

83

CONSISTÊNCIA DAS PREVISÕES DE SAFRA DE ALGODÃO, MILHO E SOJA NOS ES
TADOS UNIDOS, NO BRASIL E EM SÃO PAULO

Sebastião Nogueira Junior
Flavio Conde de Carvalho
Alfredo Tsunehiro

Governo do Estado de São Paulo
Secretaria da Agricultura
Coordenadoria Sócio-Econômica

Instituto de Economia Agrícola



CONSISTÊNCIA DAS PREVISÕES DE SAFRA DE ALGODÃO, MILHO E SOJA, NOS ESTADOS UNI
DOS, NO BRASIL E EM SÃO PAULO

Sebastião Nogueira Junior
Flavio Condé de Carvalho
Alfredo Tsunechiro



ÍNDICE

1 - INTRODUÇÃO	1
2 - SISTEMAS DE LEVANTAMENTO	2
2.1 - Estados Unidos	2
2.2 - Brasil	3
2.3 - São Paulo	3
3 - CALENDÁRIO AGRÍCOLA	5
3.1 - Estados Unidos	5
3.2 - Brasil	6
3.3 - São Paulo	7
4 - ANTECEDENTES	7
5 - METODOLOGIA	10
6 - RESULTADOS E DISCUSSÃO	12
6.1 - Algodão	13
6.2 - Milho	19
6.3 - Soja	22
7 - CONCLUSÕES	28
LITERATURA CITADA	31
RESUMO	34
SUMMARY	35

CONSISTÊNCIA DAS PREVISÕES DE SAFRA DE ALGODÃO, MILHO E SOJA NOS ESTADOS UNIDOS, NO BRASIL E EM SÃO PAULO (1)

Sebastião Nogueira Junior
Flavio Condê de Carvalho
Alfredo Tsunechiro

1 - INTRODUÇÃO

Previsões de safra constituem-se em importantes instrumentos para direcionar a comercialização agrícola, seja do ponto de vista do produtor, do consumidor, como também dos órgãos governamentais para a formulação de políticas de abastecimento.

A divulgação de estatísticas fidedignas de previsão de safra em tempo hábil permite ajustamentos de ordem sócio-econômica, sobretudo no que diz respeito às projeções de preço, possibilitando diminuição de riscos de perdas para os produtores e de gastos para os consumidores.

Embora importante, o campo de previsões de safra nos moldes atuais nos vários estados do Brasil, incluindo São Paulo, ainda apresenta falhas e, conseqüentemente, tem a fidedignidade de seus dados contestada, dada a predominância de informações subjetivas para complementação dos dados objetivos.

Nos Estados Unidos, por sua vez, os procedimentos objetivos, aliados a um rigoroso controle de qualidade dos dados obtidos, conferem elevado grau de credibilidade às estimativas periodicamente divulgadas, de ampla aceitação e utilização por entidades públicas e privadas, inclusive no exterior.

Resultados de previsão de safra de algodão, milho e soja são aguardados com grande expectativa, dada a importância desses produtos na alimentação humana e animal, e ainda porque deles advêm uma gama de derivados de grande utilização industrial. A escolha recaiu sobre esses produtos em função de sua importância e pela disponibilidade de séries temporais adequadas nas

(1) Os autores agradecem à Pesquisadora Científica Maria de Fátima Packer pela colaboração na escolha do método estatístico, à Bibliotecária Agurí Sawatani Negri pelo preparo das referências bibliográficas das versões preliminares deste estudo que foram apresentadas em eventos específicos (18), (19) e (20), e à escriturária Sonia Regina de Santana.

três instituições analisadas.

O estudo tem por objetivo proceder a uma descrição e análise da consistência dos levantamentos de previsão de safras de algodão, milho e soja, realizados por três instituições: Departamento de Agricultura dos Estados Unidos (USDA), Fundação Instituto Brasileiro de Geografia e Estatística (IBGE) e Instituto de Economia Agrícola (IEA).

2 - SISTEMAS DE LEVANTAMENTO

2.1 - Estados Unidos

Há mais de um século o Departamento de Agricultura dos Estados Unidos (USDA) por intermédio do Statistical Reporting Service (SRS), coleta e divulga informações agropecuárias. Cerca de 150 culturas e 50 atividades de origem animal são incluídas nos levantamentos, bem como sementes, rações, fertilizantes, mão-de-obra e preços pagos e recebidos pelos agricultores. Dessa forma cerca de 650 relatórios são divulgados anualmente (10).

Os relatórios de Safras (Crop Reports) incluem estimativas de intenção de plantio, área plantada, área colhida, produção, distribuição e estoques. Previsões de produtividade e produção são feitas mensalmente durante o ciclo vegetativo das culturas, totalizando seis levantamentos. São fundamentais em contagens e medidas realizadas em propriedades escolhidas por amostragem, nas condições de desenvolvimento da planta e nos rendimentos presumíveis por área.

Os questionários enviados aos informantes pelo correio constituem, ainda, o principal método de coleta de dados. Cerca de 25 mil questionários preenchidos retornam ao USDA e são agrupados por distritos, e uma vez tabulados e calculados, são usados para ponderação das médias estaduais ou totais, chamadas "indicações". Estas médias, de certa forma, são viesadas porque a listagem de informantes pode apresentar falhas e nem todos os questionários retornam. Entretanto, o viés é reduzido pela depuração dos dados via utilização de gráficos de regressão, quando os dados amostrais enviados são cotejados com os dados das estimativas conclusivas.

Previsões mensais e estimativas finais para várias culturas, entre as quais as do algodão, do milho e da soja, são também baseadas em campos de observação selecionados por amostragem, visando dar indicações sobre o provável rendimento da cultura quando o produto for colhido. Isto possibilita rever dados mensais, se necessário, e serve para a estimativa final. Para preservar a acurácia das informações, são feitas amostragens

adicionais, em estados produtores não principais. Estes dados são enviados diretamente a Crop Reporting Board (CRB), servindo para checar as informações referentes aos estados tradicionais produtores (31).

A importância e confiabilidade dos dados da previsão estadunidense requer um rigoroso esquema para sua divulgação. Até a hora do anúncio, pontualmente às 15 horas do dia pré-determinado do mês, apenas os quatro membros da CRB têm conhecimento prévio das cifras. Esse procedimento - precisão dos dados e pontualidade da divulgação - confere grande credibilidade às estatísticas do USDA, utilizadas em todo o mundo (14).

2.2 - Brasil.

O IBGE realiza mensalmente, desde 1973, o Levantamento Sistemático da Produção Agrícola (LSPA), que é uma pesquisa de previsão e acompanhamento das safras de 30 produtos agrícolas, dentre os quais algodão, milho e soja (2). O LSPA objetiva a elaboração de estimativas sobre área, produção e rendimento médio, desde a fase de intenção de plantio até o final da colheita, abrangendo todo o ciclo vegetativo de cada cultura.

A sistemática de coleta prevê o registro, em questionário próprio, dos dados obtidos junto aos informantes (entidades públicas e privadas) existentes nos municípios (13).

Além dos dados referentes à produção, área e rendimento, são coletadas informações sobre condições climáticas, espaçamento modal e preço pago ao produtor. Para o milho são feitos nove levantamentos mensais a nível nacional, desde a fase do plantio até o final da colheita; e para o algodão e a soja, sete levantamentos.

O ano civil é adotado como período de referência pelo fato da pesquisa constituir subsídio para a elaboração do Produto Interno Bruto (PIB) do setor agrícola, objetivando ainda a comparação com outros levantamentos agropecuários contínuos e censitários.

2.3 - São Paulo

As previsões e estimativas de safras no Estado de São Paulo são elaboradas pelo Instituto de Economia Agrícola, a partir de dados coletados pela Coordenadoria de Assistência Técnica Integral (CATI), mediante dois tipos de levantamentos: subjetivo e objetivo (por amostragem). No caso dos

(2) Para a rede de coleta a pesquisa é bimestral.

levantamentos subjetivos todos os municípios do Estado de São Paulo, através das Casas da Agricultura pertencentes à CATI, recebem periodicamente questionários contendo perguntas sobre a produção agrícola, que são preenchidos pelos agrônomos regionais da CATI e devolvidos ao IEA, onde os dados são depurados e tabulados.

Quanto aos levantamentos objetivos, para cada unidade de amostra (imóvel rural) são preenchidos questionários por enumeradores da CATI em entrevista com o produtor, administrador ou outra pessoa em condições de informar sobre o imóvel sorteado. Devolvidos os questionários para o IEA, seus dados são depurados e processados.

As principais diferenças entre os levantamentos subjetivos e objetivos do ponto de vista da execução do trabalho e do ponto de vista estatístico podem ser vistos com maiores detalhes em PINO; CAMARGO; JIMENEZ OSSIO (25).

Desde o início da década de 50 tem sido utilizada a técnica de amostragem para obtenção de estimativas a nível regional e estadual. A partir de 1973, com o objetivo de aprimoramento destas estatísticas, estabeleceu-se a meta de se obter estimativas para o Estado de São Paulo como um todo, com erros padrões que não ultrapassassem $\pm 5\%$ do valor das citadas estimativas.

Conforme CAMPOS & PIVA (4) o tamanho da amostra para procedimento da coleta de dados nesta segunda fase, foi determinado a partir do cadastro do INCRA referente ao ano de 1972, com um total de 257.955 propriedades rurais, distribuídas pelas 10 Divisões Regionais Agrícolas (DIRAs) do Estado, considerando doze estratos, de acordo com o tamanho dos imóveis com área superior a 3 hectares.

Os dados para o dimensionamento foram obtidos através de dois levantamentos, realizados em janeiro e março de 1974, com uma amostra de 6.996 propriedades (3).

A determinação do tamanho da amostra foi baseada na partilha de Neyman, a partir de cadastro único para as oito culturas mais importantes do Estado (algodão, amendoim, arroz, café, cana-de-açúcar, feijão, milho e soja). O número de levantamentos durante o ano agrícola também variou no tempo, tendo sido de três no passado e de cinco atualmente, realizados em setembro, novembro, fevereiro, abril e junho, sendo que o levantamento de junho corresponde ao dado final do ano agrícola ou safra considerado.

(3) A partir de setembro de 1977 a amostra foi reformulada e passou a ter 5.646 questionários e atualmente se situa em 3.622 imóveis rurais.

Em levantamentos por amostragem surge o problema, quase inevitável, de falta de resposta. Quando, por quaisquer razões, o percentual de falta de respostas for elevado, o levantamento pode ter seus resultados invalidados. Esse parece não ser o caso de São Paulo, porquanto a percentagem média de questionários respondidos tem sido elevada, com o menor índice de retorno situando-se em 82% na DIRA de São Paulo no período 1975-79 (24).

Aqueles mesmos autores também se propuseram a detectar a magnitude de dos erros não amostrais a partir do levantamento objetivo de abril de 1978, estabelecendo tipos, diferenças regionais e devidas aos enumeradores (23). O número encontrado (1,8 erro, em média) por questionário, embora pequeno, é suficiente para causar aumento na variância das estimativas, por que este aumento depende mais da magnitude do erro que do número de erros. Os tipos de erro mais frequentes referem-se a respostas incompletas e falta de somatório.

Foi detectado por este trabalho diferença significativa entre as DIRAs, quanto a percentagem de erro, tendo como prováveis causas a quantidade e qualificação dos enumeradores, infra-estrutura e recursos materiais; atividades agrícolas e infra-estrutura da região; extensão territorial e entrosamento IEA-CATI.

3 - CALENDÁRIO AGRÍCOLA

Na descrição dos sistemas de levantamento não tem sido mencionada a distribuição temporal das fases vegetativas das culturas analisadas. Devido à diversidade climática existente em um País ou mesmo em um Estado, em um determinado mês coexistem culturas em diferentes fases do processo de crescimento. Além disso, há algumas fases desse processo que podem proporcionar melhores indicações da produção ou produtividade final da cultura. Esses aspectos não têm sido discutidos com profundidade na literatura consultada.

3.1 - Estados Unidos

Nos Estados Unidos, o início do plantio do algodão ocorre em fevereiro, atingindo o pico em maio e estendendo-se até junho. A colheita começa em agosto, tem seu auge em outubro, encerrando-se em dezembro.

O plantio de milho inicia-se em março, atinge o pico em maio e vai até junho. A colheita inicia-se em agosto, com maior concentração em outubro e vai até fins de novembro, com pequena parcela ainda em colheita em dezembro.

A cultura da soja estadunidense é plantada de abril a junho, com a maior parte concentrando-se em maio. A colheita começa em setembro, acentuando-se em outubro e avança até dezembro (26).

Na falta de dados quantitativos sobre a distribuição do plantio e colheita nos Estados Unidos, optou-se pela utilização da distribuição da comercialização ao nível de produtores agrícolas durante o ano.

De acordo com o USDA (11) a maior parte da comercialização do algodão estadunidense na safra 1978-79 ocorreu no período outubro-janeiro com 68,6% do total; soja de outubro a fevereiro com 63,3%; e milho de outubro a abril com 62,9%.

3.2 - Brasil

No Brasil, o calendário agrícola apresenta-se de maneira inversa ao dos Estados Unidos e de todo o Hemisfério Norte.

De modo geral, os plantios das culturas analisadas cujas produções se concentram na Região Centro-Sul do Brasil (4) têm início em setembro, atingem o pico em outubro e avançam até novembro (algodão) e dezembro (milho), enquanto o da soja inicia-se em outubro, com maior concentração em novembro e avança até dezembro.

As colheitas por sua vez iniciam com algodão e milho em fevereiro, soja em março, atingem o auge em março-abril (algodão e soja) e abril-maio (milho) e estendem-se até junho (algodão e soja) e julho (milho) (12).

Em termos de distribuição quantitativa, de acordo com o CENSO AGROPECUÁRIO do Brasil de 1980 (6), a maior parte da área cultivada com algodão herbáceo teve seu plantio realizado no período setembro-outubro (61%), cifra relativa à Região Centro-Sul e 24% de janeiro a março (Região Norte-Nordeste). A colheita ocorreu praticamente de fevereiro a maio (84%).

Para milho, a grande concentração do plantio (81%) ocorreu em agosto a dezembro, e a quantidade colhida foi centralizada de março a julho (89%). A colheita foi quase totalmente realizada (89%) entre fevereiro e maio.

A soja, lavoura quase exclusiva da Região Centro-Sul teve, seu plantio totalmente concentrado de setembro a dezembro e a colheita de fevereiro a maio.

Dada a concentração da produção na Região Centro-Sul, para efeito de tomada de decisão as previsões de safra para algodão efetuadas em julho

(4) No caso do algodão herbáceo, o Centro-Sul responde por mais de 90% na produção brasileira; milho por 92% e soja por 99,5%.

lho e agosto; para milho de agosto a dezembro; e para soja em junho e julho não se constituem em parâmetros relevantes. Sua não exclusão da análise se justifica pela necessidade de aferição do procedimento dos levantamentos.

3.3 - São Paulo

O plantio e colheita de algodão, milho e soja no Estado de São Paulo, praticamente tem a mesma distribuição que a verificada para a Região-Sul.

O algodão tem seu plantio concentrado em setembro e outubro; milho e soja em outubro e novembro. A colheita por sua vez tem a seguinte distribuição: algodão (março a maio); milho (abril a junho) e soja (março a abril) (21).

Quanto à distribuição relativa, de acordo com o CENSO AGROPECUÁRIO de São Paulo de 1980 (7), pode-se observar que o algodão teve praticamente todo o plantio (98%) realizado de setembro a novembro e a colheita, também 98%, de fevereiro a maio.

O milho, por sua vez, com uma distribuição temporal mais extensa, teve 92% do plantio efetivado de agosto a novembro, sendo o mesmo raciocínio válido para a colheita: 92% ocorreu em 5 meses, ou seja, de fevereiro a junho.

Finalmente, a soja mostra concentração mais forte tanto na semeadura quanto na quantidade colhida: respectivamente 96% em outubro-novembro e 95% de fevereiro a abril.

4 - ANTECEDENTES

No Brasil, trabalhos que versam sobre avaliação de previsões de safra, na maioria utilizam estatísticas divulgadas pelo Instituto de Economia Agrícola (IEA) ajustando os dados via regressões lineares simples e foram realizados na sua quase totalidade por pesquisadores dessa instituição.

BRANDT; ARAUJO; SERRANO (2) realizaram pesquisa sobre a precisão e fidedignidade das previsões de safra de algodão no período de 1947/48 a 1963/64, a partir de dados levantados pelo IEA. Seus principais resultados podem ser assim formulados:

- a) na terceira previsão de safra, em junho, existe uma tendência para subestimar as safras menores que 600 mil toneladas e para superestimar as maiores que 600 mil toneladas, sendo tanto maior este viés quanto maior o distanciamento do volume de safra do valor mencionado;

b) as duas primeiras previsões, dezembro e março, são mais fidedignas que a terceira, não havendo explicação para tal com base nos conhecimentos disponíveis;

c) na medida em que os operadores de mercado de algodão tomam conhecimento e se utilizam destas informações sobre previsão de safra, os erros aleatórios verificados nas duas primeiras previsões não são suficientemente "grandes" para forçar alterações exageradas dos preços.

PEETZ & AMARO (22) avaliaram a precisão das previsões de safra de laranja realizadas no Estado de São Paulo, pelo IEA e Coordenadoria de Assistência Técnica Integral (CATI), da Secretaria da Agricultura, e, na inexistência de um controle da previsão, foi considerado o levantamento final (novembro) como o volume real colhido. Os diversos levantamentos da safra de laranja foram relacionados entre si, dois a dois, através de funções linear e exponencial, ajustadas segundo um modelo de regressão linear simples. As variáveis, em ambas as formas, apresentaram elevado coeficiente de correlação e os testes dos coeficientes de regressão apresentaram significância ao nível de 0,5%. Os resultados dos ajustamentos da função linear foram considerados como os mais adequados para os objetivos propostos. Conclui-se ser possível projetar com bom grau de precisão a produção final de laranja a partir dos dados de 2º, 3º, 4º e 5º levantamentos, permitindo sua utilização como indicadores para decisões sobre a comercialização e política governamental para o setor.

CARVALHO et alii (5) repetiram e atualizaram estudo já anteriormente realizado para algodão por BRANDT; ARAUJO; SERRANO (2). Foram comparadas as três previsões realizadas durante o ano agrícola, individualmente, com a entrada nas usinas. Constataram que a segunda e terceira equações apresentaram as constantes de regressão não estatisticamente diferentes de zero, e, da mesma forma, os coeficientes de regressões não diferentes da unidade, preenchendo ambos os requisitos de fidedignidade exigidos. A terceira previsão, como era esperado, apresentou o melhor comportamento, com o coeficiente de determinação mais elevado..

