

AGRICULTURA EM SÃO PAULO
Boletim Técnico do Instituto de Economia Agrícola

Ano 37

Tomo 1

1990

COMPARAÇÕES ENTRE DADOS DOS CENSOS AGROPECUÁRIOS E ESTIMATIVAS DAS SAFRAS AGRÍCOLAS DO ESTADO DE SÃO PAULO⁽¹⁾

José Roberto Vicente⁽²⁾
Mário P.A. Olivetti⁽³⁾
Luiz H.O. Piva⁽²⁾
Denise Viani Caser⁽²⁾
Ana M.M.P. Camargo⁽²⁾

RESUMO

O objetivo deste trabalho foi verificar a consistência de informações provenientes de três fontes distintas, referentes à produção agrícola do Estado de São Paulo. Os resultados obtidos mostraram diferenças significativas entre os dados dos censos agropecuários e os levantados pelas outras fontes, no período 1960-1980. Para as séries temporais (1976/77 a 1988/89), os dados levantados pelo IEA, pelo IBGE e pela CFP, via-de-regra, não diferiram estatisticamente. É provável que, o intercâmbio entre esses órgãos esteja permitindo homogeneizar os dados, desfazer eventuais dúvidas e aperfeiçoar as estimativas. Isso propicia aos usuários maior confiabilidade nas informações divulgadas.

Palavras-chave: levantamentos agrícolas; previsão de safras.

COMPARISONS BETWEEN THE CENSUSES AND CROP ESTIMATIONS IN THE STATE OF SÃO PAULO

SUMMARY

The purpose of this paper was to compare information proceeding from three different sources of agricultural data in the State of São Paulo, Brazil, in order to check their consistency. The results showed significant difference between the census data of 1960, 1970, 1975 and 1980 and the obtained data by other sources. The time series from 1976/77 to 1988/89, the produced data by IEA and CFP, most of time, didn't statistically differ. Probably the interchanging of information among those offices is contributing to homogenize the data, eliminate doubts and improving the estimations. This fact gives the users greater reliance to the information.

Key-words: agricultural surveys; crop forecasting.

⁽¹⁾ Trabalho referente ao projeto STPC 16-039/89. Recebido em 29/01/90. Liberado para publicação em 09/04/90.

⁽²⁾ Pesquisador Científico do Instituto de Economia Agrícola (IEA).

⁽³⁾ Bacharel em Geografia, funcionário do Instituto de Economia Agrícola (IEA).

Agricultura em São Paulo, SP, 37(1):97-104, 1990.

1 - INTRODUÇÃO

As estimativas de safras têm como finalidade principal subsidiar os formuladores de políticas agrícolas e os tomadores de decisão do setor privado ligados às atividades agropecuárias e fornecer dados para serem utilizados em pesquisas. Dados precisos sobre quantidades produzidas apresentam elevados retornos sociais, NEGRI NETO et alii (9), razão pela qual, a sociedade movimenta recursos consideráveis para sua realização.

No Estado de São Paulo, como em todo o Brasil, as estatísticas oficiais são as publicadas pela Fundação Instituto Brasileiro de Geografia e Estatísticas (IBGE). O IBGE organiza, em cada Estado, o Grupo de Coordenação de Estatísticas Agropecuárias (GCEA), no qual especialistas de diversos órgãos reúnem-se e consolidam os dados levantados pelas várias instituições divulgando-se aqueles escolhidos por consenso.

Um dos principais levantamentos de safras de São Paulo é o da própria rede do IBGE, contemplando cerca de trinta culturas, efetuado pelos agentes de coleta regionais, que consultam vários especialistas e, tem como uma de suas fontes principais os engenheiros agrônomos da Coordenadoria de Assistência Técnica Integral (CATI), sendo a unidade do levantamento o município. Como o dado oficial do IBGE é o do GCEA, muitas vezes o levantamento da rede própria é preterido.

Outro órgão que levanta dados em São Paulo, para oito culturas é a Companhia de Financiamento da Produção (CFP), cujos técnicos viajam às principais regiões produtoras, normalmente ao nível de Divisão Regional Agrícola (DIRA) ou de Delegacia Regional Agrícola, e tem nos técnicos da CATI, os responsáveis por esses locais, uma das principais fontes de informações. Posteriormente, os dados são enviados a Brasília, corrigidos e divulgados.

