

**Sistema para estimação  
de área plantada por  
amostragem – SEARA –  
proposta de delineamento  
amostral para café em  
São Paulo e para grãos  
em Minas Gerais**

## **República Federativa do Brasil**

*Luis Inácio Lula da Silva*

Presidente

Ministério da Agricultura, Pecuária e Abastecimento

*Roberto Rodrigues*

Ministro

## **Empresa Brasileira de Pesquisa Agropecuária – Embrapa**

### **Conselho de Administração**

*Luís Carlos Guedes Pinto*

Presidente

*Sílvio Crestana*

Vice-Presidente

*Alexandre Kalil Pires*

*Cláudia Assunção dos Santos Viegas*

*Ernesto Paterniani*

*Hélio Tollini*

Membros

### **Diretoria Executiva da Embrapa**

*Sílvio Crestana*

Diretor-Presidente

*José Geraldo Eugênio de França*

*Kepler Euclides Filho*

*Tatiana Deane de Abreu Sá*

Diretores-Executivos

### **Embrapa Meio Ambiente**

*Paulo Choji Kitamura*

Chefe Geral

*Ladislau Araújo Skorupa*

Chefe-Adjunto de Pesquisa e Desenvolvimento

*Maria Cristina Martins Cruz*

Chefe-Adjunto de Administração

*Ariovaldo Luchiani Junior*

Chefe-Adjunto de Comunicação e Negócios

## **Documentos 49**

# **Sistema para estimação de área plantada por amostragem – SEARA – proposta de delineamento amostral para café em São Paulo e para grãos em Minas Gerais**

***Alfredo José Barreto Luiz  
Marcos Corrêa Neves  
Sílvio Roberto Medeiros Evangelista  
Aline de Holanda Nunes Maia***

Exemplares dessa publicação podem ser solicitados à:

Embrapa Meio Ambiente  
Rodovia SP 340 - km 127,5 - Tanquinho Velho  
Caixa Postal 69 13820-000, Jaguariúna, SP  
Fone: (19) 3867-8750 Fax: (19) 3867-8740  
sac@cnpma.embrapa.br  
www.cnpma.embrapa.br

#### **Comitê de Publicação da Unidade**

Presidente: *Ladislau Araújo Skorupa*

Secretário-Executivo: *Sandro Freitas Nunes*

Bibliotecária: *Maria Amélia de Toledo Leme*

Membros: *Heloisa Ferreira Filizola, Manoel Dornelas de Souza, Cláudio César de Almeida Buschinelli, Maria Conceição Peres Young Pessoa, Osvaldo Machado R. Cabral e Marta Camargo de Assis*

Normalização Bibliográfica: *Maria Amélia de Toledo Leme*

Editoração Eletrônica: *Sandro Freitas Nunes*

**1ª edição eletrônica**  
(2006)

#### **Todos os direitos reservados.**

A reprodução não-autorizada desta publicação, no seu todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

---

Luiz, Alfredo J. Barreto

Sistema para estimação de área plantada por amostragem – SEARA – Proposta de delineamento amostral para café em São Paulo e para grãos em Minas Gerais / Alfredo José Barreto Luiz, Marcos Corrêa Neves, Sílvio Roberto Medeiros Evangelista e Aline de Holanda Nunes Maia.– Jaguariúna: Embrapa Meio Ambiente, 2006.

26p. – (Embrapa Meio Ambiente. Documentos; 49)

1. Estatística agrícola. 2. Amostragem. I. Luiz, Alfredo J. Barreto. II. Neves, Marcos Corrêa. III. Evangelista, Sílvio Roberto Medeiros. IV. Maia, Aline de Holanda Nunes. V. Título. VI. Série.

CDD 630.2195

---

© Embrapa 2006

# **Autores**

## ***Alfredo José Barreto Luiz***

Engenheiro Agrônomo, Doutor em Sensoriamento Remoto, Pesquisador da Embrapa Meio Ambiente, Rodovia SP 340 - Km 127,5 - 13.820-000, Jaguariúna, SP. E-mail: [alfredo@cnpma.embrapa.br](mailto:alfredo@cnpma.embrapa.br)

## ***Marcos Corrêa Neves***

Engenheiro Eletricista, Doutor em Sensoriamento Remoto, Pesquisador da Embrapa Meio Ambiente, Rodovia SP 340 - Km 127,5 - 13.820-000, Jaguariúna, SP. E-mail: [marcos@cnpma.embrapa.br](mailto:marcos@cnpma.embrapa.br)

## ***Silvio Roberto Medeiros Evangelista***

Bacharel em Estatística, Doutor em Engenharia de Software, Embrapa Informática Agropecuária, Av. André Tosello, 209 - Caixa Postal 6041, Barão Geraldo, 13083-970 Campinas, SP. E-mail: [silvio@cnptia.embrapa.br](mailto:silvio@cnptia.embrapa.br)

## ***Aline de Holanda Nunes Maia***

Engenheira Agrônoma, Doutora em Agronomia, Pesquisadora da Embrapa Meio Ambiente, Rodovia SP 340 - Km 127,5 - 13.820-000, Jaguariúna, SP. E-mail: [ahmaia@cnpma.embrapa.br](mailto:ahmaia@cnpma.embrapa.br)

# Sumário

Introdução .....	05
Questões técnico-científicas .....	08
Revisão de literatura .....	09
Objetivos .....	12
Métodos .....	13
Estratégia de ação .....	14
Resultados esperados .....	14
Resultados já alcançados .....	15
a) Módulo de aleatorização .....	15
b) Delineamento amostral para a cultura do café no Estado de São Paulo .....	15
c) Delineamento amostral para culturas no Estado de Minas Gerais .....	20
Conclusão .....	23
Referências bibliográficas .....	24

# **Sistema para estimação de área plantada por amostragem - SEARA- proposta de delineamento amostral para café em São Paulo e para grãos em Minas Gerais**

---

*Alfredo José Barreto Luiz*

*Marcos Côrrea Neves*

*Silvio Roberto Medeiros Evangelista*

*Aline de Holanda Nunes Maia*

## **Introdução**

A informação precisa e atualizada referente à produção agrícola brasileira é importante e estratégica, tanto do ponto de vista econômico como da segurança alimentar. O aumento da disponibilidade de imagens de sensoriamento remoto, bem como outros avanços tecnológicos recentes, como os sistemas de informação geográfica (SIGs) e os aparelhos de posicionamento global (GPS), podem facilitar a obtenção de estimativas de área plantada, um componente fundamental da previsão de safras, e permitir o acesso a uma informação essencial em aplicações ambientais: a localização espacial dos cultivos. Embora a previsão oficial de safras no Brasil ainda seja realizada de forma subjetiva, existe um esforço da Companhia Nacional de Abastecimento (CONAB) para o aperfeiçoamento metodológico do sistema de previsão, que inclui o Projeto GeoSafras. A presente pesquisa, que pretende contribuir para esse esforço, é baseada numa abordagem que associa a amostragem estatística, imagens de satélite, SIG e GPS, e que foi desenvolvida preliminarmente pela Embrapa Meio Ambiente para estimar, de forma objetiva e probabilística, a área plantada com determinada cultura em um Município. O objetivo da pesquisa em andamento é aprimorar o método de modo a otimizar os painéis amostrais para diferentes culturas, em diferentes regiões, e disponibilizar uma ferramenta computacional única que dê suporte a todas as fases do processo. Espera-se obter, como produto final, o “Sistema para estimação de área plantada por amostragem” (SEARA), que será um programa computadorizado, integrado e amigável, que permita, tanto a construção de painéis amostrais por área, como a análise dos dados amostrais coletados e a geração de estimativas de área plantada

e sua incerteza associada. Ao final, os recursos do programa SEARA, bem como as instruções para sua utilização serão disponibilizados via Internet. Neste trabalho é apresentado um primeiro módulo computacional para amostragem aleatória de pontos dentro dos limites municipais e duas propostas de estratificação: uma para a cultura de Café no Estado de São Paulo, que será a base do método a ser implementado no SEARA; e outra para grãos, em Minas Gerais, que é fruto dos estudos para a confecção do programa.