CAMARGO (3), com a finalidade de conhecer as relações existentes entre as previsões e estimativa final para o café, utilizou também regressão simples e conclui que as três equações oferecem valores não tendenciosos da produção de café, sendo todas elas adequadas à formulação de políticas e tomada de decisão pelo empresário e pelo governo, sobretudo já a partir da 2ª previsão, uma vez que a divulgação da 3ª previsão ocorre em época tardia, tornando restrito seu uso.

VICENTE; CASER; CAMARGO (30) avaliaram, para o período de 1969-83, o comportamento das previsões das safras paulistas de 21 produtos em relação

ã estimativa final, considerando a área cultivada e a produtividade.

Para o algodão pode-se prever a área com antecedência, utilizando-se os levantamentos de setembro, novembro, fevereiro e abril. No caso da produtividade, os levantamentos de fevereiro e abril permitem que se conheça o rendimento final com nível de significância de 1%.

No caso do milho os levantamentos de setembro, fevereiro e abril mostraram-se adequados para prever a área final, e o de fevereiro e abril foram satisfatórios no que diz respeito à previsão de produtividade.

Finalmente, com respeito à soja, os levantamentos (exceto setembro) mostraram-se adequados para prever a área final e os de fevereiro e abril foram satisfatórios quanto à produtividade.

Além desses trabalhos que averiguaram de maneira clássica a precisão das previsões de safra, utilizando como regressores séries históricas de levantamentos sistemáticos, existem várias tentativas de previsão de rendimentos para uma série de culturas, utilizando-se de variáveis climáticas. Nesse caso merecem destaque estudos realizados por pesquisadores do Instituto de Economia Agrícola que consideraram o balanço hídrico como variável explicativa para as variações de rendimento, para uma gama de culturas do Estado de São Paulo.

SILVA; CASER; VICENTE (27) mensuraram o efeito das condições do tempo sobre a produtividade agropecuária, utilizando-se da análise de regressão para estimar as relações entre deficiência hídrica, geada e produtividade da terra. Os resultados mostraram haver grande sensibilidade das culturas anuais e perenes à deficiência de água, no período outubro-abril e setembro-março, respectivamente. As culturas perenes mostraram-se também muito susceptíveis à ocorrência de geadas.

SILVA; CASER; VICENTE (28) formularam e testaram modelos capazes de prever a produtividade da cultura do café no Estado de São Paulo conseguindo resultados animadores. A aplicação do método fornece previsões antecipadas em relação às baseadas em levantamentos de campo e, em alguns casos, melhores que estes últimos.

SILVA; VICENTE; CASER (29) reuniram um conjunto de estudos, focalizando o problema das previsões de safras agrícolas, pelo lado dos efeitos do tempo sobre o rendimento de importantes culturas, no Estado de São Paulo. Este esforço, construindo e testando modelos que utilizam variáveis meteorológicas, preenche uma lacuna existente no Instituto de Economia Agrícola (IEA) no que diz respeito à previsão de rendimento, em que pese o pioneirismo e tradição do IEA no desenvolvimento e utilização de modelos de previsão de área cultivada.

As condições meteorológicas foram incluídas nos modelos básicos

mente através de uma variável derivada, a deficiência hídrica. Para o cálculo da deficiência hídrica foi utilizado o método denominado de balanço hídrico, que consiste em contabilizar a água no solo, num processo em que a chuva representa o abastecimento e a evapotranspiração o consumo de água no solo, em condições de ser extraída pela planta.

Os resultados obtidos possibilitam concluir que o algodão pode ter seu rendimento previsto ao final dos meses de fevereiro e março com resultados melhores e com certa antecipação em relação aos baseados em levantamentos de campo, efetuados nos meses de fevereiro e abril pelo IEA-CATI, e que devido ao processo normal de coleta, processamento, depuração e análise, levam algum tempo até tornarem-se disponíveis: o de fevereiro é geralmente publicado em abril e o de abril em junho. Para o milho obtve-se modelos para previsão do rendimento nos meses de janeiro, fevereiro e março, este último com melhores ajustes e erros menores que a previsão de campo. Com a soja não se conseguiu modelos melhores que as previsões de campo, porém os erros percentuais médios de modelos terminados no mês de abril da ordem de 7%, confirmam a boa qualidade dos ajustes; mesmo quando o período é reduzido para fevereiro, o erro médio é inferior a 10%.

Convém citar ainda, trabalho desenvolvido por MASCARENHAS (16) que permite estimar a produtividade de soja com base nas precipitações pluviais ocorridas durante períodos relacionados ao florescimento. Esse método possibilita que a estimativa seja feita com até dois meses e meio de antecedência em relação aos métodos diretos de levantamento no campo.

Finalmente, cabe relatar os métodos utilizados pela CEPLAC (8) para a previsão de safra de cacau: contagem de frutos, amostragem e avaliação subjetiva. O método baseado na contagem de frutos, apesar do custo, mostrou-se mais funcional com relação aos resultados obtidos para a safra temporã do ano agrícola 1983/84, com erro de apenas 1,17% em relação à produção final, comparado ao erro de 17,4% da previsão por amostragem e de 20,35% na utilização do critério subjetivo.

5 - METODOLOGIA

A análise da função que associa os levantamentos preliminares com o levantamento final de produção de algodão, milho e soja, para cada instituição considerada, envolve o ajustamento de um modelo de regressão linear simples, representado por:

$$Y_j = \alpha_i + \beta_i X_{ij} + M_{ij}$$

onde:

Y_j é estimativa final da produção do ano j , com j variando de 1 a n , conforme número de anos considerados; X_{ij} é a i -ésima previsão de produção, com i variando de 1 a K , conforme número de levantamento de cada instituição, no ano j ; α_i e β_i são os parâmetros de regressão; e M_{ij} é o erro aleatório.

Os parâmetros α_i e β_i serão estimados pelo método dos mínimos quadrados, descritos em HOFFMANN & VIEIRA (13).

Para o intercepto, serão testadas as seguintes hipóteses:

$$H_0 : \alpha_i = 0$$

$$H_A : \alpha_i \neq 0$$

Para o coeficiente de regressão, serão testadas as seguintes hipóteses:

$$H_0 : \beta_i = 1$$

$$H_A : \beta_i \neq 1$$

Os testes de hipóteses serão efetuadas utilizando-se a distribuição t de Student, aos níveis de 1% e 5% significância.

Uma equação de regressão fidedigna apresenta $\alpha_1 = 0$ e $\beta_1 = 1$. Existe técnica estatística apropriada ao teste conjugado dessas hipóteses. Entretanto, devido ao fato de que o conhecimento isolado de cada teste de hipótese permite fazer inferências sobre vieses introduzidos na revisão dos dados, conforme BRANDT; ARAUJO; SERRANO (2), os testes serão efetuados separadamente.

O procedimento proposto possibilita verificar a coerência entre as previsões, as quais podem estar sub ou superestimadas. De antemão sabe-se que a última informação, por uma questão de lógica deva ser a escolhida para a estimativa de safra de um produto qualquer. Por motivos técnicos, entretanto, o seu uso é mais restrito, em virtude de sua divulgação dar-se em época já tardia para tomadas de decisão, seja a nível privado ou governamental. Assim, torna-se interessante verificar qual levantamento, em cada órgão, deve ser mais convenientemente aprimorado e utilizado, e qual o nível de previsibilidade de cada levantamento.

Embora haja maior probabilidade de acerto à medida que se aproxima o período final de colheita, mesmo no caso de levantamentos subjetivos, conforme enfatiza MUELLER (17), convém lembrar que apenas o algodão permite uma aferição relativamente mais confiável ou precisa dada a possibilidade de cotejo com a entrada nas usinas de beneficiamento. Neste caso existem dados disponíveis para os Estados Unidos e para São Paulo (5).

(5) O Estado do Paraná também possui estatísticas completas sobre entrada de algodão nas usinas de beneficiamento (1).

O problema da correlação serial nos resíduos da equação ajustada será examinado com o emprego de teste de Durbin-Watson, complementado pela estatística de Theil-Nagar.

Para São Paulo, no milho e na soja, serão utilizadas duas previsões de safra (fevereiro e abril) e o dado final (junho) nos anos em que houver previsões de safra em épocas diferentes ou em número maior que o normal, a colocação dos dados em suas respectivas séries será feita mediante critério subjetivo. Os dados se referem ao período 1947-85. Essas estatísticas não se encontram publicadas, de forma definitiva, necessitando-se de um trabalho prévio de consolidação dos mesmos. Para algodão serão utilizadas três previsões (fevereiro, abril e junho) sendo considerado como dado final a entrada de algodão em caroço nas usinas de beneficiamento, conforme levantamento da CATI.

Na parte relativa à entrada nas usinas, erros podem ocorrer nas anotações das quantidades e especificação da procedência. Por exemplo, algodão de outros estados pode ser declarado pelo usineiro como de procedência paulista, por razões de ordem fiscal ou creditícia (5).

Para efeito de comparação também serão apresentadas as equações de regressão tendo como variável dependente a última previsão (estimativa de junho).

Para o Brasil, os dados encontram-se publicados no Levantamento Sistemático de Produção Agrícola (1974 a 1986) do IBGE (15). São oito levantamentos mensais, para milho (maio a dezembro); seis para algodão herbáceo (março a agosto) e seis para soja (janeiro a junho), além da estimativa final para as três culturas. O período de disponibilidade dos dados é de 1974/75 a 1984/85.

Para os Estados Unidos, para milho e soja, serão utilizadas as previsões do USDA de agosto a dezembro, além da estimativa final, do COMMODITY YEAR BOOK (9). A série utilizada cobre o período 1947/48 a 1984/85. Para algodão, as previsões referem-se ao período agosto-janeiro, tendo sido utilizado o dado de entrada nas usinas como dado definitivo.

Igualmente, a exemplo de São Paulo, para efeito de comparação serão apresentadas as equações de regressão tendo como variável dependente a última previsão (estimativa de junho).

6 - RESULTADOS E DISCUSSÃO

Dadas as diferenças entre os procedimentos de obtenção dos dados estatísticos de produção adotados pelas instituições analisadas, torna-se

difícil comparar os resultados das equações de regressão ajustadas no presente estudo.

Levantamentos baseados em opiniões de pessoas, ligadas a entidades públicas ou particulares, como os efetuados pelo IBGE, estão sujeitos à tendenciosidade, na medida em que aquelas opiniões reflitam enfoques subjetivos dos participantes. Além disso, podem apresentar maior rigidez no ajustamento às condições da realidade produtiva, dada a tendência à irreduzibilidade de na reavaliação de estimativas pelos participantes do grupo de informantes. Nesse sentido, o elevado grau de ajuste das equações pode refletir não um alto grau de qualidade do levantamento, mas sim um alto grau de apego aos números anteriormente estabelecidos. Na inexistência de um levantamento objetivo que possibilite o confronto com os dados subjetivos, não há como avaliar, a contento, a acurácia das previsões.

Por outro lado, a utilização de um método objetivo, como aquele do levantamento do IEA-CATI, defronta-se com problemas de outra ordem. A qualidade dos levantamentos, quando realizados por enumeradores não constituintes de uma equipe homogênea, deixa a desejar. O elevado número de questionários não preenchidos, incompletos ou portadores de erros diversos, torna necessário uma rigorosa revisão dos mesmos, podendo exigir o retorno ao campo para complementação, o que, normalmente, não tem ocorrido. A subjetividade pode estar presente no preenchimento de questionários, influenciando as respostas, principalmente quando o entrevistador está ligado, direta ou indiretamente, à atividade de produção agropecuária. Experiências pessoais tendem a ser generalizadas, em prejuízo da qualidade do levantamento.

Convém ressaltar que os dados primários obtidos no campo passam por um processo exaustivo de depuração e consistência interna antes de serem tabulados, conforme relatam PINO; CAMARGO; JIMENEZ OSSIO (25).

Feitas essas ressalvas, passa-se ao exame e discussão dos resultados obtidos por produto.

6.1 - Algodão

Foram seis as equações de regressão linear simples ajustadas aos valores observados das previsões de safra de algodão nos Estados Unidos, tendo como variável dependente a entrada de algodão em caroço nas usinas.

Todas as equações foram significantes ao nível de 1%, como indica o pelos valores do teste F de Snedecor (quadro 1). O poder explicativo das regressões, medido pelo coeficiente de determinação (R^2), foi relativamente elevado em todas as equações, com crescimento ao longo do tempo. O teste de Durbin-Watson aponta problema de correlação serial dos resíduos nas equa

QUADRO 1. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carço, Tendo como Variável Dependente a Entrada de Algodão em Carço nas Usinas de Beneficiamento, Estados Unidos, 1947/ 48 a 1984/85 ⁽¹⁾

Item	Previsão					
	Agosto	Setembro	Outubro	Novembro	Dezembro	Final
Constante de regressão (b_0)	-364,39	-206,83	-163,40	-122,52	-25,98	-18,52
Coefficiente de regressão (b_1)	1,113	1,057	1,040	1,024	0,994	0,993
Desvio-padrão						
s_{b_0}	207,93	195,59	129,65	73,83	37,52	30,08
s_{b_1}	0,073	0,069	0,045	0,026	0,013	0,010
Razão t de Student						
t_{b_0}	-1,753	-1,058	-1,261	-1,660	-0,693	-0,616
t_{b_1}	1,547	0,826	0,888	0,923	0,461	0,700
Teste F de Snedecor	226,90 ^a	231,60 ^a	513,19 ^a	1.541,19 ^a	5.586,71 ^a	8.644,30 ^a
Coefficiente de determinação (R^2)	0,8630	0,8654	0,9344	0,9771	0,9935	0,9958
Estatística de Durbin-Watson	1,87	1,48	1,63	1,87	1,16 ^b	0,32 ^b
Coefficiente de Theil-Nagar	0,061	0,257	0,181	0,062	0,416	0,837

⁽¹⁾ O número de observações (n) é igual a 38. As letras representam o nível de significância a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do United States Department of Agriculture (9).

ções referentes a dezembro e final, o que se reflete nos valores do coeficiente de Theil-Nagar.

O teste da hipótese nula referente à constante de regressão (b_0) não foi estatisticamente significativa, mesmo ao nível de 5%, para nenhuma das equações ajustadas, conforme demonstrado pela razão t de Student, não se rejeitando, desse modo, a hipótese de que a constante de regressão é estatisticamente igual a zero.

O coeficiente de regressão (b_1) não foi estatisticamente diferente da unidade, ao nível de significância de 5%, nas equações ajustadas, conforme os valores obtidos da razão t de Student.

A apreciação conjunta dos resultados obtidos para a constante de regressão e para o coeficiente de regressão permite admitir a hipótese de fidedignidade das equações ajustadas, qualquer que seja o mês considerado, quando usadas na previsão da entrada de algodão nas usinas estadunidenses. Também nas regressões em que a estimativa final (janeiro) foi considerada como variável dependente os resultados estatísticos mostraram-se satisfatórios (quadro 2).

Para o Brasil, as seis equações de regressão ajustadas aos valores observados das variáveis, apresentaram-se estatisticamente significantes ao nível de 1%, conforme indicado pela estatística F de Snedecor (quadro 3). O poder explicativo das equações (R^2) mostrou-se relativamente elevado e crescente ao longo dos meses. Não se constatou problema de correlação serial dos resíduos em grau elevado, conforme estatística de Durbin-Watson e o coeficiente de Theil-Nagar.

Nos testes relacionados às constantes de regressão e aos coeficientes de regressão, as hipóteses nulas não foram rejeitadas ao nível de significância de 5%. Constatou-se, assim que, em todas as equações ajustadas, as constantes e os coeficientes de regressão não diferem do ponto de vista estatístico, de zero e da unidade, respectivamente.

Esses resultados sugerem que as equações de regressão entre previsões mensais da safra de algodão no Brasil e a estimativa final podem ser utilizadas para prover previsões fidedignas da estimativa final de uma safra.

Ao se analisar as equações referentes às previsões de safra de algodão em caroço no Estado de São Paulo, constata-se a significância geral das equações, conforme o teste F de Snedecor (quadro 4).

O poder explicativo das regressões é bastante reduzido para as previsões de fevereiro e abril, elevando-se acentuadamente na previsão de junho. As estatísticas de Durbin-Watson e os coeficientes de Theil-Nagar não revelam a existência de grau elevado de autocorrelação serial nos resí

QUADRO 2. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carço, Tendo como Variável Dependente a Estimativa Final, Estados Unidos, 1947/48 a 1984/85 ⁽¹⁾

Item	Previsão				
	Agosto	Setembro	Outubro	Novembro	Dezembro
Constante de regressão (b_0)	-325,46	-174,49	-132,78	-99,66	-6,53
Coefficiente de regressão (b_1)	1,112	1,058	1,042	1,029	1,000
Desvio-padrão					
s_{b_0}	215,61	200,25	134,62	73,11	25,65
s_{b_1}	0,076	0,071	0,47	0,025	0,009
Razão t de Student					
t_{b_0}	-1,510	-0,872	-0,987	-1,364	-0,255
t_{b_1}	1,470	0,828	0,897	1,157	0,038
Teste F de Snedecor	210,70	221,60 ^a	477,95 ^a	1.587,46 ^a	12.101,49 ^a
Coefficiente de determinação (R^2)	0,8540	0,8602	0,9299	0,9778	0,9970
Estatística de Durbin-Watson	1,88	1,49	1,59	1,84	1,95
Coefficiente de Theil-Nagar	0,059	0,252	0,201	0,078	0,023

⁽¹⁾ O número de observações (n) é igual a 38. A letra a representa o nível de significância de 1%.

Fonte: Elaborado a partir de dados básicos do United States Department of Agriculture (9).

