

---

# PREVISÃO DE PRODUÇÃO AGRÍCOLA BASEADA EM REGRAS LINGUÍSTICAS E LÓGICA FUZZY

José M. Vargens\*  
vargens@inf.puc-rio.br

Ricardo Tanscheit\*  
ricardo@ele.puc-rio.br

Marley M.B.R. Vellasco\*  
marley@ele.puc-rio.br

\*Departamento de Engenharia Elétrica, PUC-Rio, Caixa Postal 38.063 22.452-970 Rio de Janeiro, RJ, Brasil

---

## ABSTRACT

A good precision in forecasting cocoa production has become significant due to the decreasing profit margin cocoa producers have experienced in the past years. This has been caused by increasing production costs and by a reduction in international prices. This work presents a fuzzy logic system that allows farmers to forecast their cocoa production with higher precision and simplicity. The method is less costly than the one currently employed and has been devised to be part of a system that should help farm management. Farmers do not need any specific technical knowledge of the model used and results obtained so far through simulation are encouraging.

**KEYWORDS:** Cocoa production; linguistic rules; fuzzy logic.

## RESUMO

Uma boa precisão na estimativa da produção passou a ser significativa na cacauicultura devido à forte diminuição da margem de comercialização nos últimos anos, provocada pela combinação do aumento dos custos da produção e pela queda da cotação internacional do produto final. Este trabalho apresenta um sistema de lógica fuzzy que permite aos fazendeiros fazer uma previsão da

sua produção de cacau com mais precisão e simplicidade. O método proposto é de baixo custo quando comparado ao método atualmente empregado e foi projetado para ser parte de um sistema que auxilie no gerenciamento da fazenda. Sua utilização não exige do produtor conhecimento técnico específico e os resultados obtidos até o momento através de simulações são animadores.

**PALAVRAS-CHAVE:** Produção de cacau; regras linguísticas, lógica fuzzy.

## 1 INTRODUÇÃO

Nos últimos dez anos a cacauicultura tem passado por uma mudança radical na sua forma de produção. A dizimação dos pés de cacau pela doença chamada Vassoura de Bruxa reduziu a produção a praticamente zero (Annual Report, 1996/1997). Ao longo deste tempo foram desenvolvidas pesquisas para resolver o problema e, atualmente, a CEPLAC propõe uma nova técnica de produção que permite a produção comercial uma vez que reduz sensivelmente os efeitos da Vassoura de Bruxa (Cocoa Newsletter, 1999). Contudo, esta técnica traz consigo um aumento muito grande do custo da produção em relação ao método tradicional empregado anteriormente.

A cacauicultura é uma produção agrícola de exportação e por isso seu preço é definido basicamente por sua cotação internacional. O preço do produto final varia dia a dia de acordo com a sua cotação nas bolsas internacionais de 'commodities' e depende do estoque in-

Artigo submetido em 20/12/00

1a. Revisão em 15/10/01; 2a. Revisão em 04/02/02

Aceito sob recomendação do Ed. Assoc. Prof. Fernando Gomide

ternacional do produto. Atualmente a cotação internacional do cacau tem estado num patamar muito baixo considerando-se o novo custo de produção da tecnologia proposta pela CEPLAC. Consequentemente, a margem do cacauicultor está reduzida de tal forma que pequenos erros na conjugação produção versus comercialização causam prejuízo.