O acesso oportuno à informação confiável tem se tornado cada dia mais importante para os processos de tomada de decisão relacionados ao tema agrícola, tanto no âmbito nacional como no internacional. Estabelecer e desenvolver programas de levantamento de dados do setor é um componente importante para qualquer sistema de informação agrícola (LUIZ, 2005). Métodos de levantamento agrícola que considerem o crescente aumento da disponibilidade de imagens de sensoriamento remoto, bem como outros avanços tecnológicos (GPS, SIGs etc.), podem vir a se tornar as formas mais práticas para um país produzir dados básicos sobre as principais mercadorias agrícolas (FAO, 1996). Gao (2002) afirma que o advento da tecnologia GPS tem aumentado a facilidade e versatilidade na aquisição de dados espaciais, como também, diversificado os enfoques pelos quais ela pode ser integrada com o sensoriamento remoto e os SIGs.

No Brasil, o Instituto Brasileiro de Geografia e Estatística (IBGE) realiza um levantamento sistemático da produção agrícola, com dados obtidos de forma subjetiva, através de consulta a especialistas, por município, e um censo agropecuário, de periodicidade variável, com informação colhida, através de entrevistas, por estabelecimento rural (COLLARES et al., 1993). Por sua vez, a Companhia Nacional de Abastecimento (CONAB), é o órgão do Ministério da Agricultura, Pecuária e do Abastecimento (MAPA), responsável pelo levantamento de dados, organização, análises e divulgação da safra brasileira. Ela possui um dos maiores bancos de dados sobre o setor agropecuário, e disponibiliza, desde 1992, parte desse acervo, por meio da Revista “Indicadores da Agropecuária”, que contempla, entre outras, informações referentes à previsão de safras agrícolas, também baseadas em coleta de dados subjetivos (CONAB, 2005).

A deficiência dessas formas de coleta de dados, embora continue sendo apontada em trabalhos mais recentes (IPPOLITI-RAMILO, 1998; PINO, 1999), já é reconhecida há muito mais tempo devido à sua subjetividade e à magnitude dos erros (BATISTA et al., 1978).

Com o objetivo de aprimorar a qualidade das suas previsões, a CONAB iniciou, em 2004, como parte do seu Programa de Revitalização, o “Projeto de Aperfeiçoamento Metodológico do Sistema de Previsão de Safras no Brasil” -

GeoSafras, que pretende desenvolver novos métodos para previsão de safras agrícolas, de forma objetiva, para proporcionar mais qualidade às estimativas de safras até então obtidas de forma subjetiva. Esses novos métodos estão fundamentados na utilização de geotecnologias, ou seja, o apoio de imagens de satélite e a utilização de SIG e GPS para desenvolver modelos estatísticos para estimação de áreas cultivadas e modelos agrometeorológicos adaptados a cada tipo de produto agrícola, seja de culturas permanentes ou temporárias, para a previsão de produtividade. A partir destas duas variáveis básicas (área cultivada e produtividade) obtém-se a estimativa de produção, que é a informação fundamental, aguardada com a maior antecedência possível pelo setor agropecuário, necessária ao planejamento de toda a cadeia produtiva. A importância do produto deste trabalho para a sociedade está no caráter oficial da informação, aceita pelo setor como isenta de qualquer tendenciosidade (CONAB, 2005).

No âmbito do GeoSafras, a CONAB, com o apoio do Instituto Nacional de Pesquisas Espaciais (INPE), tem aplicado experimentalmente, em alguns Estados e para algumas culturas, o método desenvolvido por Luiz (2003), da Embrapa Meio Ambiente. Mas, o módulo utilizado para estimar área e a etapa de construção de painéis amostrais para cada cultura de interesse, em cada Estado investigado, são executados de forma quase artesanal (não automática), com algumas etapas subjetivas de escolha do tamanho da amostra, e com a utilização de vários programas computacionais, alguns proprietários, em diferentes fases do procedimento. O desafio que se coloca, portanto, é o do estabelecimento de um aprimoramento no método que permita otimizar os painéis amostrais para as diferentes culturas e nos diferentes Estados, e disponibilizar uma ferramenta computacional única que possa ser utilizada em todas as fases e aumente o nível de automatização do processo.

A resposta a esse desafio se dará na forma de um programa computadorizado denominado “Sistema para estimação de área plantada por amostragem” (SEARA), que será integrado e amigável, e permitirá, tanto a construção de painéis amostrais por área, como a análise dos dados amostrais coletados e a geração de estimativas de área plantada e sua incerteza associada. Para a elaboração desse programa, a equipe envolvida contará com a experiência adquirida pela Embrapa Informática Agropecuária na concepção do Agritempo. O Agritempo disponibiliza, por meio da Internet, informações meteorológicas e agrometeorológicas de diversos Municípios e Estados brasileiros. Um usuário nele cadastrado tem acesso a um menu principal, contendo chamadas para operações de atualização de dados de estações meteorológicas (nome da estação, latitude, longitude, localização) e de dados

climáticos diários (temperatura máxima, temperatura mínima e precipitação), bem como de geração de boletins agrometeorológicos, gráficos e mapas (por exemplo, estiagem agrícola, temperaturas diárias, precipitação, possibilidade de geada, entre outros), obtidos dinamicamente no momento de sua execução (EVANGELISTA et al., 2003). O módulo SEARA será incorporado à mesma plataforma do Agritempo, permitindo ao usuário cadastrado planejar seu levantamento, preparar o material de campo, proceder à entrada dos dados coletados e gerar as estimativas de área e seus níveis de confiabilidade respectivos.

## **Questões técnico-científicas**

No atual estágio da pesquisa, estão sendo desenvolvidos critérios objetivos, baseados em métodos estatísticos, que permitam fornecer subsídios à tomada de decisão relacionada às seguintes questões:

- Qual o método de estratificação ótimo considerando-se a relação entre tamanho de amostra e incerteza das estimativas? Como determinar o número ótimo de estratos? Deve ser estabelecido um número mínimo de Municípios por estrato? Se sim, como determinar esse número?

- Quando se deseja estimar a área plantada por Estado, que Municípios devem compor a população: todos; todos com área plantada diferente de zero; ou deve haver uma área mínima para inclusão? Se for indicado área mínima, qual o critério para seu estabelecimento? Quais as características das culturas que influenciam nesse valor?

- Como determinar o número de Municípios por estrato para compor a amostra? Esse número deve ser igual para todos os estratos?

- Dado um mesmo tamanho de amostra (número de estratos x número de municípios por estrato x número de pixels amostrados por município), como é que a exigência de contiguidade entre os municípios constituintes dos estratos afeta a precisão das estimativas de área?