QUADRO 3. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carço, Brasil, 1974/75 a 1984/85 (1)

Item	Previsão					
	Março	Abril	Maió	Junho	Julho	Agosto
Constante de regressão (b_0)	166,523	-7.284	-79.498	-59,346	7.259	3.158
Constante de regressão (b_1)	0,908	1,016	1,060	1,051	1,005	1,013
Desvio-padrão						
s_{b_0}	195.964	199.882	123.244	80.033	60.571	54.699
s_{b_1}	0,123	0,125	2,774	0,050	0,037	0,034
Razão t de Student						
t_{b_0}	0,849	-0,037	-0,646	-0,742	0,119	0,057
t_{b_1}	0,747	0,128	0,077	1,020	0,135	0,382
Teste F de Snedecor	54,30 ^a	65,06 ^a	187,32 ^a	434,87 ^a	702,57 ^a	865,90 ^a
Coefficiente de determinação (R^2)	0,8578	0,8784	0,9541	0,9797	0,9873	0,9897
Estatística de Durbin-Watson (2)	1,90	2,27	2,46	2,07	2,75	2,65
Coefficiente de Theil-Nagar	0,047	-0,137	-0,236	-0,040	-0,378	-0,330

(1) O número de observações (n) é igual a 11. A letra a representa o nível de significância de 1%.

(2) O valor da Tabela para comparação sô é disponível para n=15.

Fonte: Elaborado a partir de dados básicos da Fundação IBGE (15).

QUADRO 4. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carço, Tendo como Variável Dependente a Entrada de Algodão em Carço nas Usinas de Beneficiamento, Estado de São Paulo, 1947/48 a 1984/85 (1)

Item	Previsão		
	Fevereiro	Abril	Junho
Constante de regressão (b_0)	282,41	172,34	-6,35
Coefficiente de regressão (b_1)	0,493	0,719	1,042
Desvio-padrão			
s_{b_0}	66,50	60,72	40,66
s_{b_1}	0,404	0,106	0,072
Razão t de Student			
t_{b_0}	4,246 ^a	2,837 ^a	-0,157
t_{b_1}	1,254	2,650 ^b	0,583
Teste F de Snedecor	19,96 ^a	45,21 ^a	208,59 ^a
Coefficiente de determinação (R^2)	0,3842	0,5855	0,8669
Estatística de Durbin-Watson	2,04	2,03	1,50
Coefficiente de Theil-Nagar	-0,024	-0,019	0,246

(1) O número de observações (n) é igual a 34. As letras representam o nível de significância: a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola e da Coordenadoria de Assistência Técnica Integral.

duos das equações.

O teste da hipótese nula de que a constante de regressão (b_0) é igual a zero levou à rejeição dessa hipótese nas equações relativas a fevereiro e abril, ao nível de 1% de significância. Na previsão de junho, essa hipótese não foi rejeitada.

O teste da hipótese nula de que o coeficiente de regressão (b_1) é igual à unidade levou à rejeição dessa hipótese ao nível de significância de 5%, na previsão de abril.

A apreciação global dos resultados desses testes de significância indica que apenas as equações da previsão de junho podem ser consideradas fidedignas na previsão de valores de entrada de algodão em caroço nas usinas de beneficiamento paulistas.

Para São Paulo, também foram ajustadas equações de regressão tendo como variável dependente o levantamento final (estimativa de junho) verificando a aferição dos dados das instituições responsáveis pelos levantamentos — IEA e CATI.

O poder explicativo das duas regressões é maior que o das equações correspondentes quando considerada a entrada nas usinas como variável dependente (quadro 5). Entretanto, os testes das hipóteses nulas, tanto para a constante (b_0) quanto para o coeficiente (b_1) de regressão, foram significativos ao nível de 1% revelando a não fidedignidade dos levantamentos. Por conseguinte, infere-se que as estimativas preliminares são menos consistentes quando se usa a estimativa final como variável dependente do que quando são usadas as entradas de algodão nas usinas de beneficiamento.

6.2 - Milho

As equações de regressão linear simples ajustadas a dados de previsão de safra de milho dos Estados Unidos apresentaram-se estatisticamente significantes ao nível de 1%, conforme indicado pelos valores da estatística F de Snedecor (quadro 6). O poder explicativo da regressão (R^2) foi bastante elevado em todas as equações. A correlação serial nos resíduos não se destacou como problema, a não ser na equação referente ao nível de 5%, o que é corroborado pelos valores do coeficiente de Theil-Nagar.

A hipótese nula $b_1 = 1$, com o emprego da razão t de Student, foi rejeitada, ao nível de, pelo menos 5%, nas equações referentes aos meses de setembro, outubro, novembro e dezembro. Além disso, a constante de regressão (b_0) foi estatisticamente diferente de zero nas equações de outubro e novembro, ao nível de, pelo menos, 5%.

Na medida em que um elevado grau de fidedignidade de uma previ

QUADRO 5. - Principais Características das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carço, Tendo como Variável Dependente a Estimativa Final, São Paulo, 1947/48 a 1984/85 (1)

Item	Previsão	
	Fevereiro	Abril
Constante de regressão (b_0)	259,30	158,92
Coefficiente de regressão (b_1)	0,498	0,706
Desvio-padrão		
s_{b_0}	52,81	43,69
s_{b_1}	0,087	0,076
Razão t de Student		
t_{b_0}	4,909 ^a	3,637 ^a
t_{b_1}	5,722 ^a	3,810 ^a
Teste F de Snedecor	32,31 ^a	84,39 ^a
Coefficiente de determinação (R^2)	0,5024	0,7250
Estatística de Durbin-Watson	2,27	2,48
Coefficiente de Theil-Nagar	-0,138	-0,242

(1) O número de observações (n) é igual a 34. A letra a representa o nível de significância de 1%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

QUADRO 6. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Milho, Estados Unidos, 1947/48 a 1984/85 ⁽¹⁾

Item	Previsão				
	Ago.	Set.	Out.	Nov.	Dez.
Constante de regressão (b_0)	-4.584	-4.784	-4.767	-3.881	-1.794
Coefficiente de regressão (b_1)	1,055	1,059	1,057	1,044	1,019
Desvio-padrão					
s_{b_0}	4.756	2.768	2.021	1.280	909
s_{b_1}	0,038	0,022	0,016	0,010	0,007
Razão t de Student					
t_{b_0}	-0,964	-1,729	-2,359 ^b	-3,032 ^a	-1,975
t_{b_1}	1,429	2,657 ^b	3,505 ^a	4,357 ^a	2,703 ^b
Teste F de Snedecor	751,62 ^a	2.231,87 ^a	4.187,78 ^a	10.311,41 ^a	19.836,06 ^a
Coefficiente de determinação (R^2)	0,9542	0,9841	0,9914	0,9965	0,9981
Estatística de Durbin-Watson	1,679	1,806	1,728	1,508	0,693 ^a
Coefficiente de Theil-Nagar	0,160	0,096	0,135	0,245	0,653

⁽¹⁾ O número de observações (n) é igual a 38. As letras representam o nível de significância: a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do United States Department of Agriculture (9).

são de safra ocorra somente na presença simultânea de constante não diferente, estatisticamente, de zero e coeficiente de regressão não diferente, estatisticamente, da unidade, tem-se que as equações de regressão linear simples ajustadas para as previsões de safra de milho, nos Estados Unidos, para os meses do último quadrimestre do ano, não atendem a esses requisitos. Esses são observados, apenas, na equação do mês de agosto. As demais equações podem dar origem a previsões dotadas de maior grau de incerteza.

Para o Brasil tem-se que as oito equações de regressão ajustadas apresentaram-se estatisticamente significantes ao nível de 1% conforme indicado pela estatística F de Snedecor (quadro 7). O poder explicativo da regressão (R^2), especialmente para os meses de julho a novembro, foi bastante elevado, notando-se uma queda no valor do coeficiente de determinação em dezembro. O teste de Durbin-Watson, utilizado para o exame de correlação serial nos resíduos, mostrou-se inconclusivo para os meses de agosto, novembro e dezembro, não sendo significativo nos demais meses.

Testando-se a hipótese nula $b_1 = 1$, a razão t de Student calculada não leva à rejeição dessa hipótese em nenhuma das equações ajustadas, ao nível de 5%. A esse mesmo nível de significância, a constante de regressão (b_0) não foi estatisticamente diferente de zero em nenhuma das oito equações de regressão ajustadas.

Desse modo, as equações de regressão de previsão de safra de milho, para o Brasil, podem ser consideradas como satisfazendo as condições necessárias da estimativa final de produção desse cereal.

Ao se analisar as duas equações de regressão linear simples ajustadas para a previsão de safra de milho no Estado de São Paulo tem-se que em função dos valores calculados para a estatística F de Snedecor, ambas apresentaram-se estatisticamente significantes ao nível de 1% (quadro 8). O poder explicativo (R^2) das duas equações foi bastante elevado. O teste de Durbin-Watson apresentou resultado positivo quanto à presença de correlação serial nos resíduos, ao nível de 5% para o mês de abril.

O coeficiente de regressão b_1 mostrou-se significativamente diferente de unidade, ao nível de 1%, para a equação de abril. O intercepto b_0 , em ambas as equações de regressão, não se mostrou estatisticamente diferente de zero, mesmo ao nível de significância de 5%.

Com esses resultados, a equação de regressão ajustada para o mês de abril não apresenta condições para que ocorra um elevado grau de fidedignidade, se usada para prever safra de milho no Estado de São Paulo.

6.3 - Soja

Para os Estados Unidos, foram ajustadas cinco equações de regres

QUADRO 7. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Milho, Brasil, 1974/75 a 1984/85 (1)

Item	Previsão							
	Mai.	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Constante de regressão (b_0)	-761.976	-690.660	-508.807	-615.933	-254.474	-136.838	9.270	-337.534
Coefficiente de regressão (b_1)	1.000	1,016	1,016	1,028	1,008	1,004	0,998	1,030
Desvio-padrão								
s_{b_0}	2.247.275	1.331.054	693.871	704.871	175.141	103.227	95.232	1.869.133
s_{b_1}	0,113	0,068	0,035	0,036	0,009	0,005	0,004	0,099
Razão t de Student								
t_{b_0}	-0,340	-0,519	-0,734	-0,874	-1,453	-1,326	0,097	-0,181
t_{b_1}	0,001	0,235	0,455	0,771	0,897	0,760	0,343	0,306
Teste F de Snedecor	78,14 ^a	221,41 ^a	800,39 ^a	784,01 ^a	12.244,32 ^a	34.824,32 ^a	40.305,01 ^a	108,27 ^a
Coefficiente de Determinação (R^2)	0,8967	0,9609	0,9888	0,9886	0,9992	0,9997	0,9997	0,9232
Estatística de Durbin-Watson ⁽²⁾	2,238	2,291	2,734	2,906 IN	2,532	2,351	1,194 IN	1,172 IN
Coefficiente de Theil-Nagar	0,120	-0,147	-0,369	-0,456	-0,268	-0,177	0,404	0,415

(1) O número de observações (n) é igual a 11. A letra a representa o nível de significância de 1%. IN = inconclusivo.

(2) Os valores calculados foram cotejados com os da Tabela para 15 observações, que é o menor número disponível.

Fonte: Elaborado a partir de dados básicos da Fundação IBGE (15).

QUADRO 8. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Milho, Estado de São Paulo, 1947/48 a 1984/85 (1)

Item	Previsão	
	Fevereiro	Abril
Constante de regressão (b_0)	-2,994	27,487
Coefficiente de regressão (b_1)	0,946	0,951
Desvio-padrão		
s_{b_0}	101,456	41,309
s_{b_1}	0,042	0,017
Razão t de Student		
t_{b_0}	-0,030	0,665
t_{b_1}	1,246	2,743 ^a
Teste F de Snedecor	490,305 ^a	2.898,476 ^a
Coefficiente de determinação (R^2)	0,9387	0,9890
Estatística de Durbin-Watson	1,761	0,989 ^b
Coefficiente de Theil-Nagar	0,119	0,505

(1) O número de observação (n) é igual a 34; as observações referentes aos anos 1954/55, 1956/57, 1957/58 e 1960/61 foram eliminadas, por incompletas. As letras representam o nível de significância: a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

são linear simples, tendo as previsões dos meses de agosto a dezembro como variáveis independentes e a estimativa final como variável dependente.

Todas as equações foram significantes ao nível de 1%, como indicado pelos elevados valores do teste F de Snedecor (quadro 9). O poder explicativo das regressões, medido pelo coeficiente de determinação (R^2), foi bastante elevado para todas as previsões. O teste de Durbin-Watson de correlação serial nos resíduos apresentou-se inconclusivo na equação de agosto, o que é corroborado pelo coeficiente de Theil-Nagar.

O teste de hipótese nula referente à constante de regressão não foi estatisticamente significativa, a nível de 5%, para nenhuma das equações ajustadas, conforme demonstrado pela razão t de Student, não se rejeitando, portanto, a hipótese de que a constante de regressão é igual a zero.

O coeficiente de regressão foi estatisticamente diferente da unidade, ao nível de significância de 5%, apenas na equação de dezembro, conforme indicado pelos valores da razão t de Student.

A apreciação conjunta dos resultados estatísticos obtidos para a constante de regressão e para o coeficiente de regressão é de molde a considerar as equações ajustadas, qualquer que seja o mês considerado, à exceção de dezembro, como fidedignas na previsão de produção de soja esta dunidense.

Para o Brasil, foram ajustadas seis equações de regressão linear simples, referentes às previsões dos meses de janeiro a junho, tendo a estimativa final como variável dependente.

Todas as equações foram significantes ao nível de 1%, conforme os valores do teste F de Snedecor (quadro 10). O poder explicativo das regressões, medido pelo coeficiente de determinação (R^2) apresentou-se em crescimento, sendo tanto maior quanto mais próximo da estimativa final for a previsão considerada. Os valores de R^2 obtidos, principalmente os dos três últimos meses, são bastante elevados. A estatística de Durbin-Watson mostrou-se inconclusiva na equação de março.

A constante de regressão e o coeficiente de regressão foram considerados do ponto de vista estatístico diferentes de zero, e da unidade, respectivamente ao nível de significância de 5%, apenas na equação referente ao mês de junho, conforme as correspondentes razões t calculadas.

A avaliação desses resultados para a constante de regressão e para o coeficiente de regressão permite indicar, como fidedignas para previsão das safras de soja no Brasil, as equações referentes a todos os meses, exceto o de junho.

As duas equações de regressão linear simples ajustadas para o

QUADRO 9. - Principais Características Estatísticas das Equações Linear Simples de Previsão de Safras de Soja, Estados Unidos, 1947/48 a 1984/85 (1)

Item	Previsão				
	Agosto	Setembro	Outubro	Novembro	Dezembro
Constante de regressão (b_0)	585,13	381,67	288,80	220,35	208,79
Coefficiente de regressão (b_1)	0,978	0,988	0,994	0,994	0,992
Desvio-padrão					
s_{b_0}	599,95	487,54	413,16	305,82	122,92
s_{b_1}	0,018	0,015	0,013	0,009	0,003
Razão t de Student					
t_{b_0}	0,975	0,782	0,699	0,720	1,698
t_{b_1}	1,159	0,778	0,421	0,576	2,023 ^b
Teste F de Snedecor	2.681,36 ^a	4.110,90 ^a	5.756,96 ^a	10.555,07 ^a	65.420,88 ^a
Coefficiente de determinação (R^2)	0,9867	0,9913	0,9937	0,9966	0,9994
Estatística de Durbin-Watson	1,42 IN	1,69	1,66	1,51	1,86
Coefficiente de Theil-Nagar	0,287	0,151	0,169	0,242	0,065

(1) O número de observações (n) é igual a 38. As letras representam o nível de significância: a = 1% e b = 5% e IN = inconclusivo.

Fonte: Elaborado a partir de dados básicos do United States Department of Agriculture (9).

QUADRO 10. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Soja, Brasil, 1974/75 a 1984/85 ⁽¹⁾

Item	Previsão					
	Janeiro	Fevereiro	Março	Abril	Maió	Junho
Constante de regressão (b_0)	779.132	1.077.512	1.005.596	905.510	678.849	911.535
Coefficiente de regressão (b_1)	0,916	0,905	0,911	0,916	0,942	0,936
Desvio-padrão						
s_{b_0}	160.076	1.074.887	956.172	756.248	463.817	301.816
s_{b_1}	0,116	0,078	0,070	0,055	0,034	0,022
Razão t de Student						
t_{b_0}	0,486	1,002	1,051	1,197	1,463	3,020 ^b
t_{b_1}	0,720	1,204	1,264	1,506	1,667	2,799 ^b
Teste F de Snedecor	62,164 ^a	131,935 ^a	168,737 ^a	274,246 ^a	755,788 ^a	1.722.228 ^a
Coefficiente de determinação (R^2)	0,8735	0,9361	0,9493	0,9682	0,9882	0,9948
Estatística de Durbin-Watson ⁽²⁾	1,525	1,561	1,010 IN	1,465	1,791	1,635
Coefficiente de Theil-Nagar	0,238	0,220	0,496	0,268	0,104	0,182

⁽¹⁾ O número de observações (n) é 11.

⁽²⁾ Os valores calculados foram cotejados com os da Tabela para 15 observações, que é o menor número disponível. As letras representam o nível de significância: a=1% e b=5%. IN = inconclusivo.

Fonte: Elaborado a partir de dados básicos da Fundação IBGE (15).

Estado de São Paulo referiram-se aos meses de fevereiro e de abril.

Ambas as equações foram significantes ao nível de 1%, conforme os valores do teste F (quadro 11). O coeficiente de determinação (R^2), foi bastante elevado nos dois casos. O teste de Durbin-Watson não indicou presença de correlação serial dos resíduos em nível que pudesse causar possíveis problemas, o que também ocorreu com o coeficiente de Theil-Nagar.

A razão t de Student, calculada para testar a hipótese nula de que o coeficiente de regressão é igual à unidade, para a previsão de abril, foi significativa ao nível de 5%, levando à rejeição dessa hipótese. Desse modo, a previsão de abril não apresenta as condições necessárias para um elevado grau de fidedignidade ao estimar a produção final paulista de soja.

7 - CONCLUSÕES

O método de análise de consistência de previsões de safra utilizando a regressão linear simples mostrou-se em geral, adequado conforme evidenciaram os valores do teste F e do coeficiente de determinação. Os resultados obtidos no estudo sugerem que a consistência das previsões de algodão é maior que a das de soja, que por sua vez é maior que a das de milho.

Dos resultados pode-se também inferir que, independentemente dos procedimentos de levantamento dos dados, o melhor desempenho pode ser atribuído aos dados do Brasil, seguindo-se os dos Estados Unidos e o do Estado de São Paulo. Essa ressalva quanto aos procedimentos de levantamento, já efetuada anteriormente, diz respeito ao levantamento subjetivo de dados do Brasil o que, segundo MUELLER (17) não permite a atribuição de margens de erro e impossibilita o estabelecimento de intervalos de confiança. Assim, mesmo que a qualidade de dados obtidos subjetivamente seja razoável, isso não pode ser comparado de maneira científica.