O mais antigo levantamento para previsões e estimativas das safras paulistas é o efetuado pela Secretaria de Agricultura e Abastecimento, atualmente, através do Instituto de Economia Agrícola (IEA) que planeja e processa os levantamentos e da CATI, que efetua a pesquisa de campo. O sistema compreende um levantamento

subjetivo (ao nível de município, para 120 culturas) e um levantamento objetivo (ao nível de imóvel rural). Este último, que consiste em amostra probabilística duplamente estratificada, CAMPOS & PIVA (4) e CAMARGO (3), retirada do cadastro do Instituto Nacional de Colonização e Reforma Agrária (INCRA), engloba uma grande variedade de questões, abordando, além das previsões de safras, demografia e mão-de-obra, salários, preços pagos e recebidos, utilização de insumos, tecnologia empregada, etc., que subsidiam diversas pesquisas sócio-econômicas realizadas pelo IEA e por outras instituições.

Outra fonte, por excelência, de dados sócio-econômicos é o Censo Agropecuário, atualmente com periodicidade quinquenal, efetuado pelo IBGE. A principal desvantagem do Censo é a demora na publicação dos resultados; por exemplo, até janeiro de 1990, não tinham ainda sido divulgados os dados do Censo de 1985.

Essas fontes alternativas raramente têm seus dados comparados com métodos adequados. O máximo que se ousou, em tal sentido, foi apontar discrepâncias ocasionais, em anos em que sua magnitude chama mais a atenção.

1.1 - Objetivos

O objetivo central do presente estudo é comparar os dados levantados e publicados pelo IBGE (Censo Agropecuário e Anuário Estatístico do Brasil), CFP e IEA, referentes ao Estado de São Paulo, procurando identificar quais das séries dos principais produtos agrícolas do Estado diferem entre si.

2 - METODOLOGIA

Para cumprir os objetivos propostos, tomar-se-á os dados publicados dos Censos Agropecuários (5) de 1960⁽⁴⁾, 1970, 1975 e 1980, referentes às culturas do algodão, amendoim, arroz, banana, batata, café, cana, cebola, feijão, laranja, mandioca, mamona, milho, soja, tomate e trigo, que serão comparados com os publicados pelo próprio IBGE (1) e pelo IEA (12).

⁽⁴⁾ O censo de 1960 traz informações sobre a produção obtida no ano de 1959. *Agricultura em São Paulo*, SP, 37(1):97-104, 1990.

Também serão utilizados os dados da CFP (8), disponíveis para o período 1976/77 a 1988/89, para comparações com os publicados pelo IEA e IBGE, referentes ao mesmo período e às culturas de algodão, arroz, milho, soja, amendoim, feijão, trigo e mamona. Por último, uma vez que a laranja, a cana e o café são os principais produtos agrícolas do Estado de São Paulo em termos de valor da produção (13), apesar de não terem seus dados publicados pela CFP, efetuar-se-á comparações entre as séries do IEA e do IBGE, também referentes ao período 1976/77 a 1988/89.

O método utilizado consiste no ajuste de modelos de regressão linear simples, com uma das séries como variável dependente e a outra como variável independente. Cada produto poderá, então, dar origem a dois modelos, já que pode-se ajustar $Y = f(X)$ e $X = g(Y)$, onde os parâmetros estimados serão diferentes. Para escolher qual dos modelos terá seus parâmetros testados, optou-se por considerar como mais adequado o que apresentar valor da estatística de Durbin-Watson mais próximo de 2, uma vez que as estimativas dos coeficientes são mais confiáveis na ausência de autocorrelação. Por esse mesmo motivo, optou-se por reajustar os modelos oriundos de mínimos quadrados ordinários com valor da estatística de Durbin-Watson maiores do que 2,3 e menores do que 1,7, pelo método de Cochran-Orcutt, KMENTA (7), exceto para os dados do Censo Agropecuário, que não constituem série contínua.

A significância dos modelos estimados foi testada, como é usual, pelo valor da estatística F.

Para verificar se as duas séries são iguais, serão testados o intercepto (a) e o coeficiente (b). Caso o primeiro seja estatisticamente igual a zero e o segundo estatisticamente igual a um, pode-se considerar que as séries não diferem.