- Para uma determinada cultura de interesse, qual a relação entre número de pixels na amostra, área do pixel, área do Município e área da cultura no Município? Como determinar, dado um nível desejado de precisão da estimativa, o número de pixels da amostra por Município?

Não menos importante é estabelecer critérios que deverão ser adotados para permitir o acesso de um usuário ao sistema SEARA:

- Será viável permitir acesso irrestrito aos usuários cadastrados? Em caso positivo, como serão considerados e tratados os dados eventualmente coletados

por usuários individuais? Como se dará a eventual divulgação desses dados?

## Revisão de Literatura

A utilização dos SIGs já se faz presente em diversos campos de aplicação, entre eles, a agropecuária (ASSAD & SANO, 1998). Entre as demais geotecnologias, o Sensoriamento Remoto tem sua utilização mais freqüente como fonte de informação do uso e cobertura da superfície da terra. Como ferramenta de apoio à construção de painéis amostrais, surpreendentemente, apenas um país o incorporou definitivamente e de forma operacional ao seu programa de estatísticas agrícolas oficiais, os EUA (LUIZ, 2003).

O uso do sensoriamento remoto, associado aos SIGs, além de facilitar a obtenção de estimativas da área plantada e da produção das principais culturas, permite o acesso a uma outra informação essencial em muitas aplicações, especialmente as ambientais, qual seja, a localização das áreas cultivadas (EPIPHANIO, 1997). Esta associação é uma ferramenta disponível para que o Estado, as Organizações não Governamentais, os Agentes Financeiros etc., possam fiscalizar o cumprimento de leis de uso do solo e dos recursos hídricos, elaborar zoneamentos agroambientais, ordenar a aplicação de recursos, e assim por diante (LUIZ, 2003).

A incorporação de avanços tecnológicos, alcançados em diversas áreas do conhecimento, aos procedimentos utilizados nos levantamentos das estatísticas agrícolas nacionais vai se configurar como um impulsionador no sentido de aprimorar, em objetividade, rapidez e confiabilidade, as estimativas oficiais da safra brasileira (LUIZ, 2003).

Para atender aos interesses dos agentes políticos e econômicos da agropecuária brasileira, que necessitam de informação a tempo de tomar decisões e intervir na cadeia produtiva, a obtenção de informação sobre área plantada e desenvolvimento das culturas deve ser um processo contínuo, periódico, rápido, de abrangência nacional, transparente, auditável e de precisão conhecida. Essas exigências impõem certas restrições aos métodos possíveis de serem adotados. A cobertura nacional, num país de mais de 5.500 municípios e 850.000.000 de hectares, só pode ser obtida através da amostragem, dificilmente pelo método censitário, exaustivo ou de mapeamento (LUIZ, 2003).

Apesar de o mapeamento temático ser uma das aplicações mais comuns do sensoriamento remoto, ainda não atingiu um estágio no qual os mapas de cobertura do terreno possam ser rotineiramente derivados de dados obtidos remotamente. Muitas são as razões que limitam o potencial do uso do

sensoriamento remoto como uma fonte de dados temáticos, tanto de cobertura do solo como de outros temas específicos. Entre elas pode-se destacar a natureza das classes, as resoluções espectrais e radiométricas dos sensores, e os métodos usados nos mapeamentos (FOODY, 1999). Adiciona-se a isso o fato de que a qualidade das imagens de satélite, atualmente disponíveis para estudo do uso do solo, são dependentes, para sua obtenção, da ausência de nuvens, o que no Brasil não é comum nos principais períodos e regiões de concentração da produção agrícola.

No caso específico do café, Moreira et al. (2004) concluíram que para mapear as lavouras por meio de imagens orbitais, é fundamental que o resultado da classificação digital seja corrigido por meio da fotointerpretação na tela do computador, para diferenciá-las de outros tipos de uso do solo com mesma resposta espectral, mas com padrão espacial diferenciado. Além disso, observaram que a reflectância do café em produção apresentou grande variabilidade entre lavouras, que pode ser atribuída à idade, espaçamento de plantas e cultivar, e que indica a necessidade de trabalho em campo para a correta identificação das lavouras de café nas imagens.

Outros trabalhos que abordam a possibilidade de se fazer o mapeamento completo da área de estudo, e não apenas amostrar parte dela, geralmente são limitados na área de abrangência, o que não atende às necessidades dos órgãos oficiais brasileiros, que precisam de números confiáveis para muitas culturas e para todo o país, ou ao menos para os principais estados produtores e não só para municípios, bacias, vales, cenas de imagens ou outras formas de delimitar regiões. Doraiswamy et al (2004), por exemplo, trabalharam sobre um mosaico de imagens do satélite MODIS (do inglês Moderate Resolution Imaging Spectroradiometer), que abrangia uma área aproximada de 900.000 ha, e apontaram que, como a área de estudo cobria apenas parcialmente o território de nove condados, não foi possível comparar seus resultados com os dados do Serviço Nacional de Estatística Agrícola dos EUA (NASS), que são disponibilizados por condado. Eles ainda citaram a dificuldade de conseguir imagens livres de nuvens (que dificulta o mapeamento). Esses autores apesar de terem obtido bons resultados na classificação automática de soja e milho, registraram péssimo resultado para alfafa e pastagem.

Já o trabalho de Lobell et al. (2003) se restringiu à área de um vale de agricultura intensiva no México e também obteve resultados com diferentes graus de precisão para distintas culturas. Enquanto a diferença encontrada entre os dados declarados pelos agricultores e os estimados por meio do mapeamento com auxílio de imagens foram pequenas para a cultura do trigo (entre 3 e 5%), foram

grandes para o milho (entre 42 e 66%). Nesse trabalho a classificação das culturas foi determinada pelo perfil temporal da presença da cultura (ciclo fenológico), estimado por uma razão simples (SR) entre bandas espectrais, e não por algum tipo de assinatura espectral das culturas. Por exemplo, pixels com alto valor para SR em outubro e julho foram identificados com o sistema de rotação milho de outono/soja de verão, enquanto que pixels com altos valores de SR em fevereiro e julho foram igualados ao sistema de rotação trigo de inverno/ soja de verão.

É preciso ainda ressaltar que as características ambientais, socioeconômicas e fundiárias no Brasil são de tal forma variáveis entre regiões espacialmente próximas, que se pode encontrar, representadas na área coberta por uma única imagem de satélite (CBERS ou Landsat, por exemplo), o cultivo de múltiplas espécies vegetais e diferentes variedades da mesma espécie, lavouras conduzidas em sistemas diversos, semeadas em diferentes datas e ocupando talhões de formas e tamanhos variados. Mesmo que isso seja mais freqüente em Santa Catarina que para o Mato Grosso, sempre se encontrará uma certa diversidade que dificulta a classificação e o mapeamento. Uma mesma cultura, como a soja, apresenta uma resposta espectral totalmente diferente aos 15, aos 30 ou aos 45 dias após a semeadura (LUIZ et al., 2001), e atualmente estão disponíveis aos agricultores brasileiros cultivares de ciclos que vão de menos de 100 a mais de 150 dias e a época de plantio se estende por 45 dias ou mais na maioria das regiões produtoras (EMBRAPA, 2004).