Como a crise se estende há dez anos, os produtores se encontram descapitalizados e a forma para se manterem no negócio é o comprometimento da produção futura, seja através de penhor ao sistema financeiro seja através da venda a futuro ao exportador. Neste contexto, saber estimar com precisão a sua produção passou a ser fator fundamental para o cacauicultor. Até hoje as formas consistentes para a previsão da produção têm sido a média da produção histórica do fazendeiro e o cálculo estatístico baseado numa amostra da plantação. A utilização da série histórica da produção tornou-se praticamente sem valor, uma vez que a Vassoura de Bruxa dizimou as plantações anteriormente existentes e as novas plantas têm características genéticas de produtividade diferentes das antigas. A previsão por amostragem teve seus parâmetros revistos pela CEPLAC para se ajustar às características de produtividade da nova plantação e poder continuar a ser utilizada. Desta forma ela é utilizada em combinação com o cálculo do potencial de produção da fazenda, baseado no número de plantas, seu tipo e área plantada, para auxiliar na tomada das decisões estratégicas do cacauicultor. No entanto, este método apresenta alguns inconvenientes que comprometem a sua eficácia. O primeiro é o fato de seus parâmetros terem sido obtidos somente no laboratório da CEPLAC, não representando, necessariamente, com exatidão as características das novas plantas no campo e acrescentando um custo adicional relacionado à coleta antecipada de alguns frutos. O segundo refere-se ao intervalo mínimo entre amostragens, que não deve ser menor do que três meses, fazendo com que fatores que afetem a produção não tenham seus efeitos estimados antes do final do período considerado. O terceiro é que esta previsão reflete, na verdade, a consequência das ações tomadas anteriormente pelo produtor, não permitindo um acompanhamento mais amigável nem a simulação antecipada de cenários.

Neste trabalho é proposto um método de previsão que complementa a previsão por amostragem e que dá origem a um sistema fuzzy que permite ao fazendeiro avaliar o impacto das suas decisões de tal forma que ele consiga maximizar sua margem de lucro. Na elaboração do sistema levou-se em consideração que a solução proposta não acarretasse aumento de custos e nem obrigasse o fazendeiro a adquirir conhecimentos específicos

para este processo além dos já dominados por ele para a plantação de cacau. O sistema fuzzy ajusta a estimativa amostral de previsão à medida que ocorram fatos que afetem a produção, de tal modo que o produtor possa reavaliar a previsão da sua produção a qualquer instante.

## 2 DESCRIÇÃO DO PROBLEMA

O problema consiste na construção de um sistema para calcular a variação da produção, causada pela conjugação de diversos fatores que ocorrem na plantação, e que possa ser utilizado como um elemento num sistema gerencial da fazenda.

A informação de entrada é a estimativa técnica da produção, baseada em características estáticas da plantação, tais como número de pés em produção, características químicas do solo e média de produção histórica. A previsão é feita no início do ano agrícola e reavaliada trimestralmente pela estimativa amostral da plantação, de acordo com procedimento padrão da CEPLAC. A esta informação são acrescentadas as práticas agrícolas definidas pelo cacauicultor. Não se têm medidas precisas a respeito da implementação destas práticas pois ocorrem falhas humanas e perda de material no processo de produção. Por exemplo, não há certeza se no processo de aplicação de cobre para o controle de doença toda a área foi coberta ou se os trabalhadores esqueceram de efetuar a aplicação em algumas plantas. O cacauicultor tem como meta seguir o procedimento técnico padrão proposto pela CEPLAC, mas faz uso de uma medida subjetiva de quão aderente a esse padrão está a sua prática na fazenda. Neste trabalho pretende-se captar essa medida subjetiva da realidade do produtor.

Baseado nas informações acima, o sistema efetua as correções e produz a previsão da produção ajustada pelos fatores que a afetam. Esta previsão ajustada será então utilizadas para redefinir as práticas agrícolas. Estas servem então como realimentação do sistema, de tal forma que uma reavaliação seja efetuada continuamente.

No projeto do sistema alguns aspectos importantes foram considerados:

- deveria ser simples e expresso em termos familiares aos produtores, de modo a reduzir o seu esforço de adaptação ao sistema;
- deveria ser projetado com o auxílio de especialistas, de modo a produzir resultados precisos e de forma a garantir uma melhora da eficiência;

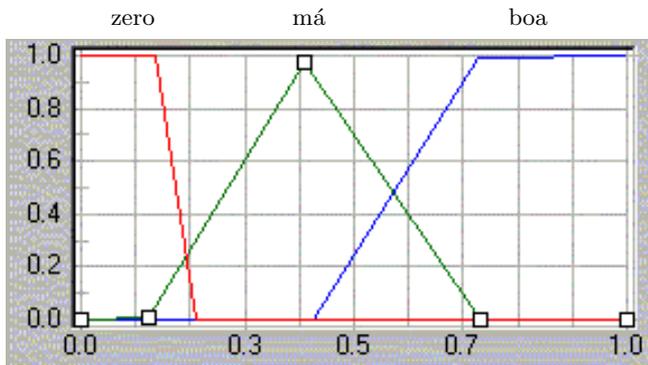


Figura 1: Conjuntos fuzzy para a variável de entrada distribuição de chuva

- deveria ser flexível o suficiente para permitir a incorporação de informações sobre novas tecnologias de produção.