Mesmo os modelos que não forem significativos terão seus parâmetros testados como descrito; todavia, caso não se recuse as hipóteses de $a = 0$ e $b = 1$, a não significância do modelo não permitirá concluir que as séries não diferem. Por outro lado, a rejeição de uma dessas hipóteses servirá para rejeitar, também, a igualdade estatística.

Dentre os tipos de testes que podem ser utilizados citam-se o "F", que verifica a exis-

tência simultânea de $a = 0$ e $b = 1$, como descrito por HOFFMANN & VIEIRA (6), e utilizados por VICENTE; CASER; CAMARGO (14) e testes "t" separados para os coeficientes estimados e que, devido à ortogonalidade no caso presente, são equivalentes ao primeiro processo. Esses últimos foram utilizados por PEETZ & AMARO (11), CAMARGO (2) e NOGUEIRA JR.; CARVALHO; TSUNECHIRO (10).

Dos programas de regressão disponíveis para microcomputadores, consegue-se, facilmente, calcular esta segunda alternativa, que por simplicidade, foi escolhida:

$$t_a = \frac{\hat{a}-0}{\hat{s}_a} \qquad t_b = \frac{\hat{b}-1}{\hat{s}_b}$$

onde

t_a e t_b são estatísticas "t" de Student, relativas ao intercepto e ao coeficiente angular, respectivamente; \hat{a} é a estimativa do intercepto; \hat{b} é a estimativa do coeficiente angular; e \hat{s}_a e \hat{s}_b são as estimativas dos desvios-padrões das estimativas dos parâmetros.

A variável escolhida para as comparações foi a produção, não apenas por ser a variável economicamente mais importante, mas também, porque as fontes diferem quanto à definição das áreas; enquanto, por exemplo, o IEA levanta a área plantada, o IBGE levanta a área colhida, que podem diferir significativamente em anos de ocorrência de fenômenos climáticos adversos severos.

O nível de significância escolhido para os testes foi o de 5% de probabilidade.

3 - RESULTADOS E DISCUSSÃO

O pequeno número de observações existentes, a partir de 1960 dos censos agropecuários, fez com que as séries cujos coeficientes de correlação não sejam bastante elevados (inferior a 0,95), propiciassem modelos estatisticamente não significativos. Isso aconteceu para quinze das 32 comparações entre séries do censo e do IEA ou do IBGE, distribuídas quase que igualmente (oito comparações não significativas

com séries do IEA e sete com séries do IBGE) (quadro 1).

Entre as dezessete comparações que propiciaram modelos significativos a 5% de probabilidade, seis delas apresentaram coeficientes angulares diferentes da unidade, e uma apresentou intercepto diferente de zero, sendo, portanto, estatisticamente distintas.

Os produtos que os testes revelaram ter dados estatisticamente não diferentes no censo e nas séries do IEA e IBGE foram a cana, o café, a soja e o tomate, enquanto que o algodão tem dados estatisticamente não diferentes nos censos e na série do IEA, e a mamona e o trigo, nos censos e na série do IBGE.

No Estado de São Paulo existe outra fonte de dados para a cultura do algodão, que é a entrada nas usinas de beneficiamento. Comparando-se os dados ali provenientes com os das outras três fontes, observa-se que, a produção declarada nos censos é sistematicamente menor, em magnitudes entre 52 e 147 mil toneladas, o que representa de 12% a 24% da quantidade produzida. Já para as séries do IEA e as publicadas pelo IBGE nos Anuários Estatísticos do Brasil, essa diferença é sensivelmente menor (quadro 2).

Quando se analisa as séries temporais referentes aos anos agrícolas 1976/77 a 1988/89, da CFP, do IBGE e do IEA, o número de observações mais elevado permite obter modelos significativos para quase todas as comparações, exceto para a mamona com as séries da CFP e do IEA (quadro 3).

Para o arroz, o milho, o amendoim, o feijão e o trigo, os testes efetuados indicaram que as três séries não diferem. O café e a cana, produtos não levantados pela CFP, não diferem nas séries do IEA e do IBGE. No caso do algodão, os dados de produção publicados pela CFP e pelo IEA são estatisticamente não diferentes, enquanto que os publicados pelo IBGE diferem de ambas as séries anteriores.