O presente trabalho está fundamentado em algumas pressuposições básicas. Em primeiro lugar, supõe-se que a combinação de imagens obtidas por satélite com o processamento digital de dados espaciais por meio de um SIG e com o uso do aparelho GPS para localização no campo propicia a aplicação do método de amostragem aleatória por pontos, equivalentes aos pixels das imagens, em levantamentos estatísticos para a obtenção de estimativas agrícolas, e que esse método permite gerar estimativas não tendenciosas de área ocupada por diferentes tipos de uso do solo. E ainda, que a precisão dessas estimativas é facilmente controlada através de alterações no tamanho da amostra. É possível definir como será coletada a amostra, qual o seu tamanho e quais métodos probabilísticos serão utilizados, de forma a garantir o caráter probabilístico de amostragem e identificar a população que essa amostra irá representar (COCHRAN, 1977).

Em segundo lugar, supõe-se que a estratificação do universo amostral aumenta a eficiência da amostragem. Existem métodos que buscam partições ótimas da população investigada em estratos, de forma que o tamanho total da amostra seja mínimo para uma dada precisão (LAVALLÉE & HIDIROGLOU, 1988), ou de maneira que, especificado um tamanho de amostra, se obtenha uma variância

mínima para o estimador (HEDLIN, 2000).

Por fim, a experiência na elaboração de sistemas computacionais robustos, adequados para projetos com requerimentos dinâmicos, equipe pequena e necessidade de desenvolvimento rápido (EVANGELISTA et al., 2003), permite assumir o desafio de implementar um programa computacional dedicado aos levantamentos de estatísticas agrícolas por amostragem, que, numa única plataforma e via Internet, permita associar dados tabulares, imagens de satélite, informações georeferenciadas, planejamento amostral, análise de agrupamentos, análise estatística de dados amostrais e geração de relatórios na forma de mapas, gráficos e tabelas.

## **Objetivos**

Os objetivos gerais do trabalho que está em desenvolvimento são:

- 1) Disponibilizar, na forma de um programa computadorizado com atributos de usabilidade, um método confiável, baseado no uso de imagens de sensoriamento remoto e em amostragem probabilística, que permita estimar a área ocupada por uma determinada cultura agrícola em escala municipal;
- 2) Estabelecer a forma de quantificar as incertezas associadas às estimativas de área plantada obtidas pelo método de amostragem proposto;
- 3) Definir a maneira como se dará a expansão do método de estimativa Municipal para a escala Estadual e implementá-lo no programa;
- 4) Divulgar os produtos desenvolvidos para os usuários potenciais, principalmente CONAB e IBGE.

Para alcançar esses objetivos gerais, serão perseguidos os seguintes objetivos específicos:

- a) Desenvolver um roteiro para a apropriação de imagens da câmera de alta resolução CCD (do inglês Charge Coupled Device – dispositivo de cargas acopladas) do Satélite Sino-Brasileiro de Recursos Terrestres - CBERS (do inglês China-Brazil Earth Resources Satellite) de forma a permitir seu uso em levantamentos agrícolas por amostragem;
- b) Operacionalizar no programa o acesso à plataforma do Google Earth, para fornecimento de imagem como alternativa aos dados CBERS;
- c) Estabelecer uma forma de extrair uma amostra aleatória de pixels de uma imagem, dentre um conjunto de pixels espacialmente delimitado;
- d) Descrever algoritmos para estimar área plantada por município, bem como suas variâncias, a partir dos dados obtidos por amostragem;

e) Definir algoritmos para realização de amostragem estratificada, tanto por município como em um conjunto de municípios, inclusive com os cálculos das estimativas e respectivas variâncias;

f) Realizar estudos de simulação com o método de amostragem estratificada, para verificar o efeito das alterações nos valores das restrições sobre as estimativas e variâncias.

## Métodos

O método de amostragem auxiliado por imagem de sensoriamento remoto já foi definido para um município (LUIZ, 2003 e 2005; LUIZ & EIPHANIO, 2001; LUIZ & GÜRTLER, 2003). A área do município é dividida em elementos quadrados, de área conhecida, com o auxílio de imagens de sensoriamento remoto georeferenciadas. Esses elementos (elementos de cena ou pixels – picture elements) compõem a população a ser amostrada. O método permite extrair uma amostra aleatória de pixels, que são localizados no campo com auxílio do GPS e das imagens de satélite. A proporção de pixels ocupados pela cultura de interesse é um estimador não viesado da proporção da área do município ocupada pela cultura. Culturas que ocupem uma pequena área relativamente à área total do município exigirão uma amostra maior de pixels para que as estimativas atinjam um grau desejado de confiabilidade. O problema é que, quanto maior o número de pontos a serem visitados no campo, maior o custo da amostragem. É preciso então um estudo, específico para cada cultura, que determine a melhor relação custo/benefício na estimação da área cultivada.

Para poder expandir o método de amostragem por pontos do âmbito municipal para o regional, estadual ou mesmo nacional, adota-se o procedimento da estratificação, que divide a população de municípios que se tem interesse de estudar em estratos, e tomam-se apenas alguns municípios de cada estrato como amostra. Será pesquisado o método mais adequado para obterem-se partições ótimas da população amostrada em estratos, sempre à procura de otimizar a relação entre tamanho da amostra e precisão das estimativas. Existem diversos métodos de estratificação disponíveis (COCHRAN, 1977), e será objeto deste trabalho definir aquele que reúna a facilidade de implementação e operação, tenha sólidos conceitos estatísticos, e seja adequado ao problema dos levantamentos de estatísticas agrícolas.

O desenvolvimento do sistema computacional é baseado na aplicação dos conceitos do método conhecido por Programação Extrema ou XP (do inglês eXtreme Programming), e utilizará a infra-estrutura do serviço World Wide Web

(WWW) da Internet. A XP foi desenvolvida para ser aplicada em projetos com times de dois a 10 programadores que não sejam severamente restringidos pelo ambiente computacional existente e no qual boa parte da execução de testes possa ser feita em pouco tempo (GERVAZONI, 2005). A XP tem como uma das suas principais características a figura do “cliente residente”, que atua constantemente com a equipe de desenvolvimento, e visa fornecer respostas rápidas no que se refere à definição e priorização dos requisitos a serem implementados. Neste caso, tal papel é exercido pelo coordenador do projeto, que mantém contato permanente com a equipe de desenvolvimento e as instituições CONAB e IBGE, potenciais clientes do programa.

## **Estratégia de ação**

Dados necessários para alimentar o programa já estão disponíveis (estatísticas municipais do IBGE, limites municipais, imagens CBERS e Google Earth). No desenvolvimento deste trabalho, buscar-se-á obter dados na melhor escala e os mais recentes disponíveis.

O principal insumo para esta pesquisa é a integração e troca de experiência entre as equipes envolvidas, por isso, o principal esforço é no sentido de propiciar o maior número possível de encontros de trabalho, o que é facilitado pela proximidade entre as Unidades da Embrapa envolvidas. O contato com os parceiros externos será mantido continuamente via correio eletrônico e, sempre que possível, fisicamente em reuniões de análise crítica. O desenvolvimento dos métodos estatísticos estará a cargo dos pesquisadores da Embrapa Meio Ambiente; já o desenvolvimento do programa computacional estará a cargo da equipe da Embrapa Informática Agropecuária. A participação de bolsistas é muito importante, pois, além de representar a formação de novos profissionais na área de estatísticas agrícolas e computação científica, permite a ampliação da capacidade de trabalho da equipe. Os técnicos da CONAB participam como colaboradores na discussão sobre os métodos em desenvolvimento e a operação do programa. O programa desenvolvido será disponibilizado na Internet.