Um sistema baseado em regras linguísticas traduzidas para termos matemáticos através da lógica fuzzy é adequado para atender os requisitos acima (Altrock, 1995) e por isso foi adotado.

O problema pode, então, ser resumido em como projetar um sistema fuzzy para ajustar a estimativa amostral de previsão da produção à medida que ocorram fatos que a afetem, permitindo ao produtor reavaliar essa previsão a qualquer instante. Assim, mantém-se a metodologia de previsão da CEPLAC, onde são consideradas variáveis tais como número de plantas, área plantada e condições do solo, enquanto o sistema fuzzy proposto efetua um ajuste dessa previsão levando em conta variáveis referentes ao dia a dia do processo de gestão do cacauicultor.

### 3 O SISTEMA FUZZY

#### 3.1 Definição dos Conjuntos Fuzzy

Na construção do sistema fuzzy as variáveis de entrada foram divididas em três categorias, conforme especificado abaixo. Os valores fuzzy para essas variáveis e os conjuntos fuzzy correspondentes são definidos em relação a recomendações técnicas padrão definidas pela CEPLAC. Quanto mais próximo se está de seguir essas recomendações, mais próximo se estará da extrema direita do universo de discurso (1.0). Por exemplo, a posição do conjunto fuzzy **alta** no universo para *incidência de mato* (Figura 2) indica que as recomendações técnicas estão distantes de ser cumpridas, o que implica que a produção será desfavoravelmente afetada. Quando a *incidência de*

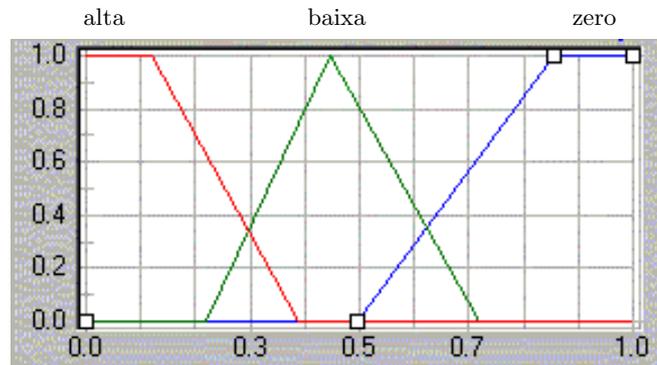


Figura 2: Conjuntos fuzzy para a variável de entrada incidência de mato

*mato* é próxima de zero (campo limpo), as recomendações são seguidas de perto e a produção não deverá ser afetada. Os formatos e suportes dos conjuntos fuzzy foram definidos num processo de tentativa e erro durante a fase de desenvolvimento do sistema. Conforme mencionado anteriormente, essas variáveis têm medida imprecisa no dia a dia do produtor e captar suas formas de expressão constitui-se em um processo subjetivo de experimentação em conjunto com os consultores. As três categorias para as variáveis de entrada são:

#### a) fatores que definem a produção:

- *distribuição de chuva*: três valores linguísticos são associados a essa variável, conforme especificado pelos conjuntos fuzzy **zero**, **má** e **boa**, cujas funções de pertinência são mostradas na Figura 1.
- *incidência de mato*: três valores linguísticos - **alta**, **baixa** e **zero**, de acordo com a Figura 2.

#### b) fatores que aumentam a produção

- *adubação*
- *calagem*

Para ambas as variáveis os valores linguísticos são **zero**, **média** e **correta**, conforme mostrado na Figura 3.