A soja apresenta as maiores diferenças entre as três fontes de dados: em todos os modelos, tanto o intercepto como o coeficiente angular apareceram diferentes, estatisticamente, de zero e um, respectivamente. Esse resultado também se verificou para a laranja, comparando-se as séries do IEA e do IBGE.

O comportamento mais diversificado foi o Agricultura em São Paulo, SP, 37(1):97-104, 1990.

verificado para a mamona: as séries do IBGE e do IEA não diferem estatisticamente, as do IBGE e da CFP são diferentes e as do IEA e da CFP sequer propiciam ajuste de modelo significativo.

4 - CONCLUSÕES E SUGESTÕES

Os resultados obtidos mostraram diferenças significativas entre os dados de produção levantados nos Censos Agropecuários e os publicados pelo IEA e pelo próprio IBGE, nos Anuários Estatísticos do Brasil. Esse fato pode ser melhor confirmado com a introdução de novos dados censitários e/ou análises com observações ao nível de microrregiões ou de Estados. Essas conclusões parecem ser corroboradas por análise da série que dispunha de maior controle, a da produção de algodão, e que indica subestimação considerável nas informações dos Censos. Esse fato é preocupante quando se sabe que os Censos Agropecuários são a única fonte para uma série de estudos, temendo-se pela qualidade de informações sobre, por exemplo, utilização de insumos e gastos efetuados já que, os produtores do Estado mais desenvolvido do País podem estar sendo incapazes de informar corretamente mesmo a produção obtida.

A análise das séries temporais, publicadas por três fontes distintas, mostrou, via-de-regra, consistência entre elas. Disso pode-se concluir que, as instituições envolvidas têm cumprido seu papel a contento, contribuindo para que se tenha conhecimento aproximado da produção obtida. Se é possível, por um lado, questionar a sobreposição dessas funções numa economia de recursos escassos, é também verdade que, esses órgãos trocam informações entre si, o que permite homogeneizar os dados, desfazer eventuais dúvidas e aperfeiçoar as estimativas.

LITERATURA CITADA

1. ANUÁRIO ESTATÍSTICO DO BRASIL. Rio de Janeiro, IBGE, 1960-1989.
2. CAMARGO, Maria de L.B. Avaliação das previsões de safras de café no Estado de

- São Paulo em relação à estimativa final. *Informações Econômicas*, São Paulo, 11(8):19-23, ago. 1981.
3. CAMARGO, Milton N. de. **Amostra para previsão e estimativa das safras agrícolas do Estado de São Paulo em vigor a partir de junho de 1981.** São Paulo, Secretaria da Agricultura, IEA, 1988. 75p. (Relatório de Pesquisa, 27/88)
 4. CAMPOS, Humberto de & 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):65-88, 1974.
 5. CENSO AGROPECUÁRIO: São Paulo. Rio de Janeiro, IBGE, 1960, 1970, 1975, 1980.
 6. HOFFMANN, Rodolfo & VIEIRA, Sergio. **Análise de regressão: uma introdução à econometria.** São Paulo, Hucitec/EDUSP, 1977. 339p.
 7. KMENTA, Jan. **Elementos de econometria.** São Paulo, Atlas, 1978. 670p.
 8. MÊS AGRÍCOLA. Brasília, Ministério da Agricultura, CFP, set. 1976-dez. 1989.
 9. NEGRI NETO, Afonso et alii. **Custo e benefício social de previsões e estimativas de produção agrícola: o valor da informação.** *Agricultura em São Paulo*, SP, 35(1):37-49, 1988.
 10. NOGUEIRA Jr., Sebastião; CARVALHO, Flavio C. de; TSUNECHIRO, Alfredo. **Consistência das previsões de safra de algodão, milho e soja nos Estados Unidos, Brasil e em São Paulo.** São Paulo, Secretaria de Agricultura, 1987. 36p. (Relatório de Pesquisa, 23/87)
 11. PEETZ, Marcia da S. & AMARO, Antonio A. **Aplicações de equações de regressão à previsão de safras de laranja no *Agricultura em São Paulo*, SP, 37(1):97-104, 1990.**
 12. **PREVISÕES E ESTIMATIVAS DAS SAFRAS AGRÍCOLAS DO ESTADO DE SÃO PAULO.** São Paulo, Secretaria de Agricultura e Abastecimento, IEA, jun. 1959-set. 1989.
 13. **PROGNÓSTICO AGRÍCOLA.** São Paulo, Secretaria da Agricultura, IEA, 1988. v.1.
 14. VICENTE, José R.; CASER, Denise V.; CAMARGO, Ana M.M.P. de. **Comportamento das previsões de safras agrícolas no Estado de São Paulo em relação à estimativa final.** São Paulo, Secretaria da Agricultura, 1987. 22p. (Relatório de Pesquisa, 05/87)