Uma informação de utilidade para os usuários de estatísticas de produção apresentada pelo trabalho, é a que especifica o mês de levantamento, a partir do qual já se tem uma previsão confiável da produção.

Para algodão, nos Estados Unidos, qualquer previsão a partir de agosto, fornece indicação fidedigna do volume de entrada nas usinas. Para o Brasil, qualquer previsão a partir de março fornece indicação segura da produção. Em São Paulo, apenas a estimativa final, efetuada em junho, fornece indicação aproveitável do volume de entrada nas usinas.

No caso do milho, nos Estados Unidos, apenas é fidedigna a primeira previsão (agosto). No Brasil, todas as previsões a partir de maio, podem ser utilizadas. No Estado de São Paulo, apenas a previsão de feverei

QUADRO 11. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Soja, Estado de Paulo, 1947/48 a 1984/85 (¹)

Item	Previsão	
	Fevereiro	Abril
Constante de regressão (b_0)	-2,516	-1,638
Coefficiente de regressão (b_1)	1,000	0,978
Desvio-padrão		
s_{b_0}	8,470	5,289
s_{b_1}	0,014	0,009
Razão t de Student		
t_{b_0}	-0,298	-0,310
t_{b_1}	0,005	2,393 ^b
Teste de Snedecor	4.467,92 ^a	11.462,23 ^a
Coefficiente de Determinação (R^2)	0,9933	0,9973
Estatística de Durbin-Watson	1,78	1,51
Coefficiente de Theil-Nagar	0,107	0,244

(¹) O número de observações (n) é 32, tendo sido excluídas da análise, por serem incompletos os dados das safras 1949/50, 1950/51, 1954/55, 1956/57, 1957/58 e 1959/60. As letras representam o nível de significância: a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

ro fornece uma indicação fidedigna de produção de milho.

No que se refere à soja, tem-se nos Estados Unidos levantamentos fidedignos no período agosto-novembro. No Brasil, são aproveitáveis as previsões de janeiro a maio não o sendo o de junho. No Estado de São Paulo, apenas a previsão de fevereiro é fidedigna.

Como os dados para o Brasil e São Paulo indicam deficiência na qualidade, é necessário que os órgãos envolvidos na sua coleta, elaboração e divulgação, atuem no sentido de aprimorá-las. No caso de São Paulo, tem-se que o cadastro utilizado para extração da amostra não é suficientemente atualizado, o que é agravado por mudanças regionais marcantes nas atividades agropecuárias, em pouco espaço de tempo.

Uma possibilidade adicional de aprimoramento dessas estatísticas é o recurso a métodos alternativos, como ocorre em outros países, não visando substituir os atuais sistemas de previsão de safras, mas sim complementá-los. Um desses métodos, bastante promissor, é o emprego de satélites artificiais para sensoriamento remoto, possibilitando estimar, com margem de erro conhecida, variáveis agropecuárias importantes para obtenção de previsões de safra mais fidedignas.

O interregno entre coleta e divulgação dos dados é elevado, o que prejudica a utilização da informação pelos interessados. A agilização dessa divulgação já constituiriam num grande avanço para os usuários.

Essa lentidão pode fazer com que o tomador de decisões, quer do governo, quer da empresa privada, necessite recorrer a outros instrumentos para obter uma previsão de produção. A possibilidade de utilização de métodos econométricos para prever a produção, em uma dada safra, não deve ser posta de lado. Nesse sentido, modelos que considerem variáveis meteorológicas como fatores explicativos da produção, podem ser acionados para obter previsões em tempo hábil, já que essas variáveis estão prontamente disponíveis nas estações meteorológicas ou mesmo, a nível de propriedade, facultando a regionalização da previsão.

LITERATURA CITADA

1. ALGODÃO: relatório, 1960/70-1984/85. Curitiba, Empresa Paranaense de Classificação de Produtos - CLASPAR, 1970-1986.
2. BRANDT, Sergio A.; ARAUJO, Paulo F.C. de; SERRANO, Ondalva. Precisão e fidedignidade das previsões de safra de algodão no Estado de São Paulo, 1947-64. Piracicaba, ESALQ/USP, 1966. 15p.
3. CAMARGO, Maria de L.B. Avaliação das previsões de safra de café, Estado de São Paulo, 1960/61 a 1979/80. Informações Econômicas, São Paulo 11(8):19-24, ago. 1981.
4. CAMPOS, Humberto & PIVA, Luiz H. de O. Dimensionamento de amostra para estimativa e previsão de safra no Estado de São Paulo. Agricultura em São Paulo, SP, 21(3):65-68, 1974.
5. CARVALHO, Flávio C. de et alii. Avaliação das previsões de safra de algodão no Estado de São Paulo, 1947/48 a 1976/77. São Paulo, Secretaria da Agricultura, IEA, 1978. 11p. (Relatório de Pesquisa, 17/78)
6. CENSO AGROPECUÁRIO: Brasil, 1980. Rio de Janeiro, FIBGE, 1984. v.2, t.3, n.1.
7. CENSO AGROPECUÁRIO: São Paulo, 1980. Rio de Janeiro, FIBGE, 1984. v.2, t.3, n.19, 1ª parte.
8. CEPLAC. Divisão Sócio-Econômica. Previsão de safras. Informe de Pesquisas de 1983, Ilhêus, 1985. p.211-215.
9. COMMODITY YEARBOOK. New York, Commodity Research Bureau, 1960-1985.
10. ESTADOS UNIDOS. Department of Agriculture. Major statistical series of the U.S. Department of Agriculture: how they are constructed and used - crop and livestock estimates. Washington, 1971 v.8. (Agriculture Handbook, 365)

11. FARM marketing of field crops, United States, 1977-78 and 1978-79. Crop Production, Washington, dec. 1979. p.8-8.
12. FUNDAÇÃO GETÚLIO VARGAS. Centro de Estudos Agrícolas. Calendário agrícola. Rio de Janeiro, 1979. 254p.
13. HOFFMANN, Rodolfo & VIEIRA, Sonia. Análise de regressão: uma introdução à econometria. São Paulo, Hucitec, 1977. 339p.
14. KOHLS, Richard L. & UHL, Joseph N. Marketing of agricultural products. 5.ed. New York, Macmillan Publishing, 1980. 612p.
15. LEVANTAMENTOS SISTEMÁTICO DA PRODUÇÃO AGRÍCOLA. Rio de Janeiro, FIBGE, 1975-1986.
16. MASCARENHAS, Hipólito A.A. Estimativa da produção de soja. In: FUNDAÇÃO CARGILL, ed. A soja no Brasil Central. 2ed. Campinas, 1982. p.394-404.
17. MUELLER, Charles C. Previsões e estimativas de safras agrícolas: questões sobre suas formas de obtenção, confiabilidade e seus impactos. Indicadores IBGE, Rio de Janeiro, 5(8):73-84, ago. 1986. Nova Série.
18. NOGUEIRA JR., Sebastião; CARVALHO, Flávio C.de; TSUNECHIRO, Alfredo. Consistência das previsões de safra de algodão em caroço nos Estados Unidos, no Brasil e em São Paulo. In: REUNIÃO NACIONAL DO ALGODÃO,4., Belém 1986. Resumos... Campina Grande, EMBRAPA/CNPA, 1986. p.121.
19. ———; ———; ———. Consistência das previsões de safra de milho nos Estados Unidos, no Brasil e em São Paulo. In: CONGRESSO NACIONAL DE MILHO E SORGO, 16., Belo Horizonte, 1986. Resumos... Sete Lagoas, EMBRAPA/CNPMS, 1986. p.95-96.
20. ———; ———; ———. Consistência das previsões de safra de soja nos Estados Unidos, no Brasil e em São Paulo. In: SEMINÁRIO NACIONAL DE PESQUISA DE SOJA, 4., Porto Alegre, 1986. Resumos e Programa ... Londrina, EMBRAPA/CNPSo, 1986. p.23-24.

21. PEDRO Jr., Mário J. et alii, eds. Instruções agrícolas para o Estado de São Paulo. 3.ed.rev.atual. Campinas, Secretaria de Agricultura e Abastecimento, Instituto Agronômico, 1986. 229p. (Boletim 200)
22. PEETZ, Marcia S. & AMARO, Antonio A. Aplicação de equação de regressão à previsão de safra de laranja no Estado de São Paulo. Informações Econômicas, São Paulo, 8 (3):1-8, mar. 1978.
23. PINO, Francisco A. & CASER, Denise V. Análise de erros não amostrais em levantamentos para previsão e estimativas de safras no Estado de São Paulo. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1984. 25p. (Relatório de Pesquisa, 10/84)
24. ———; ———. Falta de respostas em levantamento por amostragem: um estudo de caso. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1984. 25p. (Relatório de Pesquisa, 08/84)
25. ———; CAMARGO, Maria de L.B.; JIMENEZ OSSIO, Júlio H. Estatísticas de produção agrícola no Estado de São Paulo. 1942/43 a 1984/85. Informações Econômicas, 16:1-60, set. 1986. Supl. n.1.
26. SALMON, David S. The cycle of food and fiber production. Commodity Yearbook, New York, 1986. p.15-28.
27. SILVA, Gabriel L.S.P. da; CASER, Denise V.; VICENTE, José R. Efeito das condições do tempo sobre a produtividade agrícola no Estado de São Paulo. Revista de Economia Rural, Brasília, 23(1):3-19, jan./mar. 1985.
28. ———; ———; ———. Variações do tempo e da produtividade agrícola: proposta de uma metodologia de análise e sua aplicação para a cultura do café no Estado de São Paulo. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1986. 28p. (Relatório de Pesquisa, 13/86)

29. ———; VICENTE, José R.; CASER, Denise V. Variações do tempo e produtividade agrícola: um subsídio à previsão de safras no Estado de São Paulo. Campinas, Fundação Cargill, 1986. 148p.
30. VICENTE, José R.; CASER, Denise V.; CAMARGO, Ana M.M.P. Comportamento das previsões de safras agrícolas do Estado de São Paulo em relação à estimativa final. São Paulo, Secretaria da Agricultura e Abastecimento. IEA, s.d. (no prelo)
31. WIESEMEYER, Jim. How USDA puts together its crop report estimates. Futures, Iowa, 12(11):50-54, nov. 1983.

R E S U M O

O trabalho tem por objetivo descrever e analisar a consistência dos levantamentos de previsões de safra de algodão, milho e soja nos Estados Unidos, no Brasil e em São Paulo, realizados respectivamente, pelo Departamento de Agricultura dos Estados Unidos (USDA), Fundação Instituto Brasileiro de Geografia e Estatística (IBGE) e Instituto de Economia Agrícola (IEA). São apresentadas as sistemáticas dos levantamentos de dados e os calendários de plantio e colheita nas regiões consideradas.

O método de análise utilizado — regressão linear simples — que associa levantamentos preliminares com o levantamento final de produção, mostrou-se adequado. Os resultados obtidos sugerem que a consistência das previsões de algodão é maior, comparativamente, que a das de soja, que por vez é maior que a das de milho. Infere-se também pelos dados globais, independentemente dos procedimentos de levantamento, que o desempenho das previsões do Brasil é melhor que o dos Estados Unidos, que por sua vez é melhor que o do Estado de São Paulo. Contudo, são feitas ressalvas quanto aos métodos utilizados.

O estudo mostra o mês de levantamento, a partir do qual, já se tem uma previsão confiável da produção. Para o algodão, nos Estados Unidos, é o do mês de agosto; no Brasil é o do mês de março; e em São Paulo, o do mês de junho. Para o milho nos Estados Unidos, a previsão de agosto, no Brasil, a de maio e em São Paulo, a de fevereiro. Para a soja nos Estados Unidos, a previsão de agosto, no Brasil a de maio e no Estado de São Paulo, a de fevereiro.

Finalizando, os autores sugerem a utilização de métodos alternati

vos para complementação dos sistemas tradicionalmente utilizados que poderiam aprimorar e agilizar as informações de previsão de safra. Neste caso são citados o emprego de satélites artificiais para sensoriamento remoto e o emprego de modelos estatísticos que considerem variáveis meteorológicas como fatores explicativos da produção. Cumpre ressaltar que os métodos atuais não devem ser substituídos, mas sim aprimorados, já que se constituem em importantes instrumentos para tomada de decisão dos usuários.

CONSISTENCY OF THE COTTON, CORN AND SOYBEAN CROP FORECASTS IN THE UNITED STATES, BRAZIL AND SÃO PAULO STATE

S U M M A R Y

The paper objectives to describe and to analyse the consistency of the forecast surveys of cotton, corn and soybeans in the United States Department of Agriculture (USDA), Fundação Instituto Brasileiro de Geografia e Estatística (IBGE) and Instituto de Economia Agrícola (IEA). The systematic of the data surveys and the sowing and harvesting seasons are introduced in the considered regions.

The analysis method used - simple linear regression - that associates preliminary surveys with the final datum of production, revealed itself appropriated. The results suggested that the consistency of the cotton forecast is bigger, comparatively, than the soybeans ones, that by their time, is bigger than the corn ones. It can also conclude through the global data, independently of the procedures of the surveys, that the performance of the Brazilian forecasts is better than the United States one, that by this time is better than the São Paulo State one. However, exemptions are made as for the used methods.

The study shows the survey since that already it has a confident forecast of the production. For the cotton, in the United States is the August forecast; in Brazil the survey of March; and in São Paulo, the June one. For the corn, in the United States, the forecast of August, in Brazil, the May one and in São Paulo, the forecast of February. For the soybeans, in the United States, the August forecast, in Brazil, the May one and in São Paulo, the February one.

Finally, the authors suggest the utilization of alternative methods to the complementation of the traditionally used systems that could improve and become dynamics the informations of crop surveys.

In this case are mentioned the use of the artificial satellites for the re mote sensing and the use of statistical models that consider meteorological variables as explanatory factors of the production. It can emphasized that the current methods cannot be replaced, but improved since constitute them selves in important ways to decision-making of the users.

**SECRETARIA DA AGRICULTURA
INSTITUTO DE ECONOMIA AGRÍCOLA**

Comissão Editorial:

Coordenador: Flavio Condé de Carvalho

Membros: Antonio Ambrósio Amaro

Arthur Antonio Ghilardi

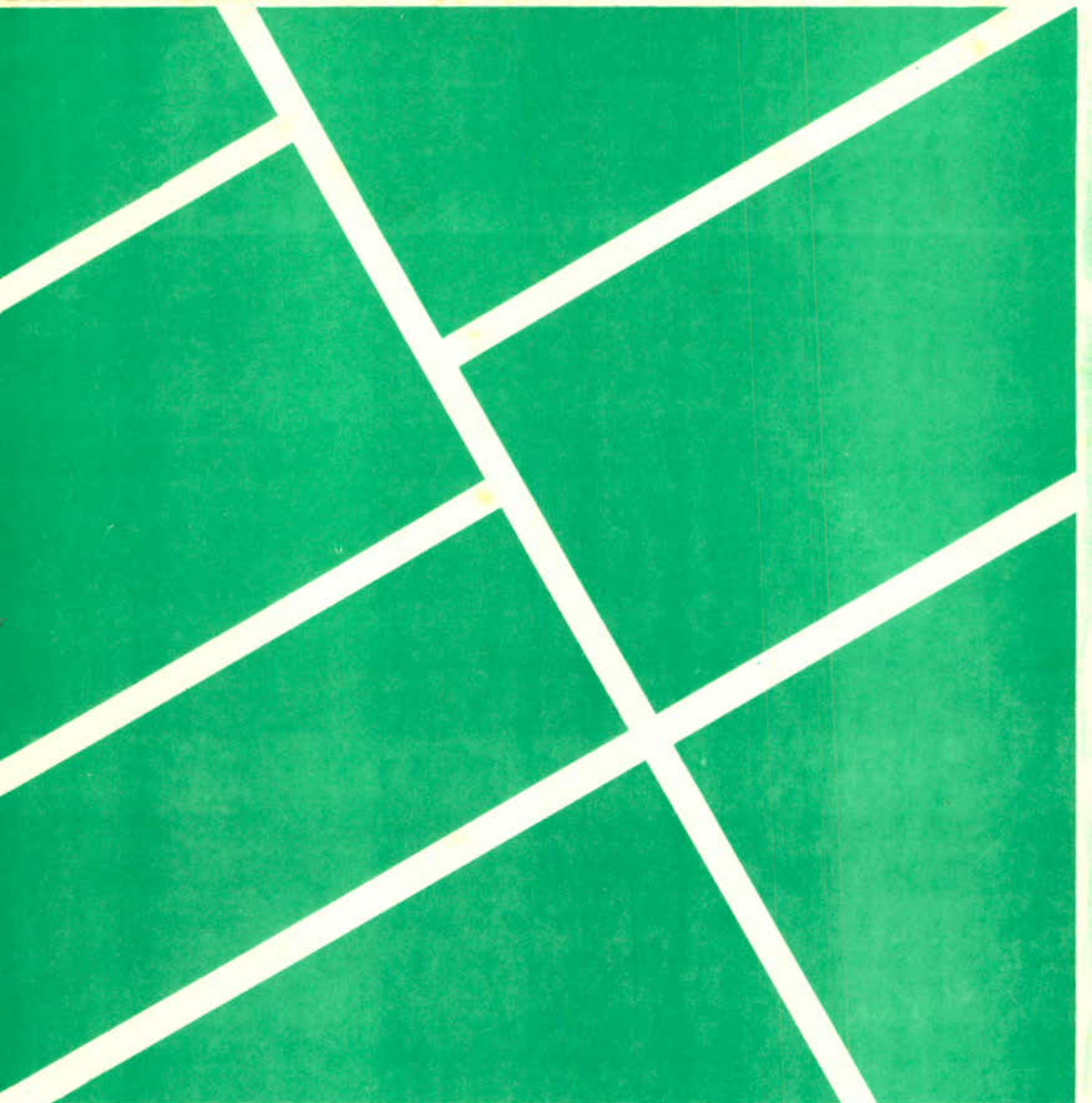
Elcio Umberto Gatti

José Luiz Teixeira Marques Vieira

Maria Carlota Meloni Vicente

Maria de Lourdes Barros Camargo

Bibliografia: Fátima Maria Martins Saldanha Faria



Relatório de Pesquisa
Nº23/87

Governo do Estado de São Paulo
Secretaria da Agricultura
Coordenadoria Sócio-Econômica

Instituto de Economia Agrícola

83

CONSISTÊNCIA DAS PREVISÕES DE SAFRA DE ALGODÃO, MILHO E SOJA NOS ES
TADOS UNIDOS, NO BRASIL E EM SÃO PAULO

Sebastião Nogueira Junior
Flavio Conde de Carvalho
Alfredo Tsunehiro

Governo do Estado de São Paulo
Secretaria da Agricultura
Coordenadoria Sócio-Econômica

Instituto de Economia Agrícola



CONSISTÊNCIA DAS PREVISÕES DE SAFRA DE ALGODÃO, MILHO E SOJA, NOS ESTADOS UNI
DOS, NO BRASIL E EM SÃO PAULO

Sebastião Nogueira Junior
Flavio Condé de Carvalho
Alfredo Tsunechiro



ÍNDICE

1 - INTRODUÇÃO	1
2 - SISTEMAS DE LEVANTAMENTO	2
2.1 - Estados Unidos	2
2.2 - Brasil	3
2.3 - São Paulo	3
3 - CALENDÁRIO AGRÍCOLA	5
3.1 - Estados Unidos	5
3.2 - Brasil	6
3.3 - São Paulo	7
4 - ANTECEDENTES	7
5 - METODOLOGIA	10
6 - RESULTADOS E DISCUSSÃO	12
6.1 - Algodão	13
6.2 - Milho	19
6.3 - Soja	22
7 - CONCLUSÕES	28
LITERATURA CITADA	31
RESUMO	34
SUMMARY	35

CONSISTÊNCIA DAS PREVISÕES DE SAFRA DE ALGODÃO, MILHO E SOJA NOS ESTADOS UNIDOS, NO BRASIL E EM SÃO PAULO (1)

Sebastião Nogueira Junior
Flavio Condê de Carvalho
Alfredo Tsunechiro

1 - INTRODUÇÃO

Previsões de safra constituem-se em importantes instrumentos para direcionar a comercialização agrícola, seja do ponto de vista do produtor, do consumidor, como também dos órgãos governamentais para a formulação de políticas de abastecimento.

