e portanto representam aproximadamente o perfil do seu comportamento final.

Comparando-se, então, até este período, o perfil da frutificação em desenvolvimento e adulta do ano agrícola 1985/86, em relação aos perfis dos anos anteriores, dentro de cada grupo, conclui-se que a maior semelhança está dentro do grupo 2, com maior tendência a se aproximar do ano agrícola 1983/84; portanto, os valores para as variáveis T e G são, respectivamente, 7 e 2. A constante  $h$  (máxima) = 0,831 e a constante  $h$  do ponto preditor tem valor 10,909, portanto, a previsão é uma extrapolação. A previsão nesse período é de 3,167 mi-

lhões de sacos de 60 kg de cacau em amêndoas secas, cujo erro de estimação, relativo ao verdadeiro valor da safra é de 1,09%; os outros dois métodos erram 7,03% e 8,82%, respectivamente.

As previsões para a safra principal são obtidas para os períodos  $j = 13$  e 16.

No período  $j = 13$  (segunda quinzena de junho), é efetuada a primeira previsão de safra principal. A produção total do ano agrícola é prevista, e a previsão da safra principal é obtida pela diferença entre o total previsto e o previsto para a safra do temporão no período  $j = 12$ . A construção amostral do modelo baseia-se na frutificação em desenvolvimento observada neste período, acrescida da frutificação madura acumulada entre os períodos  $j = 8$  e  $j = 13$ . O que corresponde aproximadamente ao total de frutos que representará a produção total do ano agrícola 1985/86, na amostra.

A análise gráfica (Fig. 3) mostra que o perfil das frutificações em desenvolvimento e adulta, até este período do ano agrícola, com nove e sete observações, respectivamente, assemelha-se aos perfis do grupo 2, entre os perfis dos anos 1978/79 e 1983/84 (segundo e sétimo anos agrícolas, respectivamente).

Temos  $h$  (máxima) = 0,983 da construção amostral do modelo preditivo. As constantes  $h$  para as predições, considerando as semelhanças com os anos agrícolas 1978/79 e 1984/84, são 1,682 e 2,096, respectivamente. Confirma-se a extrapolação nestes pontos. Observa-se, através da posição deste perfil, que a produção final do ano agrícola deverá estar entre o previsto nestes dois pontos preditores. Assim, utiliza-se a média aritmética do previsto para estes dois pontos. A previsão do ano agrícola 1985/86 é, então, de 6,161 milhões de sacos de cacau em amêndoas secas, cujo erro de estimação relativo ao verdadeiro valor é 2,22%. Portanto, o valor previsto para a safra principal é de 2,994 milhões de sacos de cacau em amêndoas secas, com um erro de estimação relativo ao verdadeiro valor da safra de 3,20%; o método do grupo Gill and Duffus erra 12,03%, e o método de amostragem não mostra resultado.

Finalizando, no período  $j = 16$  (fins de agos-

to, início de setembro), utiliza-se a frutificação em desenvolvimento para prever novamente a safra principal. Consideram-se para comparações de perfis somente as ocorrências entre os períodos  $j = 9$  e  $j = 17$  da frutificação jovem, pois nota-se empiricamente que estes perfis diferem em relação aos perfis do ano todo. Assim, o perfil da frutificação jovem da safra principal assemelha-se ao do ano agrícola 1984/85 no grupo 1, enquanto que o perfil da frutificação jovem do ano todo assemelha-se, em média, com os perfis do grupo 2. Considera-se, então, como o melhor critério de classificação, o primeiro estudo, por estar mais evidente a semelhança. A constante  $h$  máxima da construção amostral do modelo preditivo é 0,983, e a constante  $h$  do ponto preditor, 0,987; essa previsão é uma extrapolação, e seu valor 2,885 milhões de sacos de cacau em amêndoas secas, com erro de estimação relativo ao verdadeiro valor da safra igual a 1,58%; o método do Gill & Duffus erra 3,48%, e também não existe resultado para o método de amostragem.

Das quinze previsões da safra do temporão, o método de contagem perde duas vezes para o do grupo Gill & Duffus: uma vez na visita 9 do ano agrícola 1984/85; e uma vez na visita 6 do ano agrícola 1981/82. Mas, em todas as previsões obtidas nas visitas 12, que são previsões mais seguras, por se tratar de estimativas baseadas em frutos adultos - portanto com maior probabilidade de sobrevivência -, os erros de estimação sempre são os menores e não ultrapassam 2,00%. Além disso, referem-se à última previsão, que é obtida para a safra do temporão, com três meses de antecedência ao final desta safra.