QUADRO 1. - Principais Resultados da Análise de Regressão Utilizando Dados dos Censos Agropecuários (CA/IBGE), das Previsões e Estimativas das Safras Agrícolas (IEA) e dos Anuários Estatísticos do Brasil (AE/IBGE), Estado de São Paulo, Referentes a 1959, 1970, 1975 e 1980

Cultura	Variável dependente	Variável independente	a (1)	b (2)	R ² (3)	d (4)
Algodão	IEA	CA/IBGE	30.865	1,097	0,936*	1,99
	CA/IBGE	AE/IBGE	-15.540	0,862	0,896	1,91
Amendoim	IEA	CA/IBGE	-190.714	1,946*	0,986*	1,85
	AE/IBGE	CA/IBGE	109.018	2,478*	0,961*	1,22
Arroz	IEA	CA/IBGE	396.114	0,419	0,077	1,95
	AE/IBGE	CA/IBGE	478.050	0,476	0,031	1,92
Banana	IEA	CA/IBGE	-452	1,172	0,210	1,88
	AE/IBGE	CA/IBGE	281.021*	-7,511*	0,969*	1,91
Batata	CA/IBGE	IEA	326.808*	-0,092*	0,220	1,89
	AE/IBGE	CA/IBGE	291.371	0,330	0,002	1,80
Café	CA/IBGE	IEA	269.495	0,580	0,945*	2,13
	CA/IBGE	AE/IBGE	-46.951	0,969	0,982*	1,40
Canã	CA/IBGE	IEA	-1,2x10 ⁶	1,165	0,956*	2,76
	CA/IBGE	AE/IBGE	-2,8x10 ⁵	1,036	0,996*	2,22
Cebola	CA/IBGE	IEA	35.274	0,511	0,722	1,99
	CA/IBGE	AE/IBGE	38.923	0,501	0,729	1,90
Feijão	CA/IBGE	IEA	5.117	0,647	0,868	2,97
	CA/IBGE	AE/IBGE	-9.971	0,669	0,759	1,97
Laranja	IEA	CA/IBGE	565.194	1,267*	0,999*	1,96
	CA/IBGE	AE/IBGE	1,9x10 ⁵	0,716*	0,991*	2,85
Mandioca	CA/IBGE	IEA	168.691	0,086*	0,202	2,72
	CA/IBGE	AE/IBGE	113.173	0,144*	0,484	2,21
Mamona	CA/IBGE	IEA	688	0,463	0,815	1,43
	CA/IBGE	AE/IBGE	-8.005	0,669	0,973*	1,32
Milho	CA/IBGE	IEA	481.034	0,671	0,734	1,41
	CA/IBGE	AE/IBGE	242.084	0,790	0,786	1,64
Soja	IEA	CA/IBGE	-38.426	1,232	0,991*	2,96
	AE/IBGE	CA/IBGE	-26.124	1,158	0,998*	2,96
Tomate	IEA	CA/IBGE	-43.293	1,281	0,976*	2,68
	AE/IBGE	CA/IBGE	6.734	1,132	0,948*	2,89
Trigo	CA/IBGE	IEA	-5.614	0,720*	0,994*	1,83
	AE/IBGE	CA/IBGE	7.224	1,418	0,986*	2,19

(1) Estimativa do intercepto; os assinalados com asterisco são estatisticamente diferentes de zero a 5% de probabilidade.

(2) Estimativa do coeficiente angular; os assinalados com asterisco são estatisticamente diferentes da unidade a 5% de probabilidade.

(3) Coeficiente de determinação; o asterisco indica os modelos significativos a 5% de probabilidade.