## **Resultados esperados**

- a) Programa SEARA desenvolvido e disponível para uso via Internet;
- b) Melhoria na qualidade das estimativas de área plantada, pela incorporação dos critérios de objetividade e confiabilidade oferecidos pelo método;

c) Maior integração entre as equipes da Embrapa e CONAB, com benefícios mútuos no sentido de uma maior e melhor compreensão da dinâmica dos números da agricultura brasileira.

## Resultados já alcançados

Embora esteja em uma fase inicial de execução, este trabalho de pesquisa já possui alguns resultados relevantes para apresentar, fruto de testes, desenvolvimentos rápidos e demandas emergenciais, como: o módulo de amostragem probabilística de pixels; o delineamento amostral para a cultura de café no Estado de São Paulo; e o delineamento amostral para diversas culturas no Estado de Minas Gerais. Veremos a seguir alguns detalhes sobre estes resultados iniciais:

### a) Módulo de aleatorização

O módulo de amostragem aleatória dos pixels dentro do território dos municípios já está pronto e pode ser solicitado aos autores ou executado no endereço<sup>1</sup>: <http://www.agritempo.gov.br/agroclima/amostra?uf=uf>. No lugar do “uf” final, deve ser escrita a sigla do Estado ao qual pertence o município no qual se quer fazer a amostragem. Depois de aberta a página, deve ser escolhido o nome do “município” e definido o número de pixels (“tamanho da amostra”) a serem escolhidos aleatoriamente. O programa gera então os pontos e os distribui dentro dos limites do município; a descrição detalhada do processo de geração dos pontos encontra-se em Luiz & Gürtler (2003) e em Luiz & Epiphanyo (2001). É oferecida então, ao usuário, a opção de selecionar uma “imagem Google”, o que permite sobrepor os limites do município e os pixels sorteados a uma imagem obtida do Google Earth (desde que esse programa esteja instalado na máquina do usuário).

### b) Delineamento amostral para a cultura do café no Estado de São Paulo

O delineamento amostral para estimação da área plantada com café no Estado de São Paulo, Brasil, foi proposto na forma descrita a seguir. Após passar por análise crítica dos técnicos da CONAB, se aceito, o método será implantado computacionalmente como módulo do SEARA.

<sup>1</sup>Algumas configurações não permitem a execução do módulo via Internet.

Foram utilizados os dados preliminares do GCEA do IBGE para 2005. Como os dados são preliminares, eles não serão individualizados por município, mas serão apresentados de forma agrupada. De qualquer forma, esses dados foram utilizados para definir a estratificação por serem considerados os melhores e mais recentes disponíveis, o que ajuda na obtenção de estimativas de boa qualidade. Por outro lado, todos os erros eventualmente cometidos são de inteira responsabilidade dos autores, não cabendo ao IBGE nenhuma responsabilidade pelo uso dos seus dados.

Segundo o IBGE (2005a), o Estado de São Paulo possuía, em 2001, 645 municípios com área total de 24.820.942,6 hectares. As estimativas de área plantada total com café para o Estado somaram 233.878 ha em 438 municípios (207 apresentaram área zero).

Neste estudo, todos os municípios com menos de 100 hectares de café foram colocados no estrato zero, no qual não se fará a amostragem. Para o ano de 2005, além dos 207 municípios com área zero, pertencem a esse estrato mais 175 municípios com área entre um e 98 ha. Ao final, para que não haja subestimação da área de café no Estado, os valores mais recentes disponíveis para esses 175 municípios, obtidos pelo IBGE, serão considerados e somados, sem assunção de variância, ao total obtido por amostragem.

Foram considerados como população alvo, então, 263 municípios com área igual ou superior a 100 ha de café, que totalizam 228.347 ha plantados com essa cultura, ou 97,63% do total do Estado.

A amostragem é estratificada em dois estágios. A unidade amostral no primeiro estágio é o município e no segundo estágio é o pixel. A variável observada em cada sub-unidade, ou unidade do segundo estágio, é a presença (valor = 1) ou ausência (valor = 0) da cultura do café. A escolha dos elementos da amostra é realizada de acordo com a amostragem aleatória simples, sem reposição, nos dois estágios.

Para se definir o número de estratos, foram feitos exercícios com 5, 6 e 7 agrupamentos, que resultaram, ao utilizar o mesmo procedimento relatado a seguir, em variâncias mais heterogêneas entre os estratos nos casos com 5 e 7 agrupamentos. No caso de 5 estratos, a maior variância (77.773,87) foi equivalente a mais de 5 vezes a menor (14.498,31). No caso de 7 estratos, a diferença era ainda maior, pois o valor no estrato de menor variância (703,58) era quase 45 vezes menor que no estrato de maior variância (31.623,61). Já com 6 estratos, a razão entre a maior e a menor variância foi menor que 2,5, o que levou à sua escolha.

Havia, a priori, uma certa experiência por parte da CONAB que indicava uma

restrição gerencial em relação ao número total de municípios a serem incluídos na amostra, que deveria ficar próximo a 30. Como o número mínimo de municípios por estrato deve ser dois, para permitir o cálculo da variância, e se esperava trabalhar com um estrato certo<sup>2</sup>, para incluir os municípios mais importantes em área, com probabilidade de seleção igual a um, se o número de estratos aumentasse muito, as restrições gerenciais para o número de municípios não poderiam ser atendidas. Por esse motivo, não foram realizados exercícios de agrupamento para números maiores de estratos.

Foi feita uma primeira análise de agrupamento, que usou a variável área plantada, a distância euclidiana, o método do centróide, e fixou o número máximo de agrupamentos igual a seis.

No método do centróide, a distância entre dois agrupamentos (neste caso a distância euclidiana) é medida entre os seus centróides, que são os valores médios de área plantada entre todos os municípios em cada um dos agrupamentos. Neste método, toda vez que os municípios são agrupados, um novo centróide é calculado. Os centróides dos agrupamentos mudam de lugar (ou migram) toda vez que ocorre um rearranjo ou mistura de agrupamentos. Em outras palavras, há uma mudança no centróide de um agrupamento toda vez que um novo município ou grupo de municípios é adicionado a ele. A vantagem deste método é que ele é pouco afetado por valores extremos (HAIR et al., 1998).

O resultado separou os dois primeiros municípios em área plantada, cada um em um agrupamento e formou um terceiro agrupamento com apenas 3 municípios. Como o menor número a ser amostrado por estrato é 2, apenas estratos com 4 ou mais municípios serão considerados<sup>3</sup>. Assim, esses 5 municípios de maior área foram separados para comporem o estrato certo e uma segunda análise de agrupamento é feita, sob os mesmos critérios.

Nessa análise, mais um município é isolado como estrato certo e se faz uma terceira análise com os restantes.

Dessa vez, nenhum município ficou isolado, mas os dois primeiros estratos em área plantada possuíam apenas 2 e 3 municípios, respectivamente. Ao seguir a regra de número mínimo por estrato, mais esses 5 municípios são colocados no estrato certo e uma quarta análise é realizada.

Os resultados indicaram que todos os 6 estratos possuíam mais de 4 municípios e apresentavam variâncias homogêneas (a maior é cerca de 2,5 vezes a menor). Essa classificação obtida é, então, considerada aceitável e os estratos ficam definidos conforme apresentado na Tabela 1.