#### c) fatores que mantêm a produção:

- *controle de praga*, considerando que haja *incidência de praga*
- *controle de doença*, considerando que haja *incidência de doença*

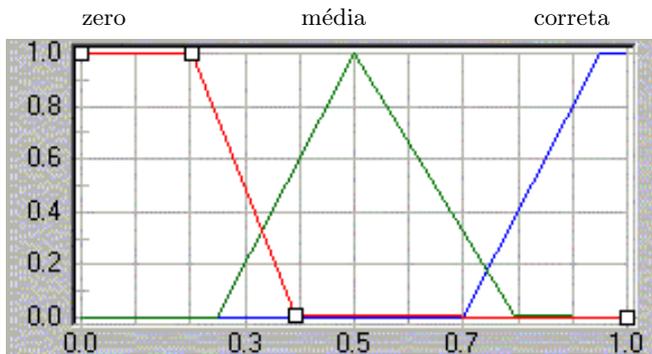


Figura 3: Conjuntos fuzzy para as variáveis de entrada adubação e calagem

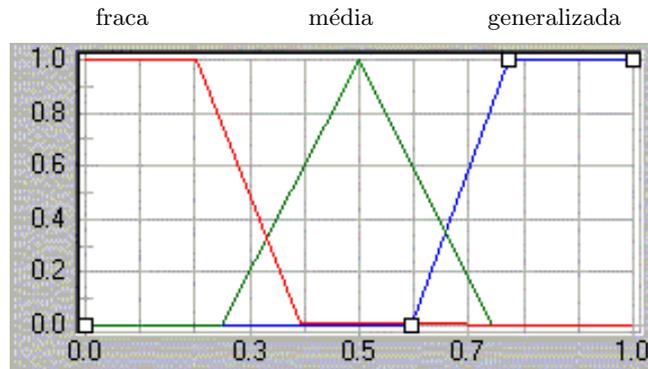


Figura 5: Conjuntos fuzzy para as variáveis de entrada incidência de praga e incidência de doença

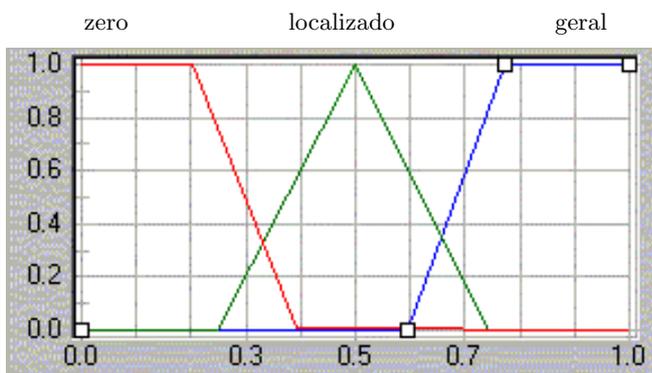


Figura 4: Conjuntos fuzzy para as variáveis de entrada controle de praga e controle de doença

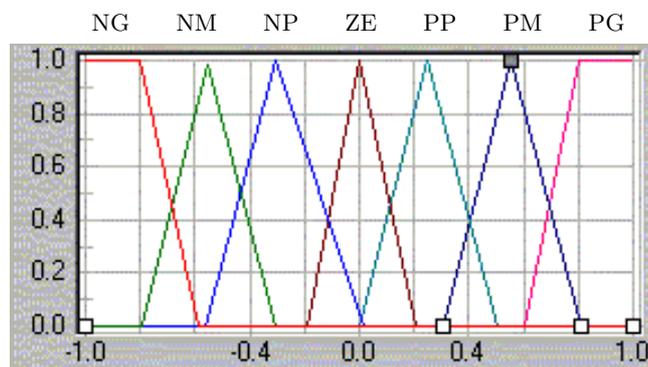


Figura 6: Conjuntos fuzzy para a variável de saída variação da produção

Para essas variáveis, as distribuições dos conjuntos nos universos, suas formas e domínios são mostrados nas Figuras 4 e 5. Nos casos de *controle de praga* e *controle de doença* as denominações são: **zero**, **localizado** e **geral** onde **geral** significa o controle feito sobre toda a plantação; para *incidência de praga* e *incidência de doença* são: **fraca**, **média** e **generalizada**.

É importante observar que a recomendação da CEPLAC para estes casos é um balanceamento da área a ser controlada tendo em vista a área onde há a incidência da doença ou praga. Como o universo de discurso faz o cortejamento da realidade em relação à recomendação da CEPLAC (extremo direito do universo – 1.0) devem ser consideradas ambas as variáveis na construção das regras.