A divulgação de estatísticas fidedignas de previsão de safra em tempo hábil permite ajustamentos de ordem sócio-econômica, sobretudo no que diz respeito às projeções de preço, possibilitando diminuição de riscos de perdas para os produtores e de gastos para os consumidores.

Embora importante, o campo de previsões de safra nos moldes atuais nos vários estados do Brasil, incluindo São Paulo, ainda apresenta falhas e, conseqüentemente, tem a fidedignidade de seus dados contestada, dada a predominância de informações subjetivas para complementação dos dados objetivos.

Nos Estados Unidos, por sua vez, os procedimentos objetivos, aliados a um rigoroso controle de qualidade dos dados obtidos, conferem elevado grau de credibilidade às estimativas periodicamente divulgadas, de ampla aceitação e utilização por entidades públicas e privadas, inclusive no exterior.

Resultados de previsão de safra de algodão, milho e soja são aguardados com grande expectativa, dada a importância desses produtos na alimentação humana e animal, e ainda porque deles advêm uma gama de derivados de grande utilização industrial. A escolha recaiu sobre esses produtos em função de sua importância e pela disponibilidade de séries temporais adequadas nas

(1) Os autores agradecem à Pesquisadora Científica Maria de Fátima Packer pela colaboração na escolha do método estatístico, à Bibliotecária Agurí Sawatani Negri pelo preparo das referências bibliográficas das versões preliminares deste estudo que foram apresentadas em eventos específicos (18), (19) e (20), e à escriturária Sonia Regina de Santana.

três instituições analisadas.

O estudo tem por objetivo proceder a uma descrição e análise da consistência dos levantamentos de previsão de safras de algodão, milho e soja, realizados por três instituições: Departamento de Agricultura dos Estados Unidos (USDA), Fundação Instituto Brasileiro de Geografia e Estatística (IBGE) e Instituto de Economia Agrícola (IEA).

2 - SISTEMAS DE LEVANTAMENTO

2.1 - Estados Unidos

Há mais de um século o Departamento de Agricultura dos Estados Unidos (USDA) por intermédio do Statistical Reporting Service (SRS), coleta e divulga informações agropecuárias. Cerca de 150 culturas e 50 atividades de origem animal são incluídas nos levantamentos, bem como sementes, rações, fertilizantes, mão-de-obra e preços pagos e recebidos pelos agricultores. Dessa forma cerca de 650 relatórios são divulgados anualmente (10).

Os relatórios de Safras (Crop Reports) incluem estimativas de intenção de plantio, área plantada, área colhida, produção, distribuição e estoques. Previsões de produtividade e produção são feitas mensalmente durante o ciclo vegetativo das culturas, totalizando seis levantamentos. São fundamentais em contagens e medidas realizadas em propriedades escolhidas por amostragem, nas condições de desenvolvimento da planta e nos rendimentos presumíveis por área.

Os questionários enviados aos informantes pelo correio constituem, ainda, o principal método de coleta de dados. Cerca de 25 mil questionários preenchidos retornam ao USDA e são agrupados por distritos, e uma vez tabulados e calculados, são usados para ponderação das médias estaduais ou totais, chamadas "indicações". Estas médias, de certa forma, são viesadas porque a listagem de informantes pode apresentar falhas e nem todos os questionários retornam. Entretanto, o viés é reduzido pela depuração dos dados via utilização de gráficos de regressão, quando os dados amostrais enviados são cotejados com os dados das estimativas conclusivas.

Previsões mensais e estimativas finais para várias culturas, entre as quais as do algodão, do milho e da soja, são também baseadas em campos de observação selecionados por amostragem, visando dar indicações sobre o provável rendimento da cultura quando o produto for colhido. Isto possibilita rever dados mensais, se necessário, e serve para a estimativa final. Para preservar a acurácia das informações, são feitas amostragens

adicionais, em estados produtores não principais. Estes dados são enviados diretamente a Crop Reporting Board (CRB), servindo para checar as informações referentes aos estados tradicionais produtores (31).

A importância e confiabilidade dos dados da previsão estadunidense requer um rigoroso esquema para sua divulgação. Até a hora do anúncio, pontualmente às 15 horas do dia pré-determinado do mês, apenas os quatro membros da CRB têm conhecimento prévio das cifras. Esse procedimento - precisão dos dados e pontualidade da divulgação - confere grande credibilidade às estatísticas do USDA, utilizadas em todo o mundo (14).

2.2 - Brasil.

O IBGE realiza mensalmente, desde 1973, o Levantamento Sistemático da Produção Agrícola (LSPA), que é uma pesquisa de previsão e acompanhamento das safras de 30 produtos agrícolas, dentre os quais algodão, milho e soja (2). O LSPA objetiva a elaboração de estimativas sobre área, produção e rendimento médio, desde a fase de intenção de plantio até o final da colheita, abrangendo todo o ciclo vegetativo de cada cultura.

A sistemática de coleta prevê o registro, em questionário próprio, dos dados obtidos junto aos informantes (entidades públicas e privadas) existentes nos municípios (13).

Além dos dados referentes à produção, área e rendimento, são coletadas informações sobre condições climáticas, espaçamento modal e preço pago ao produtor. Para o milho são feitos nove levantamentos mensais a nível nacional, desde a fase do plantio até o final da colheita; e para o algodão e a soja, sete levantamentos.

O ano civil é adotado como período de referência pelo fato da pesquisa constituir subsídio para a elaboração do Produto Interno Bruto (PIB) do setor agrícola, objetivando ainda a comparação com outros levantamentos agropecuários contínuos e censitários.

2.3 - São Paulo

As previsões e estimativas de safras no Estado de São Paulo são elaboradas pelo Instituto de Economia Agrícola, a partir de dados coletados pela Coordenadoria de Assistência Técnica Integral (CATI), mediante dois tipos de levantamentos: subjetivo e objetivo (por amostragem). No caso dos

(2) Para a rede de coleta a pesquisa é bimestral.

levantamentos subjetivos todos os municípios do Estado de São Paulo, através das Casas da Agricultura pertencentes à CATI, recebem periodicamente questionários contendo perguntas sobre a produção agrícola, que são preenchidos pelos agrônomos regionais da CATI e devolvidos ao IEA, onde os dados são depurados e tabulados.

Quanto aos levantamentos objetivos, para cada unidade de amostra (imóvel rural) são preenchidos questionários por enumeradores da CATI em entrevista com o produtor, administrador ou outra pessoa em condições de informar sobre o imóvel sorteado. Devolvidos os questionários para o IEA, seus dados são depurados e processados.

As principais diferenças entre os levantamentos subjetivos e objetivos do ponto de vista da execução do trabalho e do ponto de vista estatístico podem ser vistos com maiores detalhes em PINO; CAMARGO; JIMENEZ OSSIO (25).

Desde o início da década de 50 tem sido utilizada a técnica de amostragem para obtenção de estimativas a nível regional e estadual. A partir de 1973, com o objetivo de aprimoramento destas estatísticas, estabeleceu-se a meta de se obter estimativas para o Estado de São Paulo como um todo, com erros padrões que não ultrapassassem $\pm 5\%$ do valor das citadas estimativas.

Conforme CAMPOS & PIVA (4) o tamanho da amostra para procedimento da coleta de dados nesta segunda fase, foi determinado a partir do cadastro do INCRA referente ao ano de 1972, com um total de 257.955 propriedades rurais, distribuídas pelas 10 Divisões Regionais Agrícolas (DIRAs) do Estado, considerando doze estratos, de acordo com o tamanho dos imóveis com área superior a 3 hectares.

Os dados para o dimensionamento foram obtidos através de dois levantamentos, realizados em janeiro e março de 1974, com uma amostra de 6.996 propriedades (3).

A determinação do tamanho da amostra foi baseada na partilha de Neyman, a partir de cadastro único para as oito culturas mais importantes do Estado (algodão, amendoim, arroz, café, cana-de-açúcar, feijão, milho e soja). O número de levantamentos durante o ano agrícola também variou no tempo, tendo sido de três no passado e de cinco atualmente, realizados em setembro, novembro, fevereiro, abril e junho, sendo que o levantamento de junho corresponde ao dado final do ano agrícola ou safra considerado.

(3) A partir de setembro de 1977 a amostra foi reformulada e passou a ter 5.646 questionários e atualmente se situa em 3.622 imóveis rurais.

Em levantamentos por amostragem surge o problema, quase inevitável, de falta de resposta. Quando, por quaisquer razões, o percentual de falta de respostas for elevado, o levantamento pode ter seus resultados invalidados. Esse parece não ser o caso de São Paulo, porquanto a percentagem média de questionários respondidos tem sido elevada, com o menor índice de retorno situando-se em 82% na DIRA de São Paulo no período 1975-79 (24).

Aqueles mesmos autores também se propuseram a detectar a magnitude de dos erros não amostrais a partir do levantamento objetivo de abril de 1978, estabelecendo tipos, diferenças regionais e devidas aos enumeradores (23). O número encontrado (1,8 erro, em média) por questionário, embora pequeno, é suficiente para causar aumento na variância das estimativas, por que este aumento depende mais da magnitude do erro que do número de erros. Os tipos de erro mais frequentes referem-se a respostas incompletas e falta de somatório.

Foi detectado por este trabalho diferença significativa entre as DIRAs, quanto a percentagem de erro, tendo como prováveis causas a quantidade e qualificação dos enumeradores, infra-estrutura e recursos materiais; atividades agrícolas e infra-estrutura da região; extensão territorial e entrosamento IEA-CATI.

3 - CALENDÁRIO AGRÍCOLA

Na descrição dos sistemas de levantamento não tem sido mencionada a distribuição temporal das fases vegetativas das culturas analisadas. Devido à diversidade climática existente em um País ou mesmo em um Estado, em um determinado mês coexistem culturas em diferentes fases do processo de crescimento. Além disso, há algumas fases desse processo que podem proporcionar melhores indicações da produção ou produtividade final da cultura. Esses aspectos não têm sido discutidos com profundidade na literatura consultada.

3.1 - Estados Unidos

Nos Estados Unidos, o início do plantio do algodão ocorre em fevereiro, atingindo o pico em maio e estendendo-se até junho. A colheita começa em agosto, tem seu auge em outubro, encerrando-se em dezembro.

O plantio de milho inicia-se em março, atinge o pico em maio e vai até junho. A colheita inicia-se em agosto, com maior concentração em outubro e vai até fins de novembro, com pequena parcela ainda em colheita em dezembro.

A cultura da soja estadunidense é plantada de abril a junho, com a maior parte concentrando-se em maio. A colheita começa em setembro, acentuando-se em outubro e avança até dezembro (26).

Na falta de dados quantitativos sobre a distribuição do plantio e colheita nos Estados Unidos, optou-se pela utilização da distribuição da comercialização ao nível de produtores agrícolas durante o ano.

De acordo com o USDA (11) a maior parte da comercialização do algodão estadunidense na safra 1978-79 ocorreu no período outubro-janeiro com 68,6% do total; soja de outubro a fevereiro com 63,3%; e milho de outubro a abril com 62,9%.

3.2 - Brasil

No Brasil, o calendário agrícola apresenta-se de maneira inversa ao dos Estados Unidos e de todo o Hemisfério Norte.

De modo geral, os plantios das culturas analisadas cujas produções se concentram na Região Centro-Sul do Brasil (4) têm início em setembro, atingem o pico em outubro e avançam até novembro (algodão) e dezembro (milho), enquanto o da soja inicia-se em outubro, com maior concentração em novembro e avança até dezembro.

As colheitas por sua vez iniciam com algodão e milho em fevereiro, soja em março, atingem o auge em março-abril (algodão e soja) e abril-maio (milho) e estendem-se até junho (algodão e soja) e julho (milho) (12).

Em termos de distribuição quantitativa, de acordo com o CENSO AGROPECUÁRIO do Brasil de 1980 (6), a maior parte da área cultivada com algodão herbáceo teve seu plantio realizado no período setembro-outubro (61%), cifra relativa à Região Centro-Sul e 24% de janeiro a março (Região Norte-Nordeste). A colheita ocorreu praticamente de fevereiro a maio (84%).

Para milho, a grande concentração do plantio (81%) ocorreu em agosto a dezembro, e a quantidade colhida foi centralizada de março a julho (89%). A colheita foi quase totalmente realizada (89%) entre fevereiro e maio.

A soja, lavoura quase exclusiva da Região Centro-Sul teve, seu plantio totalmente concentrado de setembro a dezembro e a colheita de fevereiro a maio.

Dada a concentração da produção na Região Centro-Sul, para efeito de tomada de decisão as previsões de safra para algodão efetuadas em julho

(4) No caso do algodão herbáceo, o Centro-Sul responde por mais de 90% na produção brasileira; milho por 92% e soja por 99,5%.

lho e agosto; para milho de agosto a dezembro; e para soja em junho e julho não se constituem em parâmetros relevantes. Sua não exclusão da análise se justifica pela necessidade de aferição do procedimento dos levantamentos.

3.3 - São Paulo

O plantio e colheita de algodão, milho e soja no Estado de São Paulo, praticamente tem a mesma distribuição que a verificada para a Região-Sul.

O algodão tem seu plantio concentrado em setembro e outubro; milho e soja em outubro e novembro. A colheita por sua vez tem a seguinte distribuição: algodão (março a maio); milho (abril a junho) e soja (março a abril) (21).

Quanto à distribuição relativa, de acordo com o CENSO AGROPECUÁRIO de São Paulo de 1980 (7), pode-se observar que o algodão teve praticamente todo o plantio (98%) realizado de setembro a novembro e a colheita, também 98%, de fevereiro a maio.

O milho, por sua vez, com uma distribuição temporal mais extensa, teve 92% do plantio efetivado de agosto a novembro, sendo o mesmo raciocínio válido para a colheita: 92% ocorreu em 5 meses, ou seja, de fevereiro a junho.

Finalmente, a soja mostra concentração mais forte tanto na semeadura quanto na quantidade colhida: respectivamente 96% em outubro-novembro e 95% de fevereiro a abril.

4 - ANTECEDENTES

No Brasil, trabalhos que versam sobre avaliação de previsões de safra, na maioria utilizam estatísticas divulgadas pelo Instituto de Economia Agrícola (IEA) ajustando os dados via regressões lineares simples e foram realizados na sua quase totalidade por pesquisadores dessa instituição.

BRANDT; ARAUJO; SERRANO (2) realizaram pesquisa sobre a precisão e fidedignidade das previsões de safra de algodão no período de 1947/48 a 1963/64, a partir de dados levantados pelo IEA. Seus principais resultados podem ser assim formulados:

- a) na terceira previsão de safra, em junho, existe uma tendência para subestimar as safras menores que 600 mil toneladas e para superestimar as maiores que 600 mil toneladas, sendo tanto maior este viés quanto maior o distanciamento do volume de safra do valor mencionado;

b) as duas primeiras previsões, dezembro e março, são mais fidedignas que a terceira, não havendo explicação para tal com base nos conhecimentos disponíveis;

c) na medida em que os operadores de mercado de algodão tomam conhecimento e se utilizam destas informações sobre previsão de safra, os erros aleatórios verificados nas duas primeiras previsões não são suficientemente "grandes" para forçar alterações exageradas dos preços.

PEETZ & AMARO (22) avaliaram a precisão das previsões de safra de laranja realizadas no Estado de São Paulo, pelo IEA e Coordenadoria de Assistência Técnica Integral (CATI), da Secretaria da Agricultura, e, na inexistência de um controle da previsão, foi considerado o levantamento final (novembro) como o volume real colhido. Os diversos levantamentos da safra de laranja foram relacionados entre si, dois a dois, através de funções linear e exponencial, ajustadas segundo um modelo de regressão linear simples. As variáveis, em ambas as formas, apresentaram elevado coeficiente de correlação e os testes dos coeficientes de regressão apresentaram significância ao nível de 0,5%. Os resultados dos ajustamentos da função linear foram considerados como os mais adequados para os objetivos propostos. Conclui-se ser possível projetar com bom grau de precisão a produção final de laranja a partir dos dados de 2º, 3º, 4º e 5º levantamentos, permitindo sua utilização como indicadores para decisões sobre a comercialização e política governamental para o setor.

CARVALHO et alii (5) repetiram e atualizaram estudo já anteriormente realizado para algodão por BRANDT; ARAUJO; SERRANO (2). Foram comparadas as três previsões realizadas durante o ano agrícola, individualmente, com a entrada nas usinas. Constataram que a segunda e terceira equações apresentaram as constantes de regressão não estatisticamente diferentes de zero, e, da mesma forma, os coeficientes de regressões não diferentes da unidade, preenchendo ambos os requisitos de fidedignidade exigidos. A terceira previsão, como era esperado, apresentou o melhor comportamento, com o coeficiente de determinação mais elevado..