<sup>2</sup>Segundo Hedlin, 2000, o estrato certo (também chamado de estrato auto-representativo, estrato de enumeração completa, ou estrato todo-selecionado) é aquele no qual todas as unidades são selecionadas para observação.

<sup>3</sup>Se fossem permitidos estratos com 3 municípios, a amostra mínima de tamanho 2 seria maior que a metade da população, o que não justifica a amostragem.

**Tabela 1.** Distribuição dos municípios paulistas em estratos, de acordo com a área plantada com a cultura do café (em ha), segundo dados preliminares do GCEA/IBGE-2005.

Estrato (h)	CV (%)	Número de municípios (Nh)	Área média com café	Variância	Desvio padrão (sh)	Área total com café (Yh)	% estadual de café
0 (não amostrado)	79,8	382	14	133,76	11,6	5.534	2,37
1 (certo)	-	11	6.819	-	-	75.008	32,07
2	5,2	6	3.242	28.384,0	168,5	19.452	8,32
3	5,6	5	2.462	18.859,2	137,3	12.311	5,26
4	8,7	10	1.857	25.886,7	160,9	18.567	7,94
5	13,5	24	1.284	29.902,1	172,9	30.824	13,18
6	21,3	47	736	24.492,3	156,5	34.583	14,79
7	46,9	160	235	12.171,2	110,3	37.601	16,08
Total no Estado	289,6	645	363	1.102.600,8	1.050,0	233.880	100,00
Total nos estratos amostrados	173,8	263	868	2.276.060,0	1.508,7	228.347	97,63
Total nos estratos com sorteio	110,1	252	608	448.670,5	669,8	153.339	65,56

O próximo passo foi determinar o número de municípios a serem escolhidos para compor a amostra em cada um dos seis estratos (do 2 ao 7). Uma limitante gerencial imposta pela CONAB é que o coeficiente de variação da estimativa não seja superior a 7%. Assim, adaptado de Cochran (1977), foi calculado o tamanho da amostra em cada estrato para garantir o coeficiente de variação final desejado, pelas seguintes fórmulas:

$$nh^* = \{[\sum (Nh \times sh)]^2 / [0,07 \times \sum (Yh)]^2\} \times \{(Nh \times sh) / [\sum (Nh \times sh)]\} \quad (1)$$

$$nh = \text{Máximo} (nh^*; 2) \quad (2)$$

Nas expressões acima:

h é o número do estrato, e varia de 2 a 7;

nh é o número de municípios a serem escolhidos no estrato h;

Nh é o número total de municípios no estrato h;

sh é o desvio padrão da variável área com café no estrato h; e

Yh é a área total de café no estrato h.

Os valores utilizados nos cálculos são apresentados na Tabela 2. Na fórmula (2), a expressão “Máximo (nh\*; 2)” significa que o valor de nh será o maior valor entre o valor do nh\*, calculado pela expressão (1), e o número 2. Isso garante que nenhum estrato terá menos que 2 municípios selecionados para a amostra, mesmo que teoricamente não fosse preciso esse número para se obter o valor desejado para o coeficiente de variação.

**Tabela 2.** Etapas para o cálculo do tamanho da amostra por estrato, em São Paulo.

Estrato	Yh	Nh	sh	Nh x sh	nh
2	19.452	6	168,5	1.010,9	2
3	12.311	5	137,3	686,6	2
4	18.567	10	160,9	1.608,9	2
5	30.824	24	172,9	4.150,1	2
6	34.583	47	156,5	7.355,5	2
7	37.601	160	110,3	17.651,7	5
$\Sigma (Yh) = 153.339$		$\Sigma (Nh \times sh) = 32.463,7$			
$[0,07 \times \Sigma (Yh)]^2 = 115.212.344$		$[\Sigma (Nh \times sh)]^2 = 1.053.893.859$			

A Tabela 3 apresenta os limites inferiores e superiores de cada estrato, em termos de área com café, em hectares. Esses limites foram estabelecidos com base nos resultados das análises de agrupamentos utilizadas para definição dos estratos.

**Tabela 3.** Limites inferiores e superiores, em hectares com café, dos estratos propostos para estimativa de área no Estado de São Paulo.

Estrato	Limite inferior (ha)	Limite superior (ha)
0 (não amostrado)	-	99
1(certo)	3.825	-
2	2.775	3.824
3	2.160	2.774
4	1.600	2.159
5	1.025	1.599
6	471	1.024
7	100	470

Definidos os estratos, os técnicos da CONAB podem realizar o sorteio das amostras em cada estrato e programar o roteiro para irem a campo. Entretanto, ainda falta determinar o número de pixels a serem verificados em cada município da amostra. O procedimento adotado para essa determinação será explicado a seguir.

O número  $m_i$  representa a quantidade de pixels que serão visitados no município  $i$ . Ele é determinado pela seguinte expressão:

$$m_i^* = \{[10.000 \times (1-p_i)] / (p_i \times Ach^2)\} \quad (3)$$

$$m_i = \text{Mínimo}(m_i^*; 200) \quad (4)$$

Nas expressões (3) e (4):

$p_i$  é a proporção da área do município  $i$  que está ocupada por café; e

$Ach$  é a percentagem acumulada de área de café até o estrato  $h$ .

Na fórmula (4), a expressão “Mínimo ( $mi^*$ ; 200)” significa que o valor de  $mi$  será o menor valor entre o valor do  $mi^*$ , calculado pela expressão (3), e o número 200. Isso garante que nenhum município terá mais que 200 pixels selecionados para a amostra, mesmo que teoricamente fosse preciso esse número obtido na expressão (3). Essa é mais uma restrição gerencial, pois o ganho em precisão nesses casos é muito pequeno com o aumento do tamanho da amostra e, em alguns casos, esse número poderia ser muito alto, o que provocaria um esgotamento da capacidade de coleta de dados sem o equivalente retorno em ganho na qualidade das estimativas.

### **c) Delineamento amostral para culturas no Estado de Minas Gerais**

O delineamento amostral para estimativa da área plantada com grãos no Estado de Minas Gerais, Brasil, foi proposto na forma descrita a seguir. Nesse caso específico, os estratos não servirão para a aplicação do método objetivo de estimativa de área por pixels e, sim, para a aplicação do método subjetivo, conforme método próprio já aplicado pela CONAB.

O objetivo aqui é conferir um caráter probabilístico ao levantamento subjetivo, que têm muitos aspectos positivos e que nunca poderão ser substituídos por uma simples observação de pontos. No trabalho realizado pelos técnicos da CONAB, o conhecimento adquirido por eles, ao longo de anos, por meio das visitas aos principais atores das cadeias produtivas agrícolas, permite que sejam obtidas informações sobre tendências, mudanças tecnológicas, preços, impactos do clima, intenção de plantio, uso de insumos, e produtividade esperada, entre outros. Esse é o conhecimento que o mercado valoriza e que confere um diferencial aos números de safra produzidos pela CONAB.

No intuito de melhorar ainda mais esses dados, buscou-se preparar os roteiros de visita a campo, em Minas Gerais, para a primeira estimativa de tendência de plantio para a safra de grãos 2005/2006, com base em critérios estatísticos bem definidos, de forma a se poder estimar as incertezas associadas às estimativas produzidas.

Os dados utilizados foram os disponíveis no relatório PAM 2004 do IBGE (2005b). As culturas utilizadas para a definição do estrato certo foram: milho, soja, feijão, algodão, arroz, amendoim, mamona, trigo e sorgo, que serão alvo dos

levantamentos.