Os valores linguísticos para a saída do sistema (*variação da produção*) são definidos pelos conjuntos fuzzy mostrados na Figura 6. As denominações são: **negativa grande** (NG), **negativa média** (NM), **negativa**

**pequena** (NP), **zero** (ZE), **positiva pequena** (PP), **positiva média** (PM), **positiva grande** (PG).

### 3.2 Definição das Regras

A especificação do conjunto de regras foi baseada nas opiniões de especialistas sobre a influência que cada fator tem sobre a produção como um todo. Trabalhamos como consultores neste projeto três engenheiros agrônomos e um técnico agrícola, todos com mais de dez anos de experiência em cacauicultura sendo dois deles engenheiros da CEPLAC e outros dois da empresa PLANAGRO, que atua na região em projetos agrícolas.

O primeiro passo foi obter um consenso entre os consultores sobre o impacto das entradas na produção. Suas opiniões, consolidadas na Tabela 1, ajudam a estabelecer os consequentes para diferentes situações, definidas pelos antecedentes das regras. Pode ser visto, por exemplo, que uma alta *incidência de mato* pode ter um

Tabela 1: Opiniões de especialistas sobre o impacto de cada fator na *variação da produção*

Variação da produção	Distribuição de chuva	Incidência de mato	Adubação	Calagem	Não Fazer Controle de praga	Não Fazer Controle de doença
Maior Variação Negativa	70%	30%			10%	60%
Var. Negativa Intermediária	35%	15%			5%	30%
Pouca Variação Negativa		10%			2%	10%
Nenhuma Variação	0%	0%	0%	0%	0%	0%
Pouca Variação Positiva			30%	5%		
Var. Positiva Intermediária			70%	10%		
Maior Variação Positiva			150%	15%		

impacto de no máximo 30% na redução na produção. Assim, ao se escrever as regras para esse fator, deve-se ter em mente que a variação na produção nunca será extremamente afetada. A Tabela 1 também mostra que o uso de fertilizantes ou calagem somente afetará favoravelmente a produção. Se o controle de pragas e doenças seguir as recomendações técnicas, a produção deverá se manter dentro dos valores previstos inicialmente.

Alternativamente, os conseqüentes das regras poderiam ser estabelecidos sem levar em conta os diferentes impactos que as variáveis têm sobre a produção. Nesse caso, continuar-se-ia a usar a Tabela 1 como guia, mas os resultados obtidos para cada fator seriam então ponderados numa composição final. Contudo, o primeiro procedimento descrito acima foi o preferido dos especialistas consultados.

O grande número de variáveis de entrada (8) torna praticamente impossível para os especialistas definir um conjunto de regras adequado e coerente. Assim, a definição das regras é efetuada a partir de uma avaliação do impacto individual de cada fator. Esta estratégia é baseada no conceito de *regra incompleta* (Mendel, 1995), onde variáveis que não aparecem em uma dada regra são tomadas como tendo grau de pertinência igual a 1. Neste caso utiliza-se o operador *min* para computar o nível de disparo do antecedente da regra.

Foi também efetuada a adequação dos valores linguísticos da variável de saída à expectativa de variação da produção determinada pelas variáveis de entrada expres-

sas na Tabela 1; o resultado está apresentado na Tabela 2. As regras foram então definidas de modo que seus conseqüentes fossem coerentes com o impacto causado na produção. Por exemplo, comparando-se as regras (I) e (II) abaixo, observa-se que, apesar de ambas se referirem ao valor extremo do seu universo de discurso, seus conseqüentes são diferentes porque os reflexos na produção são diferentes. Assim, a regra (I), que considera a *distribuição de chuva*, tem como seu maior valor negativo 70% (conforme a Tabela 1), implicando em um conseqüente NG (conforme a Tabela 2).

Tabela 2: Guia de auxílio para adequar as opiniões de especialistas *ao universo de discurso da saída*

Valores linguísticos da saída	Variação da produção
NG	-60%
NM	-30%
NP	-10%
ZE	0%
PP	10%
PM	30%
PG	60%

A regra (II), que se refere à *incidência de mato*, tem 30% como seu maior valor negativo na Tabela 1 o que se traduz no conseqüente NM.