CAMARGO (3), com a finalidade de conhecer as relações existentes entre as previsões e estimativa final para o café, utilizou também regressão simples e conclui que as três equações oferecem valores não tendenciosos da produção de café, sendo todas elas adequadas à formulação de políticas e tomada de decisão pelo empresário e pelo governo, sobretudo já a partir da 2ª previsão, uma vez que a divulgação da 3ª previsão ocorre em época tardia, tornando restrito seu uso.

VICENTE; CASER; CAMARGO (30) avaliaram, para o período de 1969-83, o comportamento das previsões das safras paulistas de 21 produtos em relação

ã estimativa final, considerando a área cultivada e a produtividade.

Para o algodão pode-se prever a área com antecedência, utilizando-se os levantamentos de setembro, novembro, fevereiro e abril. No caso da produtividade, os levantamentos de fevereiro e abril permitem que se conheça o rendimento final com nível de significância de 1%.

No caso do milho os levantamentos de setembro, fevereiro e abril mostraram-se adequados para prever a área final, e o de fevereiro e abril foram satisfatórios no que diz respeito à previsão de produtividade.

Finalmente, com respeito à soja, os levantamentos (exceto setembro) mostraram-se adequados para prever a área final e os de fevereiro e abril foram satisfatórios quanto à produtividade.

Além desses trabalhos que averiguaram de maneira clássica a precisão das previsões de safra, utilizando como regressores séries históricas de levantamentos sistemáticos, existem várias tentativas de previsão de rendimentos para uma série de culturas, utilizando-se de variáveis climáticas. Nesse caso merecem destaque estudos realizados por pesquisadores do Instituto de Economia Agrícola que consideraram o balanço hídrico como variável explicativa para as variações de rendimento, para uma gama de culturas do Estado de São Paulo.

SILVA; CASER; VICENTE (27) mensuraram o efeito das condições do tempo sobre a produtividade agropecuária, utilizando-se da análise de regressão para estimar as relações entre deficiência hídrica, geada e produtividade da terra. Os resultados mostraram haver grande sensibilidade das culturas anuais e perenes à deficiência de água, no período outubro-abril e setembro-março, respectivamente. As culturas perenes mostraram-se também muito susceptíveis à ocorrência de geadas.

SILVA; CASER; VICENTE (28) formularam e testaram modelos capazes de prever a produtividade da cultura do café no Estado de São Paulo conseguindo resultados animadores. A aplicação do método fornece previsões antecipadas em relação às baseadas em levantamentos de campo e, em alguns casos, melhores que estes últimos.

SILVA; VICENTE; CASER (29) reuniram um conjunto de estudos, focalizando o problema das previsões de safras agrícolas, pelo lado dos efeitos do tempo sobre o rendimento de importantes culturas, no Estado de São Paulo. Este esforço, construindo e testando modelos que utilizam variáveis meteorológicas, preenche uma lacuna existente no Instituto de Economia Agrícola (IEA) no que diz respeito à previsão de rendimento, em que pese o pioneirismo e tradição do IEA no desenvolvimento e utilização de modelos de previsão de área cultivada.

As condições meteorológicas foram incluídas nos modelos básicos

mente através de uma variável derivada, a deficiência hídrica. Para o cálculo da deficiência hídrica foi utilizado o método denominado de balanço hídrico, que consiste em contabilizar a água no solo, num processo em que a chuva representa o abastecimento e a evapotranspiração o consumo de água no solo, em condições de ser extraída pela planta.

Os resultados obtidos possibilitam concluir que o algodão pode ter seu rendimento previsto ao final dos meses de fevereiro e março com resultados melhores e com certa antecipação em relação aos baseados em levantamentos de campo, efetuados nos meses de fevereiro e abril pelo IEA-CATI, e que devido ao processo normal de coleta, processamento, depuração e análise, levam algum tempo até tornarem-se disponíveis: o de fevereiro é geralmente publicado em abril e o de abril em junho. Para o milho obtve-se modelos para previsão do rendimento nos meses de janeiro, fevereiro e março, este último com melhores ajustes e erros menores que a previsão de campo. Com a soja não se conseguiu modelos melhores que as previsões de campo, porém os erros percentuais médios de modelos terminados no mês de abril da ordem de 7%, confirmam a boa qualidade dos ajustes; mesmo quando o período é reduzido para fevereiro, o erro médio é inferior a 10%.

Convém citar ainda, trabalho desenvolvido por MASCARENHAS (16) que permite estimar a produtividade de soja com base nas precipitações pluviais ocorridas durante períodos relacionados ao florescimento. Esse método possibilita que a estimativa seja feita com até dois meses e meio de antecedência em relação aos métodos diretos de levantamento no campo.

Finalmente, cabe relatar os métodos utilizados pela CEPLAC (8) para a previsão de safra de cacau: contagem de frutos, amostragem e avaliação subjetiva. O método baseado na contagem de frutos, apesar do custo, mostrou-se mais funcional com relação aos resultados obtidos para a safra temporã do ano agrícola 1983/84, com erro de apenas 1,17% em relação à produção final, comparado ao erro de 17,4% da previsão por amostragem e de 20,35% na utilização do critério subjetivo.

5 - METODOLOGIA

A análise da função que associa os levantamentos preliminares com o levantamento final de produção de algodão, milho e soja, para cada instituição considerada, envolve o ajustamento de um modelo de regressão linear simples, representado por:

$$Y_j = \alpha_i + \beta_i X_{ij} + M_{ij}$$

onde:

Y_j é estimativa final da produção do ano j , com j variando de 1 a n , conforme número de anos considerados; X_{ij} é a i -ésima previsão de produção, com i variando de 1 a K , conforme número de levantamento de cada instituição, no ano j ; α_i e β_i são os parâmetros de regressão; e M_{ij} é o erro aleatório.

Os parâmetros α_i e β_i serão estimados pelo método dos mínimos quadrados, descritos em HOFFMANN & VIEIRA (13).

Para o intercepto, serão testadas as seguintes hipóteses:

$$H_0 : \alpha_i = 0$$

$$H_A : \alpha_i \neq 0$$

Para o coeficiente de regressão, serão testadas as seguintes hipóteses:

$$H_0 : \beta_i = 1$$

$$H_A : \beta_i \neq 1$$

Os testes de hipóteses serão efetuadas utilizando-se a distribuição t de Student, aos níveis de 1% e 5% significância.

Uma equação de regressão fidedigna apresenta $\alpha_1 = 0$ e $\beta_1 = 1$. Existe técnica estatística apropriada ao teste conjugado dessas hipóteses. Entretanto, devido ao fato de que o conhecimento isolado de cada teste de hipótese permite fazer inferências sobre vieses introduzidos na revisão dos dados, conforme BRANDT; ARAUJO; SERRANO (2), os testes serão efetuados separadamente.

O procedimento proposto possibilita verificar a coerência entre as previsões, as quais podem estar sub ou superestimadas. De antemão sabe-se que a última informação, por uma questão de lógica deva ser a escolhida para a estimativa de safra de um produto qualquer. Por motivos técnicos, entretanto, o seu uso é mais restrito, em virtude de sua divulgação dar-se em época já tardia para tomadas de decisão, seja a nível privado ou governamental. Assim, torna-se interessante verificar qual levantamento, em cada órgão, deve ser mais convenientemente aprimorado e utilizado, e qual o nível de previsibilidade de cada levantamento.

Embora haja maior probabilidade de acerto à medida que se aproxima o período final de colheita, mesmo no caso de levantamentos subjetivos, conforme enfatiza MUELLER (17), convém lembrar que apenas o algodão permite uma aferição relativamente mais confiável ou precisa dada a possibilidade de cotejo com a entrada nas usinas de beneficiamento. Neste caso existem dados disponíveis para os Estados Unidos e para São Paulo (5).

(5) O Estado do Paraná também possui estatísticas completas sobre entrada de algodão nas usinas de beneficiamento (1).

O problema da correlação serial nos resíduos da equação ajustada será examinado com o emprego de teste de Durbin-Watson, complementado pela estatística de Theil-Nagar.

Para São Paulo, no milho e na soja, serão utilizadas duas previsões de safra (fevereiro e abril) e o dado final (junho) nos anos em que houver previsões de safra em épocas diferentes ou em número maior que o normal, a colocação dos dados em suas respectivas séries será feita mediante critério subjetivo. Os dados se referem ao período 1947-85. Essas estatísticas não se encontram publicadas, de forma definitiva, necessitando-se de um trabalho prévio de consolidação dos mesmos. Para algodão serão utilizadas três previsões (fevereiro, abril e junho) sendo considerado como dado final a entrada de algodão em caroço nas usinas de beneficiamento, conforme levantamento da CATI.

Na parte relativa à entrada nas usinas, erros podem ocorrer nas anotações das quantidades e especificação da procedência. Por exemplo, algodão de outros estados pode ser declarado pelo usineiro como de procedência paulista, por razões de ordem fiscal ou creditícia (5).

Para efeito de comparação também serão apresentadas as equações de regressão tendo como variável dependente a última previsão (estimativa de junho).

Para o Brasil, os dados encontram-se publicados no Levantamento Sistemático de Produção Agrícola (1974 a 1986) do IBGE (15). São oito levantamentos mensais, para milho (maio a dezembro); seis para algodão herbáceo (março a agosto) e seis para soja (janeiro a junho), além da estimativa final para as três culturas. O período de disponibilidade dos dados é de 1974/75 a 1984/85.

Para os Estados Unidos, para milho e soja, serão utilizadas as previsões do USDA de agosto a dezembro, além da estimativa final, do COMMODITY YEAR BOOK (9). A série utilizada cobre o período 1947/48 a 1984/85. Para algodão, as previsões referem-se ao período agosto-janeiro, tendo sido utilizado o dado de entrada nas usinas como dado definitivo.

Igualmente, a exemplo de São Paulo, para efeito de comparação serão apresentadas as equações de regressão tendo como variável dependente a última previsão (estimativa de junho).

6 - RESULTADOS E DISCUSSÃO

Dadas as diferenças entre os procedimentos de obtenção dos dados estatísticos de produção adotados pelas instituições analisadas, torna-se

difícil comparar os resultados das equações de regressão ajustadas no presente estudo.

Levantamentos baseados em opiniões de pessoas, ligadas a entidades públicas ou particulares, como os efetuados pelo IBGE, estão sujeitos à tendenciosidade, na medida em que aquelas opiniões reflitam enfoques subjetivos dos participantes. Além disso, podem apresentar maior rigidez no ajustamento às condições da realidade produtiva, dada a tendência à irredutibilidade de na reavaliação de estimativas pelos participantes do grupo de informantes. Nesse sentido, o elevado grau de ajuste das equações pode refletir não um alto grau de qualidade do levantamento, mas sim um alto grau de apego aos números anteriormente estabelecidos. Na inexistência de um levantamento objetivo que possibilite o confronto com os dados subjetivos, não há como avaliar, a contento, a acurácia das previsões.

Por outro lado, a utilização de um método objetivo, como aquele do levantamento do IEA-CATI, defronta-se com problemas de outra ordem. A qualidade dos levantamentos, quando realizados por enumeradores não constituintes de uma equipe homogênea, deixa a desejar. O elevado número de questionários não preenchidos, incompletos ou portadores de erros diversos, torna necessário uma rigorosa revisão dos mesmos, podendo exigir o retorno ao campo para complementação, o que, normalmente, não tem ocorrido. A subjetividade pode estar presente no preenchimento de questionários, influenciando as respostas, principalmente quando o entrevistador está ligado, direta ou indiretamente, à atividade de produção agropecuária. Experiências pessoais tendem a ser generalizadas, em prejuízo da qualidade do levantamento.

Convém ressaltar que os dados primários obtidos no campo passam por um processo exaustivo de depuração e consistência interna antes de serem tabulados, conforme relatam PINO; CAMARGO; JIMENEZ OSSIO (25).

Feitas essas ressalvas, passa-se ao exame e discussão dos resultados obtidos por produto.

6.1 - Algodão

Foram seis as equações de regressão linear simples ajustadas aos valores observados das previsões de safra de algodão nos Estados Unidos, tendo como variável dependente a entrada de algodão em caroço nas usinas.

Todas as equações foram significantes ao nível de 1%, como indica o pelos valores do teste F de Snedecor (quadro 1). O poder explicativo das regressões, medido pelo coeficiente de determinação (R^2), foi relativamente elevado em todas as equações, com crescimento ao longo do tempo. O teste de Durbin-Watson aponta problema de correlação serial dos resíduos nas equa

QUADRO 1. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Caroço, Tendo como Variável Dependente a Entrada de Algodão em Caroço nas Usinas de Beneficiamento, Estados Unidos, 1947/ 48 a 1984/85 ⁽¹⁾

Item	Previsão					
	Agosto	Setembro	Outubro	Novembro	Dezembro	Final
Constante de regressão (b_0)	-364,39	-206,83	-163,40	-122,52	-25,98	-18,52
Coefficiente de regressão (b_1)	1,113	1,057	1,040	1,024	0,994	0,993
Desvio-padrão						
s_{b_0}	207,93	195,59	129,65	73,83	37,52	30,08
s_{b_1}	0,073	0,069	0,045	0,026	0,013	0,010
Razão t de Student						
t_{b_0}	-1,753	-1,058	-1,261	-1,660	-0,693	-0,616
t_{b_1}	1,547	0,826	0,888	0,923	0,461	0,700
Teste F de Snedecor	226,90 ^a	231,60 ^a	513,19 ^a	1.541,19 ^a	5.586,71 ^a	8.644,30 ^a
Coefficiente de determinação (R^2)	0,8630	0,8654	0,9344	0,9771	0,9935	0,9958
Estatística de Durbin-Watson	1,87	1,48	1,63	1,87	1,16 ^b	0,32 ^b
Coefficiente de Theil-Nagar	0,061	0,257	0,181	0,062	0,416	0,837

⁽¹⁾ O número de observações (n) é igual a 38. As letras representam o nível de significância a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do United States Department of Agriculture (9).

ções referentes a dezembro e final, o que se reflete nos valores do coeficiente de Theil-Nagar.

O teste da hipótese nula referente à constante de regressão (b_0) não foi estatisticamente significativa, mesmo ao nível de 5%, para nenhuma das equações ajustadas, conforme demonstrado pela razão t de Student, não se rejeitando, desse modo, a hipótese de que a constante de regressão é estatisticamente igual a zero.

O coeficiente de regressão (b_1) não foi estatisticamente diferente da unidade, ao nível de significância de 5%, nas equações ajustadas, conforme os valores obtidos da razão t de Student.

A apreciação conjunta dos resultados obtidos para a constante de regressão e para o coeficiente de regressão permite admitir a hipótese de fidedignidade das equações ajustadas, qualquer que seja o mês considerado, quando usadas na previsão da entrada de algodão nas usinas estadunidenses. Também nas regressões em que a estimativa final (janeiro) foi considerada como variável dependente os resultados estatísticos mostraram-se satisfatórios (quadro 2).

Para o Brasil, as seis equações de regressão ajustadas aos valores observados das variáveis, apresentaram-se estatisticamente significantes ao nível de 1%, conforme indicado pela estatística F de Snedecor (quadro 3). O poder explicativo das equações (R^2) mostrou-se relativamente elevado e crescente ao longo dos meses. Não se constatou problema de correlação serial dos resíduos em grau elevado, conforme estatística de Durbin-Watson e o coeficiente de Theil-Nagar.

Nos testes relacionados às constantes de regressão e aos coeficientes de regressão, as hipóteses nulas não foram rejeitadas ao nível de significância de 5%. Constatou-se, assim que, em todas as equações ajustadas, as constantes e os coeficientes de regressão não diferem do ponto de vista estatístico, de zero e da unidade, respectivamente.

Esses resultados sugerem que as equações de regressão entre previsões mensais da safra de algodão no Brasil e a estimativa final podem ser utilizadas para prover previsões fidedignas da estimativa final de uma safra.

Ao se analisar as equações referentes às previsões de safra de algodão em caroço no Estado de São Paulo, constata-se a significância geral das equações, conforme o teste F de Snedecor (quadro 4).

O poder explicativo das regressões é bastante reduzido para as previsões de fevereiro e abril, elevando-se acentuadamente na previsão de junho. As estatísticas de Durbin-Watson e os coeficientes de Theil-Nagar não revelam a existência de grau elevado de autocorrelação serial nos resí

QUADRO 2. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carço, Tendo como Variável Dependente a Estimativa Final, Estados Unidos, 1947/48 a 1984/85 ⁽¹⁾

Item	Previsão				
	Agosto	Setembro	Outubro	Novembro	Dezembro
Constante de regressão (b_0)	-325,46	-174,49	-132,78	-99,66	-6,53
Coefficiente de regressão (b_1)	1,112	1,058	1,042	1,029	1,000
Desvio-padrão					
s_{b_0}	215,61	200,25	134,62	73,11	25,65
s_{b_1}	0,076	0,071	0,47	0,025	0,009
Razão t de Student					
t_{b_0}	-1,510	-0,872	-0,987	-1,364	-0,255
t_{b_1}	1,470	0,828	0,897	1,157	0,038
Teste F de Snedecor	210,70	221,60 ^a	477,95 ^a	1.587,46 ^a	12.101,49 ^a
Coefficiente de determinação (R^2)	0,8540	0,8602	0,9299	0,9778	0,9970
Estatística de Durbin-Watson	1,88	1,49	1,59	1,84	1,95
Coefficiente de Theil-Nagar	0,059	0,252	0,201	0,078	0,023

⁽¹⁾ O número de observações (n) é igual a 38. A letra a representa o nível de significância de 1%.

Fonte: Elaborado a partir de dados básicos do United States Department of Agriculture (9).

QUADRO 3. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carvão, Brasil, 1974/75 a 1984/85 (1)

Item	Previsão					
	Março	Abril	Mai	Junho	Julho	Agosto
Constante de regressão (b_0)	166,523	-7.284	-79.498	-59,346	7.259	3.158
Constante de regressão (b_1)	0,908	1,016	1,060	1,051	1,005	1,013
Desvio-padrão						
s_{b_0}	195.964	199.882	123.244	80.033	60.571	54.699
s_{b_1}	0,123	0,125	2,774	0,050	0,037	0,034
Razão t de Student						
t_{b_0}	0,849	-0,037	-0,646	-0,742	0,119	0,057
t_{b_1}	0,747	0,128	0,077	1,020	0,135	0,382
Teste F de Snedecor	54,30 ^a	65,06 ^a	187,32 ^a	434,87 ^a	702,57 ^a	865,90 ^a
Coefficiente de determinação (R^2)	0,8578	0,8784	0,9541	0,9797	0,9873	0,9897
Estatística de Durbin-Watson (2)	1,90	2,27	2,46	2,07	2,75	2,65
Coefficiente de Theil-Nagar	0,047	-0,137	-0,236	-0,040	-0,378	-0,330

(1) O número de observações (n) é igual a 11. A letra a representa o nível de significância de 1%.