(4) Valor da estatística de Durbin-Watson.

Fonte: Elaborado a partir de dados básicos da Fundação Instituto Brasileiro de Geografia e Estatística (1, 5) e do Instituto de Economia Agrícola (12).
Agricultura em São Paulo, SP, 37(1):97-104, 1990.

QUADRO 2. - Produção de Algodão em Carço, Segundo Fontes Diferentes, Estado de São Paulo, 1959, 1970, 1975 e 1980

(em tonelada)

Ano	IBGE		IEA	Usinas de beneficiamento
	Censo	Anuário estatístico		
1959	442.046	574.431	525.000	502.200
1970	610.222	707.810	705.000	757.300
1975	394.736	488.600	488.550	489.600
1980	433.388	482.635	466.950	485.198

Fonte: Fundação Instituto Brasileiro de Geografia e Estatística (1, 5) e Instituto de Economia Agrícola (12).

QUADRO 3. - Principais Resultados da Análise de Regressão Utilizando Dados do Instituto de Economia Agrícola(IEA), do Instituto Brasileiro de Geografia e Estatística (IBGE) e da Companhia de Financiamento da Produção (CFP), Estado de São Paulo, 1976/77 a 1988/89

Cultura	Variável dependente	Variável independente	a (1)	b (2)	R ² (3)	d (4)
Algodão	IEA	IBGE	74.592	0,794*	0,920*	1,93
	CFP	IEA	68.543	0,913	0,708*	2,14 ⁽⁵⁾
	CFP	IBGE	118.570	0,761*	0,794*	1,73 ⁽⁵⁾
Arroz	IEA	IBGE	10.175	0,964	0,974*	1,90 ⁽⁵⁾
	CFP	IBGE	20.782	0,943	0,952*	1,86 ⁽⁵⁾
	CFP	IEA	19.251	0,959	0,961*	2,08
Milho	IBGE	IEA	75.578	0,967	0,994*	1,90 ⁽⁵⁾
	CFP	IEA	155.483	0,949	0,877*	1,57 ⁽⁵⁾
	CFP	IBGE	221.343	0,931	0,973*	1,72
Soja	IBGE	IEA	252.788*	0,698*	0,803*	1,98
	CFP	IEA	348.944*	0,589*	0,872*	1,88
	IBGE	CFP	-192.898*	1,219*	0,975*	1,76
Laranja	IEA	IBGE	6.149*	0,855*	0,961*	1,73
Café	IEA	IBGE	29.666	0,986	0,933*	1,77 ⁽⁵⁾
Cana	IBGE	IEA	-6.317	1,057	0,975*	1,71
Amendoim	IBGE	IEA	29.225	0,827	0,832*	1,96
	CFP	IEA	20.179	0,875	0,829*	1,84
	IBGE	CFP	12.828	0,933	0,891*	2,13 ⁽⁵⁾
Feijão	IEA	IBGE	-10.88	1,051	0,794*	1,78
	IEA	CFP	37.002	0,898	0,516*	1,92
	IBGE	CFP	49.171	0,843	0,632*	1,83
Trigo	IEA	IBGE	2.875	0,998	0,992*	2,23
	CFP	IEA	-9.510	0,943	0,900*	2,00
	CFP	IBGE	-11.163	0,961	0,931*	1,98
Mamona	IBGE	IEA	3.321	0,966	0,431*	1,70
	CFP	IEA	0.886	1,102	0,276	1,86
	IBGE	CFP	14.216*	0,414*	0,349*	2,44

(1) Estimativa do intercepto; os assinalados com asterisco são estatisticamente diferentes de zero a 5% de probabilidade.

(2) Estimativa do coeficiente angular; os assinalados com asterisco são estatisticamente diferentes da unidade a 5% de probabilidade.

(3) Coeficiente de determinação; o asterisco indica os modelos significativos a 5% de probabilidade.

(4) Valor da estatística de Durbin-Watson.

(5) Modelo estimado pelo método de Cochrane-Orcutt.

Fonte: Elaborado a partir de dados básicos da Fundação Instituto Brasileiro de Geografia e Estatística (1, 5) e do Instituto de Economia Agrícola (12).

Agricultura em São Paulo, SP, 37(1):97-104, 1990.