As regras a seguir são agrupadas segundo os fatores já especificados em (a), (b) e (c):

Se *distribuição de chuva* é boa  
então *variação da produção* é ZE

Se *distribuição de chuva* é má  
então *variação da produção* é NM

Se *distribuição de chuva* é zero  
então *variação da produção* é NG (I)

Se *incidência de mato* é zero  
então *variação da produção* é ZE

Se *incidência de mato* é baixa  
então *variação da produção* é NP

Se *incidência de mato* é alta  
então *variação da produção* é NM (II)

Se *adubação* é zero  
então *variação na produção* é ZE

Se *adubação* é correta  
então *variação na produção* é PG

Se *adubação* é média  
então *variação na produção* é PM

Se *calagem* é zero  
então *variação da produção* é ZE

Se *calagem* é correta  
então *variação da produção* é PM

Se *calagem* é média  
então *variação da produção* é PP

Nas regras abaixo *incidência de praga*, *incidência de doença*, *controle de praga* e *controle de doença* são denotados por *i.p.*, *i.d.*, *c.p.* e *c.d.*:

Se *i.p.* é fraca e *c.p.* é zero  
então *variação da produção* é NP

Se *i.p.* é fraca e *c.p.* é localizado  
então *variação da produção* é ZE

Se *i.p.* é fraca e *c.p.* é geral  
então *variação da produção* é ZE

Se *i.p.* é média e *c.p.* é zero  
então *variação da produção* é NM

Se *i.p.* é média e *c.p.* é localizado  
então *variação da produção* é NP

Se *i.p.* é média e *c.p.* é geral  
então *variação da produção* é ZE

Se *i.p.* é generalizada e *c.p.* é zero  
então *variação da produção* é NM

Se *i.p.* é generalizada e *c.p.* é localizado  
então *variação da produção* é NP

Se *i.p.* é generalizada e *c.p.* é geral  
então *variação da produção* é ZE

Se *i.d.* é fraca e *c.d.* é zero  
então *variação da produção* é NM

Se *i.d.* é fraca e *c.d.* é localizado  
então *variação da produção* é NP

Se *i.d.* é fraca e *c.d.* é geral  
então *variação da produção* é ZE

Se *i.d.* é média e *c.d.* é zero  
então *variação da produção* é NM

Se *i.d.* é média e *c.d.* é localizado  
então *variação da produção* é NP

Se *i.d.* é média e *c.d.* é geral  
então *variação da produção* é ZE

Se *i.d.* é generalizada e *c.d.* é zero  
então *variação da produção* é NG

Se *i.d.* é generalizada e *c.d.* é localizado  
então *variação da produção* é NM

Se *i.d.* é generalizada e *c.d.* é geral  
então *variação da produção* é ZE

### 3.3 Inferência e Defuzzificação

O sistema utiliza o operador de implicação *min* de Mamdani (Driankov et al., 1993) e o método de composição da soma limitada (Cox, 1994), o qual foi escolhido devido à natureza das regras. Como cada uma define um incremento ou decremento na produção de cacau, é desejável que todas as regras ativadas tenham influência no resultado final. O método da soma limitada calcula a saída fuzzy através da soma de todos os valores de pertinência resultantes para um conjunto fuzzy de saída específico, limitando essa soma a 1. Quando comparado ao operador *max*, que considera apenas o valor de pertinência máximo, o método da soma limitada mostra ser mais adequado para esta aplicação específica.

O valor 'crisp' para *variação na produção* é obtido através do método de defuzzificação do Centróide (Driankov et al., 1993). O intervalo [-1,1] para o universo de discurso da variável de saída (cf. Figura 6) indica a variação real obtida pelo sistema fuzzy. O valor -1,0 indica produção nenhuma, enquanto que +1,0 indica o dobro de produção. Esta variação é então aplicada à previsão amostral calculada no início do período, resultando na nova previsão.