Com relação à safra principal, das dez previsões obtidas, o método de contagem é melhor seis vezes, com exceção dos resultados da visita 16, do ano agrícola 1982/83, quando o erro da estimação é de 0,51%, para o método de amostragem; para os anos agrícolas 1983/84, 1984/85 e 1985/86, não ultrapassam 2,04%; e para o ano 1981/82 é de 8,82%. As observações desta visita têm o mesmo valor que as observações da visita 12, no que diz respeito à safra principal. E a

previsão é obtida seis meses antes do final desta safra.

## CONCLUSÕES

1. Pelo fato de existirem poucos anos de observação, o estudo não estabelece um modelo preditivo único. No entanto, propõe uma pesquisa sobre o comportamento do perfil da frutificação de um ano agrícola em relação aos outros anos.

2. O estudo de perfis da frutificação fica estabelecido como fonte principal de informação para um prognóstico da safra. E o segundo lugar cabe à estimativa do modelo preditivo adequado, baseada no conjunto de pontos amostrais. Reunindo-se estas fontes de informações, fica fortalecido o intervalo de validade da previsão.

3. Para isto, hipóteses de semelhança são formuladas empiricamente quanto ao comportamento da frutificação. Tais hipóteses auxiliam a seleção do ponto preditivo, de modo que este seja um ponto de interpolação. Também, auxiliam o modelo preditivo estimado a ser relevante fora da construção amostral, no caso de o ponto preditivo ser um ponto de extrapolação, tornando-o mais confiável.

4. Pelas previsões com baixos erros de estimação, obtidas com antecedência de três a seis meses do final de cada safra, e pelo bom prognóstico da produção total do ano agrícola, com oito meses de antecedência, acredita-se ser este método de previsão da safra agrícola de cacau, com base em contagem de frutos, aliado ao estudo do perfil de frutificação de cada ano agrícola, até o momento, o único confiável para este cultivo.

## REFERÊNCIAS

- ALVIM, P. T.; NASCIMENTO, I. F.; PEREIRA, C. P. Ciclo entre a polinização e a maturação dos frutos em cacauzeiro. *Informe Técnico do Centro de Pesquisas do Cacau*. Ilhéus, 1965. p.16-18.
- ARROYO VERGARA, J. R.; CAMPOS, H.; MONTEIRO, A.; BORDONI, O. F. G. Dimensionamento de amostra para a previsão de safras do



- cacau do Estado da Bahia e resultados iniciais. *Revista Theobroma*, Ilhéus, v.10, n.2, p.79-80, 1980a.
- ARROYO VERGARA, J. R.; MONTEIRO, A.; ALVES, F. A.; BORDONI, O. F. G. Aspectos da distribuição de produção e área de cacau do Estado da Bahia e resultados iniciais. Ilhéus: Centro de Pesquisas do Cacau, 1980b. 21p. (Boletim Técnico, 75).
- FAO (Roma, Itália). Comitê de Problemas de Produtos Básicos. Grupo de Estudio sobre el Cacao. Comitê de Estadística. Técnicas de pronósticos relativas a la producción de cacao em Ghana. Roma, 1967. 3p.
- FAO (Roma, Itália). Comitê de Problemas de Produtos Básicos. Grupo de Estudio sobre el Cacao. Comitê de Estadística. Novedades en las técnicas y pronósticos de la producción de cacao. Roma, 1968. 3p.
- GILL & DUFFUS COCOA MARKET REPORT, Londres, n.292-324, 1981/1986.
- HARMAN, H. H. *Modern factor analysis*. 3.ed. Chicago: The University of Chicago Press, 1976. 487p.
- MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. *Multivariate analysis*. London: Academic Press, 1979. 521p.
- MONTGOMERY, D. C.; PECK, E. A. *Introduction to linear regression analysis*. Nova York: Wiley, 1982. 504p.
- PEREIRA, C. P. *Estimativas e previsões preliminares de colheitas de cacau por amostragem*. Piracicaba: ESALQ-USP, 1972. 47p. Tese de Mestrado.
- RAU, A. A. *Report on FAO regional project on cocoa statistics for evolving suitable technique to forecast the cocoa crop production in West Africa*. Roma: FAO, 1965. 30p.
- RIPAILLES, C.; ROSSION, J. *Prévision de récolte pour le cacao*. *Café Cacao Thé*, Paris, v.12, n.2, p.103-113, 1968.
- TITTERINGTON, D. M. Estimation of correlation coefficients by ellipsoidal trimming. *Applied Statistics*, London, v.27, p.227-234, 1978.
- WATERS, H. B.; HUNTER, T. Measurement of rate of development of cacao pods. *Bulletin of the Department of Agriculture, Gold Coast, Acra*, v.16, p.121-128, 1928.
- WEISBERG, S. *Applied linear regression*. 2.ed. Nova York: Wiley, 1985. 324p.
- ZARONI, M. M. H. *Previsão de safras de cacau no Estado da Bahia com base no número de frutos do primeiro estágio de frutificação*. Piracicaba: ESALQ-USP, 1981. 157p. Tese de Mestrado.