(2) O valor da Tabela para comparação só é disponível para n=15.

Fonte: Elaborado a partir de dados básicos da Fundação IBGE (15).

QUADRO 4. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carço, Tendo como Variável Dependente a Entrada de Algodão em Carço nas Usinas de Beneficiamento, Estado de São Paulo, 1947/48 a 1984/85 (1)

Item	Previsão		
	Fevereiro	Abril	Junho
Constante de regressão (b_0)	282,41	172,34	-6,35
Coefficiente de regressão (b_1)	0,493	0,719	1,042
Desvio-padrão			
s_{b_0}	66,50	60,72	40,66
s_{b_1}	0,404	0,106	0,072
Razão t de Student			
t_{b_0}	4,246 ^a	2,837 ^a	-0,157
t_{b_1}	1,254	2,650 ^b	0,583
Teste F de Snedecor	19,96 ^a	45,21 ^a	208,59 ^a
Coefficiente de determinação (R^2)	0,3842	0,5855	0,8669
Estatística de Durbin-Watson	2,04	2,03	1,50
Coefficiente de Theil-Nagar	-0,024	-0,019	0,246

(1) O número de observações (n) é igual a 34. As letras representam o nível de significância: a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola e da Coordenadoria de Assistência Técnica Integral.

duos das equações.

O teste da hipótese nula de que a constante de regressão (b_0) é igual a zero levou à rejeição dessa hipótese nas equações relativas a fevereiro e abril, ao nível de 1% de significância. Na previsão de junho, essa hipótese não foi rejeitada.

O teste da hipótese nula de que o coeficiente de regressão (b_1) é igual à unidade levou à rejeição dessa hipótese ao nível de significância de 5%, na previsão de abril.

A apreciação global dos resultados desses testes de significância indica que apenas as equações da previsão de junho podem ser consideradas fidedignas na previsão de valores de entrada de algodão em caroço nas usinas de beneficiamento paulistas.

Para São Paulo, também foram ajustadas equações de regressão tendo como variável dependente o levantamento final (estimativa de junho) verificando a aferição dos dados das instituições responsáveis pelos levantamentos — IEA e CATI.

O poder explicativo das duas regressões é maior que o das equações correspondentes quando considerada a entrada nas usinas como variável dependente (quadro 5). Entretanto, os testes das hipóteses nulas, tanto para a constante (b_0) quanto para o coeficiente (b_1) de regressão, foram significativos ao nível de 1% revelando a não fidedignidade dos levantamentos. Por conseguinte, infere-se que as estimativas preliminares são menos consistentes quando se usa a estimativa final como variável dependente do que quando são usadas as entradas de algodão nas usinas de beneficiamento.

6.2 - Milho

As equações de regressão linear simples ajustadas a dados de previsão de safra de milho dos Estados Unidos apresentaram-se estatisticamente significantes ao nível de 1%, conforme indicado pelos valores da estatística F de Snedecor (quadro 6). O poder explicativo da regressão (R^2) foi bastante elevado em todas as equações. A correlação serial nos resíduos não se destacou como problema, a não ser na equação referente ao nível de 5%, o que é corroborado pelos valores do coeficiente de Theil-Nagar.

A hipótese nula $b_1 = 1$, com o emprego da razão t de Student, foi rejeitada, ao nível de, pelo menos 5%, nas equações referentes aos meses de setembro, outubro, novembro e dezembro. Além disso, a constante de regressão (b_0) foi estatisticamente diferente de zero nas equações de outubro e novembro, ao nível de, pelo menos, 5%.

Na medida em que um elevado grau de fidedignidade de uma previ

QUADRO 5. - Principais Características das Equações de Regressão Linear Simples de Previsão de Safras de Algodão em Carço, Tendo como Variável Dependente a Estimativa Final, São Paulo, 1947/48 a 1984/85 (1)

Item	Previsão	
	Fevereiro	Abril
Constante de regressão (b_0)	259,30	158,92
Coefficiente de regressão (b_1)	0,498	0,706
Desvio-padrão		
s_{b_0}	52,81	43,69
s_{b_1}	0,087	0,076
Razão t de Student		
t_{b_0}	4,909 ^a	3,637 ^a
t_{b_1}	5,722 ^a	3,810 ^a
Teste F de Snedecor	32,31 ^a	84,39 ^a
Coefficiente de determinação (R^2)	0,5024	0,7250
Estatística de Durbin-Watson	2,27	2,48
Coefficiente de Theil-Nagar	-0,138	-0,242

(1) O número de observações (n) é igual a 34. A letra a representa o nível de significância de 1%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

QUADRO 6. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Milho, Estados Unidos, 1947/48 a 1984/85 (1)

Item	Previsão				
	Ago.	Set.	Out.	Nov.	Dez.
Constante de regressão (b_0)	-4.584	-4.784	-4.767	-3.881	-1.794
Coefficiente de regressão (b_1)	1,055	1,059	1,057	1,044	1,019
Desvio-padrão					
s_{b_0}	4.756	2.768	2.021	1.280	909
s_{b_1}	0,038	0,022	0,016	0,010	0,007
Razão t de Student					
t_{b_0}	-0,964	-1,729	-2,359 ^b	-3,032 ^a	-1,975
t_{b_1}	1,429	2,657 ^b	3,505 ^a	4,357 ^a	2,703 ^b
Teste F de Snedecor	751,62 ^a	2.231,87 ^a	4.187,78 ^a	10.311,41 ^a	19.836,06 ^a
Coefficiente de determinação (R^2)	0,9542	0,9841	0,9914	0,9965	0,9981
Estatística de Durbin-Watson	1,679	1,806	1,728	1,508	0,693 ^a
Coefficiente de Theil-Nagar	0,160	0,096	0,135	0,245	0,653

(1) O número de observações (n) é igual a 38. As letras representam o nível de significância: a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do United States Department of Agriculture (9).

são de safra ocorra somente na presença simultânea de constante não diferente, estatisticamente, de zero e coeficiente de regressão não diferente, estatisticamente, da unidade, tem-se que as equações de regressão linear simples ajustadas para as previsões de safra de milho, nos Estados Unidos, para os meses do último quadrimestre do ano, não atendem a esses requisitos. Esses são observados, apenas, na equação do mês de agosto. As demais equações podem dar origem a previsões dotadas de maior grau de incerteza.

Para o Brasil tem-se que as oito equações de regressão ajustadas apresentaram-se estatisticamente significantes ao nível de 1% conforme indicado pela estatística F de Snedecor (quadro 7). O poder explicativo da regressão (R^2), especialmente para os meses de julho a novembro, foi bastante elevado, notando-se uma queda no valor do coeficiente de determinação em dezembro. O teste de Durbin-Watson, utilizado para o exame de correlação serial nos resíduos, mostrou-se inconclusivo para os meses de agosto, novembro e dezembro, não sendo significativo nos demais meses.

Testando-se a hipótese nula $b_1 = 1$, a razão t de Student calculada não leva à rejeição dessa hipótese em nenhuma das equações ajustadas, ao nível de 5%. A esse mesmo nível de significância, a constante de regressão (b_0) não foi estatisticamente diferente de zero em nenhuma das oito equações de regressão ajustadas.

Desse modo, as equações de regressão de previsão de safra de milho, para o Brasil, podem ser consideradas como satisfazendo as condições necessárias da estimativa final de produção desse cereal.

Ao se analisar as duas equações de regressão linear simples ajustadas para a previsão de safra de milho no Estado de São Paulo tem-se que em função dos valores calculados para a estatística F de Snedecor, ambas apresentaram-se estatisticamente significantes ao nível de 1% (quadro 8). O poder explicativo (R^2) das duas equações foi bastante elevado. O teste de Durbin-Watson apresentou resultado positivo quanto à presença de correlação serial nos resíduos, ao nível de 5% para o mês de abril.

O coeficiente de regressão b_1 mostrou-se significativamente diferente de unidade, ao nível de 1%, para a equação de abril. O intercepto b_0 , em ambas as equações de regressão, não se mostrou estatisticamente diferente de zero, mesmo ao nível de significância de 5%.

Com esses resultados, a equação de regressão ajustada para o mês de abril não apresenta condições para que ocorra um elevado grau de fidedignidade, se usada para prever safra de milho no Estado de São Paulo.

6.3 - Soja

Para os Estados Unidos, foram ajustadas cinco equações de regres

QUADRO 7. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Milho, Brasil, 1974/75 a 1984/85 (1)

Item	Previsão							
	Mai.	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
Constante de regressão (b_0)	-761.976	-690.660	-508.807	-615.933	-254.474	-136.838	9.270	-337.534
Coefficiente de regressão (b_1)	1.000	1,016	1,016	1,028	1,008	1,004	0,998	1,030
Desvio-padrão								
s_{b_0}	2.247.275	1.331.054	693.871	704.871	175.141	103.227	95.232	1.869.133
s_{b_1}	0,113	0,068	0,035	0,036	0,009	0,005	0,004	0,099
Razão t de Student								
t_{b_0}	-0,340	-0,519	-0,734	-0,874	-1,453	-1,326	0,097	-0,181
t_{b_1}	0,001	0,235	0,455	0,771	0,897	0,760	0,343	0,306
Teste F de Snedecor	78,14 ^a	221,41 ^a	800,39 ^a	784,01 ^a	12.244,32 ^a	34.824,32 ^a	40.305,01 ^a	108,27 ^a
Coefficiente de Determinação (R^2)	0,8967	0,9609	0,9888	0,9886	0,9992	0,9997	0,9997	0,9232
Estatística de Durbin-Watson ⁽²⁾	2,238	2,291	2,734	2,906 IN	2,532	2,351	1,194 IN	1,172 IN
Coefficiente de Theil-Nagar	0,120	-0,147	-0,369	-0,456	-0,268	-0,177	0,404	0,415

(1) O número de observações (n) é igual a 11. A letra a representa o nível de significância de 1%. IN = inconclusivo.

(2) Os valores calculados foram cotejados com os da Tabela para 15 observações, que é o menor número disponível.

Fonte: Elaborado a partir de dados básicos da Fundação IBGE (15).

QUADRO 8. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Milho, Estado de São Paulo, 1947/48 a 1984/85 (1)

Item	Previsão	
	Fevereiro	Abril
Constante de regressão (b_0)	-2,994	27,487
Coefficiente de regressão (b_1)	0,946	0,951
Desvio-padrão		
s_{b_0}	101,456	41,309
s_{b_1}	0,042	0,017
Razão t de Student		
t_{b_0}	-0,030	0,665
t_{b_1}	1,246	2,743 ^a
Teste F de Snedecor	490,305 ^a	2.898,476 ^a
Coefficiente de determinação (R^2)	0,9387	0,9890
Estatística de Durbin-Watson	1,761	0,989 ^b
Coefficiente de Theil-Nagar	0,119	0,505

(1) O número de observação (n) é igual a 34; as observações referentes aos anos 1954/55, 1956/57, 1957/58 e 1960/61 foram eliminadas, por incompletas. As letras representam o nível de significância: a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

são linear simples, tendo as previsões dos meses de agosto a dezembro como variáveis independentes e a estimativa final como variável dependente.

Todas as equações foram significantes ao nível de 1%, como indicado pelos elevados valores do teste F de Snedecor (quadro 9). O poder explicativo das regressões, medido pelo coeficiente de determinação (R^2), foi bastante elevado para todas as previsões. O teste de Durbin-Watson de correlação serial nos resíduos apresentou-se inconclusivo na equação de agosto, o que é corroborado pelo coeficiente de Theil-Nagar.

O teste de hipótese nula referente à constante de regressão não foi estatisticamente significativa, a nível de 5%, para nenhuma das equações ajustadas, conforme demonstrado pela razão t de Student, não se rejeitando, portanto, a hipótese de que a constante de regressão é igual a zero.

O coeficiente de regressão foi estatisticamente diferente da unidade, ao nível de significância de 5%, apenas na equação de dezembro, conforme indicado pelos valores da razão t de Student.

A apreciação conjunta dos resultados estatísticos obtidos para a constante de regressão e para o coeficiente de regressão é de molde a considerar as equações ajustadas, qualquer que seja o mês considerado, à exceção de dezembro, como fidedignas na previsão de produção de soja esta dunidense.

Para o Brasil, foram ajustadas seis equações de regressão linear simples, referentes às previsões dos meses de janeiro a junho, tendo a estimativa final como variável dependente.

Todas as equações foram significantes ao nível de 1%, conforme os valores do teste F de Snedecor (quadro 10). O poder explicativo das regressões, medido pelo coeficiente de determinação (R^2) apresentou-se em crescimento, sendo tanto maior quanto mais próximo da estimativa final for a previsão considerada. Os valores de R^2 obtidos, principalmente os dos três últimos meses, são bastante elevados. A estatística de Durbin-Watson mostrou-se inconclusiva na equação de março.

A constante de regressão e o coeficiente de regressão foram considerados do ponto de vista estatístico diferentes de zero, e da unidade, respectivamente ao nível de significância de 5%, apenas na equação referente ao mês de junho, conforme as correspondentes razões t calculadas.

A avaliação desses resultados para a constante de regressão e para o coeficiente de regressão permite indicar, como fidedignas para previsão das safras de soja no Brasil, as equações referentes a todos os meses, exceto o de junho.

As duas equações de regressão linear simples ajustadas para o

QUADRO 9. - Principais Características Estatísticas das Equações Linear Simples de Previsão de Safras de Soja, Estados Unidos, 1947/48 a 1984/85 (1)

Item	Previsão				
	Agosto	Setembro	Outubro	Novembro	Dezembro
Constante de regressão (b_0)	585,13	381,67	288,80	220,35	208,79
Coefficiente de regressão (b_1)	0,978	0,988	0,994	0,994	0,992
Desvio-padrão					
s_{b_0}	599,95	487,54	413,16	305,82	122,92
s_{b_1}	0,018	0,015	0,013	0,009	0,003
Razão t de Student					
t_{b_0}	0,975	0,782	0,699	0,720	1,698
t_{b_1}	1,159	0,778	0,421	0,576	2,023 ^b
Teste F de Snedecor	2.681,36 ^a	4.110,90 ^a	5.756,96 ^a	10.555,07 ^a	65.420,88 ^a
Coefficiente de determinação (R^2)	0,9867	0,9913	0,9937	0,9966	0,9994
Estatística de Durbin-Watson	1,42 IN	1,69	1,66	1,51	1,86
Coefficiente de Theil-Nagar	0,287	0,151	0,169	0,242	0,065

(1) O número de observações (n) é igual a 38. As letras representam o nível de significância: a = 1% e b = 5% e IN = inconclusivo.

Fonte: Elaborado a partir de dados básicos do United States Department of Agriculture (9).

QUADRO 10. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Soja, Brasil, 1974/75 a 1984/85 (1)

Item	Previsão					
	Janeiro	Fevereiro	Março	Abril	Maió	Junho
Constante de regressão (b_0)	779.132	1.077.512	1.005.596	905.510	678.849	911.535
Coefficiente de regressão (b_1)	0,916	0,905	0,911	0,916	0,942	0,936
Desvio-padrão						
s_{b_0}	160.076	1.074.887	956.172	756.248	463.817	301.816
s_{b_1}	0,116	0,078	0,070	0,055	0,034	0,022
Razão t de Student						
t_{b_0}	0,486	1,002	1,051	1,197	1,463	3,020 ^b
t_{b_1}	0,720	1,204	1,264	1,506	1,667	2,799 ^b
Teste F de Snedecor	62,164 ^a	131,935 ^a	168,737 ^a	274,246 ^a	755,788 ^a	1.722.228 ^a
Coefficiente de determinação (R^2)	0,8735	0,9361	0,9493	0,9682	0,9882	0,9948
Estatística de Durbin-Watson (2)	1,525	1,561	1,010 IN	1,465	1,791	1,635
Coefficiente de Theil-Nagar	0,238	0,220	0,496	0,268	0,104	0,182

(1) O número de observações (n) é 11.

(2) Os valores calculados foram cotejados com os da Tabela para 15 observações, que é o menor número disponível. As letras representam o nível de significância: a=1% e b=5%. IN = inconclusivo.

Fonte: Elaborado a partir de dados básicos da Fundação IBGE (15).

Estado de São Paulo referiram-se aos meses de fevereiro e de abril.

Ambas as equações foram significantes ao nível de 1%, conforme os valores do teste F (quadro 11). O coeficiente de determinação (R^2), foi bastante elevado nos dois casos. O teste de Durbin-Watson não indicou presença de correlação serial dos resíduos em nível que pudesse causar possíveis problemas, o que também ocorreu com o coeficiente de Theil-Nagar.

A razão t de Student, calculada para testar a hipótese nula de que o coeficiente de regressão é igual à unidade, para a previsão de abril, foi significativa ao nível de 5%, levando à rejeição dessa hipótese. Desse modo, a previsão de abril não apresenta as condições necessárias para um elevado grau de fidedignidade ao estimar a produção final paulista de soja.

7 - CONCLUSÕES

O método de análise de consistência de previsões de safra utilizando a regressão linear simples mostrou-se em geral, adequado conforme evidenciaram os valores do teste F e do coeficiente de determinação. Os resultados obtidos no estudo sugerem que a consistência das previsões de algodão é maior que a das de soja, que por sua vez é maior que a das de milho.

Dos resultados pode-se também inferir que, independentemente dos procedimentos de levantamento dos dados, o melhor desempenho pode ser atribuído aos dados do Brasil, seguindo-se os dos Estados Unidos e o do Estado de São Paulo. Essa ressalva quanto aos procedimentos de levantamento, já efetuada anteriormente, diz respeito ao levantamento subjetivo de dados do Brasil o que, segundo MUELLER (17) não permite a atribuição de margens de erro e impossibilita o estabelecimento de intervalos de confiança. Assim, mesmo que a qualidade de dados obtidos subjetivamente seja razoável, isso não pode ser comparado de maneira científica.