## 4 ESTUDO DE CASOS

O sistema fuzzy proposto foi implementado com o auxílio do Fuzzy Toolbox<sup>©</sup> do Matlab<sup>©</sup> e testado com dife-

Tabela 3: Cenários e resultados para estudo de casos

Variáveis de Entrada	Caso 1	Caso 2	Caso 3	Caso 4
<i>Distribuição de chuva</i>	0.8	0.2	0.2	0.6
<i>Incidência de mato</i>	0.8	0.2	0.8	0.5
<i>Adubação</i>	0.8	0.2	0.8	0.75
<i>Calagem</i>	0.8	0.2	0.2	0.3
<i>Controle de praga</i>	0.8	0.2	0.1	0.3
<i>Incidência de praga</i>	0.2	0.8	0.05	0.35
<i>Controle de doença</i>	0.8	0.2	0.2	0.7
<i>Incidência de doença</i>	0.1	0.8	0.05	0.7
<b><i>Variação da produção</i></b>	<b>1.5165</b>	<b>0.4299</b>	<b>0.9481</b>	<b>0.9401</b>

rentes cenários, com o objetivo de se avaliar sua precisão e consistência. Quatro destes cenários e os resultados obtidos pelo sistema fuzzy para cada um são apresentados na Tabela 3 acima.

A leitura dos resultados deve ser feita como indicado nas análises de cada caso apresentadas a seguir. Os Casos 1 e 2 foram considerados extremos para o estudo, uma vez que seria irrealista a aplicação precisa de toda a recomendação técnica; por outro lado, o abandono total da plantação não justificaria o uso do sistema. Os Casos 3 e 4 refletem situações mais próximas do cotidiano dos produtores.

O Caso 1 apresenta um cenário onde as recomendações técnicas são seguidas corretamente, resultando num aumento de produção superior a 50%. O Caso 2, por outro lado, corresponde a um cenário onde todas as recomendações são negligenciadas; o resultado é um decréscimo na produção da ordem de 60%. O Caso 3 reflete uma situação em que uma má *distribuição de chuva*, que pode afetar desfavoravelmente a produção, é compensada mantendo-se a *incidência de mato* baixa e usando uma quantidade próxima da correta de *adubação*. Finalmente, o Caso 4 examina um cenário onde: variáveis que definem a produção (*distribuição de chuva* e *incidência de mato*) estão na média, variáveis que aumentam a produção compensam-se mutuamente (*adubação* e *calagem*) e variáveis que mantêm a produção estão razoavelmente sob controle. A produção resultante, como

esperado, quase não é afetada.

De acordo com os especialistas consultados, todos os resultados situaram-se dentro das expectativas para os cenários considerados, o que demonstra que o sistema é suficientemente consistente para ser empregado em situações reais. De qualquer forma, caso necessário, é fácil efetuar ajustes nos conjuntos fuzzy e na base de regras.

## 5 CONCLUSÕES

O objetivo principal deste trabalho foi fornecer a fazendeiros meios razoavelmente precisos e de baixo custo para prever variações na produção de cacau, de modo a auxiliá-los na tomada de decisões de caráter estratégico. Para atingir esse objetivo foi desenvolvido um sistema baseado na lógica fuzzy que calcula as variações na produção a partir de informações sobre fatores que podem afetá-la. O sistema proposto foi avaliado favoravelmente por especialistas e o próximo passo será testá-lo em condições reais, de modo a confirmar sua aplicabilidade.

O procedimento descrito neste trabalho pode ser executado no equipamento do produtor, permitindo uma atualização de sua previsão de produção cada vez que surgir alguma nova informação relevante. Além disso, o programa desenvolvido pode ser utilizado para simulações de casos, proporcionando uma avaliação do impacto de diversas alternativas de uso dos recursos financeiro e servindo, desta forma, como um instrumento de apoio à decisão para o produtor.

## REFERÊNCIAS

- Annual Report (1996/1997), International Cocoa Organization.
- Cocoa Newsletter (1999), No. 17, International Cocoa Organization.
- Altrock, C. (1995). Fuzzy Logic & Neuro-Fuzzy Applications Explained, Prentice-Hall.
- Mendel, J.M. (1995). Fuzzy Logic Systems for Engineering: a Tutorial, Proc. IEEE, **3**: 345-376.
- Driankov, D.; Hellendorn, H.; Rheinfrank, M. (1993). An Introduction to Fuzzy Control, Springer-Verlag.
- Cox, E. (1994). The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, and Maintaining Fuzzy Systems, A.P Professional.