Diferentemente do caso anterior, ao invés de uma única cultura, será feito um único plano para todas as culturas citadas. Esse é um complicador em muitos sentidos, a começar pelo critério de escolha da variável ou variáveis a serem utilizadas na definição dos estratos. Por outro lado, o número de municípios que comumente são visitados nesse tipo de levantamento é muito superior ao caso do café. Aqui a restrição gerencial estabelecida foi um número em torno de 90 municípios.

A partir desse cenário, e considerando a tradição dos levantamentos, optou-se por algumas definições iniciais. Em primeiro lugar, ficou estabelecido que todos os municípios fariam parte do universo amostral. Segundo o IBGE eles somavam 853 em 2001, conforme a base utilizada. Adotou-se também uma estratificação geográfica, coincidente com as Meso-Regiões estabelecidas pelo IBGE para Minas Gerais, apenas agregando em um só estrato as Meso-Regiões do Jequitinhonha e Mucuri e, em outro, as Meso-Regiões Central e Metropolitana de Belo Horizonte, totalizando 10 estratos no Estado.

A cultura do milho, por ser a que está presente na quase totalidade dos municípios de Minas (845 dos 853) e por ocupar a maior área no Estado ( 1.319.380 ha em 2004) foi a escolhida para ser a principal determinante para seleção do tamanho das amostras e seleção de municípios para compor o estrato certo.

Os municípios foram separados por Meso-Região e ordenados, conforme a área colhida de milho em 2004 (IBGE, 2005b), do maior para o menor. Após o ordenamento, foram somadas, por município, as áreas de todas as culturas já citadas. Em reunião com participação dos técnicos locais da CONAB, foram selecionados, em cada Meso-Região, os municípios que se destacavam com muito maior área que a média dos demais ou que eram fundamentais para qualquer uma das demais culturas abrangidas pelo levantamento. Dessa forma foram escolhidos 33 municípios para compor o estrato certo (Tabela 4).

Para os municípios restantes nas Meso-Regiões, com o objetivo de determinar o número de municípios que seriam selecionados para compor a amostra em cada estrato, foi aplicado exatamente o mesmo método que foi anteriormente descrito para o café em São Paulo, exceto que foi estabelecido um coeficiente de variação máximo de 13 %. O resultado pode ser observado na Tabela 5.

Foi realizado o sorteio dos municípios de cada estrato, de forma inteiramente ao acaso, sem reposição, ou seja, nenhum município poderia ser escolhido duas vezes.

A lista dos municípios selecionados por Meso-Região, para o estrato certo e para o estrato com sorteio aleatório, é apresentada na Tabela 6.

**Tabela 4.** Distribuição dos municípios mineiros em estratos, e dados de área colhida com a cultura do milho, segundo dados PAM/IBGE-2004.

Estrato (h)	Número de municípios (Nh)	Área média com milho	Variância	Desvio padrão (sh)	Área com milho (Yh)	% no estado
Certo	33	10.744	-	-	354.568	26,9
Alto Paranaíba	26	3.941	11.124.785	3.335	102.466	7,8
Noroeste	16	4.266	11.913.048	3.452	68.256	5,2
Triângulo	30	2.981	7.767.794	2.787	89.430	6,8
Centro Oeste	53	1.696	2.140.080	1.463	89.888	6,8
Sul	152	1.730	2.810.477	1.676	262.960	19,9
Central/Metropolitana	153	693	543.211	737	106.029	8,0
Rio Doce	99	628	443.607	666	62.172	4,7
Zona da Mata	138	566	357.909	598	78.108	5,9
Jequitinhonha/Mucuri	66	295	137.450	371	19.470	1,5
Norte	87	992	1.138.501	1.067	86.304	6,5
<b>Total no Estado</b>	<b>853</b>	<b>1.547</b>			<b>1.319.380</b>	<b>100</b>

**Tabela 5.** Informação utilizada no cálculo do tamanho da amostra por estrato, em Minas Gerais.

Estrato	Yh	Nh	sh	Nh x sh	nh
Alto Paranaíba	102.466	26	3.335	86.720	5
Noroeste	68.256	16	3.452	55.224	3
Triângulo	89.430	30	2.787	83.612	5
Centro Oeste	89.888	53	1.463	77.534	5
Sul	262.960	152	1.676	254.820	15
Central/Metropolitana	106.029	153	737	112.765	7
Rio Doce	62.172	99	666	65.938	4
Zona da Mata	78.108	138	598	82.559	5
Jequitinhonha/Mucuri	19.470	66	371	24.469	2
Norte	86.304	87	1.067	92.829	6

$$\Sigma (Yh) = 965.083$$

$$\Sigma (Nh \times sh) = 936.471,4$$

$$[0,07 \times \Sigma (Yh)]^2 = 876.978.624.451$$

$$[\Sigma (Nh \times sh)]^2 = 15.740.409.827$$

**Tabela 6.** Municípios selecionados para serem visitados para o levantamento de safra de grãos 2005/2006, pela CONAB, no Estado de Minas Gerais.

Meso-Região	Estrato Certo	Estrato com sorteio
Alto Paranaíba	Perdizes, Nova Ponte, Sacramento, Ibiá, Coromandel	Araxá, Romaria, Monte Carmelo, Carmo do Paranaíba, Abadia dos Dourados
Noroeste	Unaí, Buritis, Paracatu	Guarda-Mor, Arinos, Brasilândia
Triângulo	Uberaba, Uberlândia, Conceição das Alagoas, Monte Alegre de Minas, Tupaciguara	Conquista, Canápolis, Frutal, Veríssimo, União de Minas
Centro Oeste	Formiga, Piumhi, Bambuí	Carmo do Cajuru, Moema, Carmópolis de Minas, Santo Antônio do Monte, Córrego Danta
Sul	Três Corações, Passos, Carmo do Rio Claro	Cambuquira, Campanha, São João da Mata, Ingá, Pratápolis, Pouso Alegre, Ilicínea, Cabo Verde, Luminárias, Claraval, Brasópolis, Piranguçu, Toledo, Inconfidentes, São João Batista do Glória
Central/ Metropolitana	Entre Rios de Minas, Madre de Deus de Minas, Carandaí, Lagoa Dourada, São João Del Rei	Itaguara, Ouro Preto, Papagaios, Fortuna de Minas, Araçá, Igarapé, Taquaraçu de Minas
Rio Doce	Aimorés, Água Boa, Mutum	Ipanema, José Raydan, Virgolândia, Dolores de Guanhães
Zona da Mata	Alto Rio Doce, Ervália, Piranga, Santa Margarida	Simonésia, Pedra Bonita, Bias Fortes, Piraúba, Bicas
Jequitinhonha/ Mucuri	-	Comercinho, Frei Gaspar
Norte	Buritzeiro, Chapada Gaúcha	Montes Claros, Icarai de Minas, Montalvânia, Japonvar, Santa Fé de Minas, Verdelândia

## Conclusão

Na fase atual do trabalho, há ainda questões técnicas importantes a serem estudadas até chegar-se a um método bem definido, que atenda todos os requisitos para a previsão de safras e possa modificar os processos atuais com elevação nos níveis de confiabilidade dos dados. Por outro lado, a demanda clara por parte da CONAB, a interação de diversas instituições e pesquisadores com conhecimentos complementares e, por fim, o uso integrado de diversas técnicas

(SIG, GPS, Sensoriamento Remoto e amostragem) aponta para um avanço concreto no processo de previsão de safras, objetivo final do esforço empreendido pela equipe deste trabalho. Esta percepção é reforçada pelos resultados iniciais aqui apresentados, que demonstram a viabilidade técnica da solução apresentada e pela identificação clara dos pontos obscuros a serem esclarecidos com o andamento dos trabalhos.