Uma informação de utilidade para os usuários de estatísticas de produção apresentada pelo trabalho, é a que especifica o mês de levantamento, a partir do qual já se tem uma previsão confiável da produção.

Para algodão, nos Estados Unidos, qualquer previsão a partir de agosto, fornece indicação fidedigna do volume de entrada nas usinas. Para o Brasil, qualquer previsão a partir de março fornece indicação segura da produção. Em São Paulo, apenas a estimativa final, efetuada em junho, fornece indicação aproveitável do volume de entrada nas usinas.

No caso do milho, nos Estados Unidos, apenas é fidedigna a primeira previsão (agosto). No Brasil, todas as previsões a partir de maio, podem ser utilizadas. No Estado de São Paulo, apenas a previsão de feverei

QUADRO 11. - Principais Características Estatísticas das Equações de Regressão Linear Simples de Previsão de Safras de Soja, Estado de Paulo, 1947/48 a 1984/85 (¹)

Item	Previsão	
	Fevereiro	Abril
Constante de regressão (b_0)	-2,516	-1,638
Coefficiente de regressão (b_1)	1,000	0,978
Desvio-padrão		
s_{b_0}	8,470	5,289
s_{b_1}	0,014	0,009
Razão t de Student		
t_{b_0}	-0,298	-0,310
t_{b_1}	0,005	2,393 ^b
Teste de Snedecor	4.467,92 ^a	11.462,23 ^a
Coefficiente de Determinação (R^2)	0,9933	0,9973
Estatística de Durbin-Watson	1,78	1,51
Coefficiente de Theil-Nagar	0,107	0,244

(¹) O número de observações (n) é 32, tendo sido excluídas da análise, por serem incompletos os dados das safras 1949/50, 1950/51, 1954/55, 1956/57, 1957/58 e 1959/60. As letras representam o nível de significância: a = 1% e b = 5%.

Fonte: Elaborado a partir de dados básicos do Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

ro fornece uma indicação fidedigna de produção de milho.

No que se refere à soja, tem-se nos Estados Unidos levantamentos fidedignos no período agosto-novembro. No Brasil, são aproveitáveis as previsões de janeiro a maio não o sendo o de junho. No Estado de São Paulo, apenas a previsão de fevereiro é fidedigna.

Como os dados para o Brasil e São Paulo indicam deficiência na qualidade, é necessário que os órgãos envolvidos na sua coleta, elaboração e divulgação, atuem no sentido de aprimorá-las. No caso de São Paulo, tem-se que o cadastro utilizado para extração da amostra não é suficientemente atualizado, o que é agravado por mudanças regionais marcantes nas atividades agropecuárias, em pouco espaço de tempo.

Uma possibilidade adicional de aprimoramento dessas estatísticas é o recurso a métodos alternativos, como ocorre em outros países, não visando substituir os atuais sistemas de previsão de safras, mas sim complementá-los. Um desses métodos, bastante promissor, é o emprego de satélites artificiais para sensoriamento remoto, possibilitando estimar, com margem de erro conhecida, variáveis agropecuárias importantes para obtenção de previsões de safra mais fidedignas.

O interregno entre coleta e divulgação dos dados é elevado, o que prejudica a utilização da informação pelos interessados. A agilização dessa divulgação já constituiriam num grande avanço para os usuários.

Essa lentidão pode fazer com que o tomador de decisões, quer do governo, quer da empresa privada, necessite recorrer a outros instrumentos para obter uma previsão de produção. A possibilidade de utilização de métodos econométricos para prever a produção, em uma dada safra, não deve ser posta de lado. Nesse sentido, modelos que considerem variáveis meteorológicas como fatores explicativos da produção, podem ser acionados para obter previsões em tempo hábil, já que essas variáveis estão prontamente disponíveis nas estações meteorológicas ou mesmo, a nível de propriedade, facultando a regionalização da previsão.

LITERATURA CITADA

1. ALGODÃO: relatório, 1960/70-1984/85. Curitiba, Empresa Paranaense de Classificação de Produtos - CLASPAR, 1970-1986.
2. BRANDT, Sergio A.; ARAUJO, Paulo F.C. de; SERRANO, Ondalva. Precisão e fidedignidade das previsões de safra de algodão no Estado de São Paulo, 1947-64. Piracicaba, ESALQ/USP, 1966. 15p.
3. CAMARGO, Maria de L.B. Avaliação das previsões de safra de café, Estado de São Paulo, 1960/61 a 1979/80. Informações Econômicas, São Paulo 11(8):19-24, ago. 1981.
4. CAMPOS, Humberto & PIVA, Luiz H. de O. Dimensionamento de amostra para estimativa e previsão de safra no Estado de São Paulo. Agricultura em São Paulo, SP, 21(3):65-68, 1974.
5. CARVALHO, Flávio C. de et alii. Avaliação das previsões de safra de algodão no Estado de São Paulo, 1947/48 a 1976/77. São Paulo, Secretaria da Agricultura, IEA, 1978. 11p. (Relatório de Pesquisa, 17/78)
6. CENSO AGROPECUÁRIO: Brasil, 1980. Rio de Janeiro, FIBGE, 1984. v.2, t.3, n.1.
7. CENSO AGROPECUÁRIO: São Paulo, 1980. Rio de Janeiro, FIBGE, 1984. v.2, t.3, n.19, 1ª parte.
8. CEPLAC. Divisão Sócio-Econômica. Previsão de safras. Informe de Pesquisas de 1983, Ilhêus, 1985. p.211-215.
9. COMMODITY YEARBOOK. New York, Commodity Research Bureau, 1960-1985.
10. ESTADOS UNIDOS. Department of Agriculture. Major statistical series of the U.S. Departmente of Agriculture: how they are constructed and used - crop and livestock estimates. Washington, 1971 v.8. (Agriculture Handbook, 365)

11. FARM marketing of field crops, United States, 1977-78 and 1978-79. Crop Production, Washington, dec. 1979. p.8-8.
12. FUNDAÇÃO GETÚLIO VARGAS. Centro de Estudos Agrícolas. Calendário agrícola. Rio de Janeiro, 1979. 254p.
13. HOFFMANN, Rodolfo & VIEIRA, Sonia. Análise de regressão: uma introdução à econometria. São Paulo, Hucitec, 1977. 339p.
14. KOHLS, Richard L. & UHL, Joseph N. Marketing of agricultural products. 5.ed. New York, Macmillan Publishing, 1980. 612p.
15. LEVANTAMENTOS SISTEMÁTICO DA PRODUÇÃO AGRÍCOLA. Rio de Janeiro, FIBGE, 1975-1986.
16. MASCARENHAS, Hipólito A.A. Estimativa da produção de soja. In: FUNDAÇÃO CARGILL, ed. A soja no Brasil Central. 2ed. Campinas, 1982. p.394-404.
17. MUELLER, Charles C. Previsões e estimativas de safras agrícolas: questões sobre suas formas de obtenção, confiabilidade e seus impactos. Indicadores IBGE, Rio de Janeiro, 5(8):73-84, ago. 1986. Nova Série.
18. NOGUEIRA JR., Sebastião; CARVALHO, Flávio C.de; TSUNECHIRO, Alfredo. Consistência das previsões de safra de algodão em caroço nos Estados Unidos, no Brasil e em São Paulo. In: REUNIÃO NACIONAL DO ALGODÃO,4., Belém 1986. Resumos... Campina Grande, EMBRAPA/CNPA, 1986. p.121.
19. ———; ———; ———. Consistência das previsões de safra de milho nos Estados Unidos, no Brasil e em São Paulo. In: CONGRESSO NACIONAL DE MILHO E SORGO, 16., Belo Horizonte, 1986. Resumos... Sete Lagoas, EMBRAPA/CNPMS, 1986. p.95-96.
20. ———; ———; ———. Consistência das previsões de safra de soja nos Estados Unidos, no Brasil e em São Paulo. In: SEMINÁRIO NACIONAL DE PESQUISA DE SOJA, 4., Porto Alegre, 1986. Resumos e Programa ... Londrina, EMBRAPA/CNPSo, 1986. p.23-24.

21. PEDRO Jr., Mário J. et alii, eds. Instruções agrícolas para o Estado de São Paulo. 3.ed.rev.atual. Campinas, Secretaria de Agricultura e Abastecimento, Instituto Agronômico, 1986. 229p. (Boletim 200)
22. PEETZ, Marcia S. & AMARO, Antonio A. Aplicação de equação de regressão à previsão de safra de laranja no Estado de São Paulo. Informações Econômicas, São Paulo, 8 (3):1-8, mar. 1978.
23. PINO, Francisco A. & CASER, Denise V. Análise de erros não amostrais em levantamentos para previsão e estimativas de safras no Estado de São Paulo. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1984. 25p. (Relatório de Pesquisa, 10/84)
24. ———; ———. Falta de respostas em levantamento por amostragem: um estudo de caso. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1984. 25p. (Relatório de Pesquisa, 08/84)
25. ———; CAMARGO, Maria de L.B.; JIMENEZ OSSIO, Júlio H. Estatísticas de produção agrícola no Estado de São Paulo. 1942/43 a 1984/85. Informações Econômicas, 16:1-60, set. 1986. Supl. n.1.
26. SALMON, David S. The cycle of food and fiber production. Commodity Yearbook, New York, 1986. p.15-28.
27. SILVA, Gabriel L.S.P. da; CASER, Denise V.; VICENTE, José R. Efeito das condições do tempo sobre a produtividade agrícola no Estado de São Paulo. Revista de Economia Rural, Brasília, 23(1):3-19, jan./mar. 1985.
28. ———; ———; ———. Variações do tempo e da produtividade agrícola: proposta de uma metodologia de análise e sua aplicação para a cultura do café no Estado de São Paulo. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1986. 28p. (Relatório de Pesquisa, 13/86)

29. ———; VICENTE, José R.; CASER, Denise V. Variações do tempo e produtividade agrícola: um subsídio à previsão de safras no Estado de São Paulo. Campinas, Fundação Cargill, 1986. 148p.
30. VICENTE, José R.; CASER, Denise V.; CAMARGO, Ana M.M.P. Comportamento das previsões de safras agrícolas do Estado de São Paulo em relação à estimativa final. São Paulo, Secretaria da Agricultura e Abastecimento. IEA, s.d. (no prelo)
31. WIESEMEYER, Jim. How USDA puts together its crop report estimates. Futures, Iowa, 12(11):50-54, nov. 1983.

R E S U M O

O trabalho tem por objetivo descrever e analisar a consistência dos levantamentos de previsões de safra de algodão, milho e soja nos Estados Unidos, no Brasil e em São Paulo, realizados respectivamente, pelo Departamento de Agricultura dos Estados Unidos (USDA), Fundação Instituto Brasileiro de Geografia e Estatística (IBGE) e Instituto de Economia Agrícola (IEA). São apresentadas as sistemáticas dos levantamentos de dados e os calendários de plantio e colheita nas regiões consideradas.

O método de análise utilizado — regressão linear simples — que associa levantamentos preliminares com o levantamento final de produção, mostrou-se adequado. Os resultados obtidos sugerem que a consistência das previsões de algodão é maior, comparativamente, que a das de soja, que por vez é maior que a das de milho. Infere-se também pelos dados globais, independentemente dos procedimentos de levantamento, que o desempenho das previsões do Brasil é melhor que o dos Estados Unidos, que por sua vez é melhor que o do Estado de São Paulo. Contudo, são feitas ressalvas quanto aos métodos utilizados.

O estudo mostra o mês de levantamento, a partir do qual, já se tem uma previsão confiável da produção. Para o algodão, nos Estados Unidos, é o do mês de agosto; no Brasil é o do mês de março; e em São Paulo, o do mês de junho. Para o milho nos Estados Unidos, a previsão de agosto, no Brasil, a de maio e em São Paulo, a de fevereiro. Para a soja nos Estados Unidos, a previsão de agosto, no Brasil a de maio e no Estado de São Paulo, a de fevereiro.

Finalizando, os autores sugerem a utilização de métodos alternati

vos para complementação dos sistemas tradicionalmente utilizados que poderiam aprimorar e agilizar as informações de previsão de safra. Neste caso são citados o emprego de satélites artificiais para sensoriamento remoto e o emprego de modelos estatísticos que considerem variáveis meteorológicas como fatores explicativos da produção. Cumpre ressaltar que os métodos atuais não devem ser substituídos, mas sim aprimorados, já que se constituem em importantes instrumentos para tomada de decisão dos usuários.

CONSISTENCY OF THE COTTON, CORN AND SOYBEAN CROP FORECASTS IN THE UNITED STATES, BRAZIL AND SÃO PAULO STATE

S U M M A R Y

The paper objectives to describe and to analyse the consistency of the forecast surveys of cotton, corn and soybeans in the United States Department of Agriculture (USDA), Fundação Instituto Brasileiro de Geografia e Estatística (IBGE) and Instituto de Economia Agrícola (IEA). The systematic of the data surveys and the sowing and harvesting seasons are introduced in the considered regions.

The analysis method used - simple linear regression - that associates preliminary surveys with the final datum of production, revealed itself appropriated. The results suggested that the consistency of the cotton forecast is bigger, comparatively, than the soybeans ones, that by their time, is bigger than the corn ones. It can also conclude through the global data, independently of the procedures of the surveys, that the performance of the Brazilian forecasts is better than the United States one, that by this time is better than the São Paulo State one. However, exemptions are made as for the used methods.

The study shows the survey since that already it has a confident forecast of the production. For the cotton, in the United States is the August forecast; in Brazil the survey of March; and in São Paulo, the June one. For the corn, in the United States, the forecast of August, in Brazil, the May one and in São Paulo, the forecast of February. For the soybeans, in the United States, the August forecast, in Brazil, the May one and in São Paulo, the February one.

Finally, the authors suggest the utilization of alternative methods to the complementation of the traditionally used systems that could improve and become dynamics the informations of crop surveys.

In this case are mentioned the use of the artificial satellites for the re mote sensing and the use of statistical models that consider meteorological variables as explanatory factors of the production. It can emphasized that the current methods cannot be replaced, but improved since constitute them selves in important ways to decision-making of the users.

**SECRETARIA DA AGRICULTURA
INSTITUTO DE ECONOMIA AGRÍCOLA**

Comissão Editorial:

Coordenador: Flavio Condé de Carvalho

Membros: Antonio Ambrósio Amaro

Arthur Antonio Ghilardi

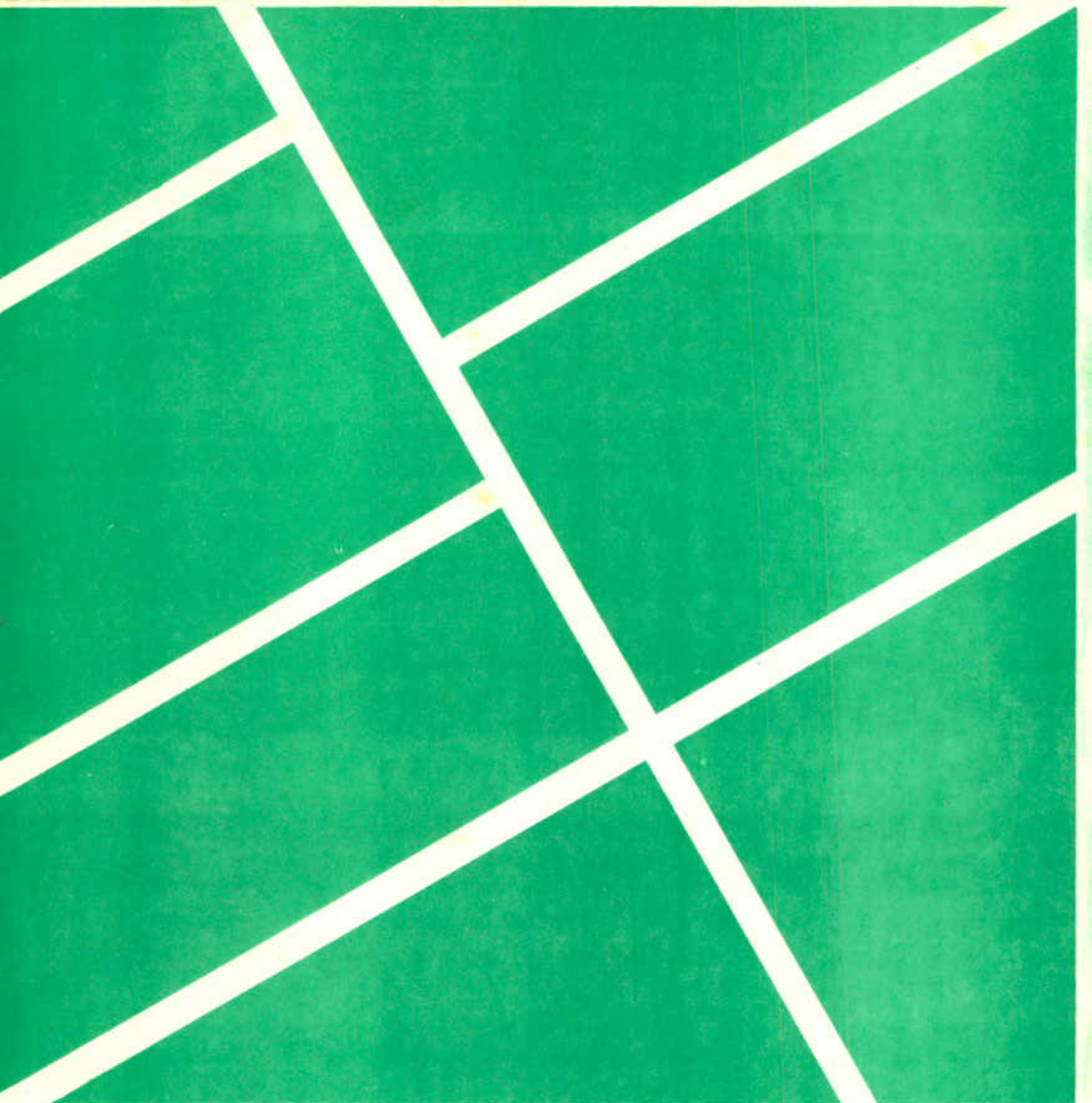
Elcio Umberto Gatti

José Luiz Teixeira Marques Vieira

Maria Carlota Meloni Vicente

Maria de Lourdes Barros Camargo

Bibliografia: Fátima Maria Martins Saldanha Faria



Relatório de Pesquisa
Nº23/87

Governo do Estado de São Paulo
Secretaria da Agricultura
Coordenadoria Sócio-Econômica

Instituto de Economia Agrícola