## **Referências Bibliográficas**

**ASSAD, E.D.; SANO, E.E. (Ed.).** Sistema de informações geográficas: aplicações na agricultura. 2.ed. Brasília: Embrapa SPI, 1998. 434p.

**BATISTA, G.T.; MENDONÇA, F.J.; LEE, D.C.L.; TARDIN, A.T.; CHEN, S.C.; NOVAES, R.A.** Uso de sensores remotos a bordo de satélite e aeronave na identificação e avaliação de áreas de culturas para fins de previsão de safras. São José dos Campos: INPE, 1978. 41p. (INPE-1229-NTI/103).

**COCHRAN, W.G.** Sampling techniques. 3.ed. New York: John Wiley & Sons, 1977. 428p.

**COLLARES, J.E.R.; LAURIA, C.A.; CARRILHO, M.M.** Pesquisa de previsão e acompanhamento de safras baseada em painéis de amostras de áreas. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 7., 1993, Curitiba. Anais... São José dos Campos: INPE, 1993. v.4, p.450-453.

**COMPANHIA NACIONAL DE ABASTECIMENTO (CONAB).** Indicadores da Agropecuária, v.14, n.7, jul. 2005. 61p. Disponível em: <<http://www.conab.gov.br/download/indicadores/pubindicadores.pdf>>. Acesso em: 16 ago. 2005.

**DORAISWAMY, P.C.; HATFIELD, J.L.; JACKSON, T.J.; AKHMEDOV, B.; PRUEGER, J.; STERN, A.** Crop condition and yield simulations using Landsat and MODIS. Remote Sensing of Environment, v.92, p.548-559, 2004.

**EMBRAPA.** Tecnologia de produção de soja – região central do Brasil – 2005. Londrina: Embrapa Soja, 2004. 239p. (Sistemas de Produção, 6).

**EIPHANIO, J.C.N.** Monitoreo ambiental y sistemas de información para la evaluación de la relación agricultura-medio ambiente. In: **VIGLIZZO, E.** (Coord.). Libro verde: elementos para una política agroambiental en el Cono Sur. Montevideo: IICA – PROCISUR, 1997. Cap.3, p.35-51.

**EVANGELISTA, S.R.M.; TERNES, S.; SANTOS, E.H. dos; ASSAD, E.D.; ROMANI, L.A.S.; OTAVIAN, A.F.** Monitoramento agroclimatológico pela Web. In: **CONGRESSO BRASILEIRO DA SOCIEDADE BRASILEIRA DE INFORMÁTICA APLICADA À AGROPECUÁRIA E AGROINDÚSTRIA (SBI-Agro)**, 4., Porto Seguro, 2003. Anais... Lavras: UFLa, 2003.

**FOODY, G.M.** The continuum of classification fuzziness in thematic mapping. Photogrammetric Engineering and Remote Sensing, v.65, n.4, p.443-451, 1999.

**FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO).** Multiple frame agricultural surveys: volume 1 current survey based on area and list sampling methods. Rome: FAO, 1996. 119p. (FAO Statistical Development Series, 7)

**GAO, J.** Integration of GPS with remote sensing and GIS: reality and prospect. Photogrammetric Engineering and Remote Sensing, v.68, n.5, p.447-453, 2002.

**GERVAZONI, T.P.** XP – Extreme Programming – Parte 1. Disponível em: <[http://www.linhadecodigo.com.br/artigos.asp?id\\_ac=764](http://www.linhadecodigo.com.br/artigos.asp?id_ac=764)>. Acesso em: 28 dez. 2005.

**HAIR JR., J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C.** Multivariate data analysis. 5.ed. Upper Saddle River: Prentice Hall, 1998. 762p.

**HEDLIN, D.** A procedure for stratification by an extended Ekman rule. Journal of Official Statistics, v.16, n.1, p.15-29, 2000.

**IBGE.** Cartografia. Disponível em: <[http://www.ibge.gov.br/home/geociencias/areaterritorial/pdf/areas\\_2001\\_35.pdf](http://www.ibge.gov.br/home/geociencias/areaterritorial/pdf/areas_2001_35.pdf)>. Acesso em: 08 set. 2005. 2005a.

**IBGE.** Produção agrícola municipal: cereais, leguminosas e oleaginosas: 2004. Rio de Janeiro: IBGE/ Coordenação de Agropecuária, 2005b. 57p. Acompanha CD-ROM.

**IPPOLITI-RAMILO, G.A.** Imagens TM/Landsat-5 da época de pré-plantio para a previsão da área de culturas de verão. **1998. 182p. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 1998. (INPE-7116-TDI/668).**

**LAVALLÉE, P.; HIDIROGLOU, M. A.** On the stratification of skewed populations. *Survey Methodology*, v.14, n.1, p.33-43, 1988.

**LOBELL, D.B.; ASNER, G.P.; ORTIZ-MONASTERIO, J.I.; BENNING, T.L.** Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. *Agriculture, Ecosystems and Environment*, v.94, p.205–220, 2003.

**LUIZ, A.J.B.; BARROS, M.G.M.; FORMAGGIO, A.R.; EIPHANIO, J.C.N.** Comportamento espectral associado a parâmetros agronômicos de soja (*Glycine max*) e feijão (*Phaseolus vulgaris*). In: **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 10., 2001, Foz do Iguaçu. Anais... São José dos Campos: INPE, 2001. CD-ROM**

**LUIZ, A.J.B.; EIPHANIO, J.C.N.** Amostragem por pontos em imagens de sensoriamento remoto para estimativa de área plantada por município. In: **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 10., 2001, Foz do Iguaçu. Anais... São José dos Campos: INPE, 2001. Disponível na biblioteca digital URLib: <<http://iris.sid.inpe.br:1908/rep/dpi.inpe.br/lise/2001/09.13.10.55>>.**

**LUIZ, A.J.B.** Estatísticas agrícolas por amostragem auxiliadas pelo sensoriamento remoto. **2003. 116p. Tese (Doutorado) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2003.**

**LUIZ, A.J.B.** Estatísticas agrícolas por amostragem auxiliadas pelo sensoriamento remoto. In: **SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 12., 2005, Goiânia. Anais... São José dos Campos: INPE, 2005. p.181-188.**

**LUIZ, A.J.B., GÜRTLER, S.** Aleatorização de pontos no território de um município, usando o **SPRING**, para a estimativa de área agrícola por amostragem. In: **CONGRESSO BRASILEIRO DA SOCIEDADE BRASILEIRA DE INFORMÁTICA APLICADA À AGROPECUÁRIA E AGROINDÚSTRIA (SBI-Agro), 4., Porto Seguro,**

**2003.** Anais... Lavras: UFLa, 2003. v.1, p.9-12.

PINO, F.A. Estatísticas agrícolas para o século XXI. **Agricultura em São Paulo**, v.46, n.2, p.71-105, 1999.

**Embrapa**

---

*Meio Ambiente*

Ministério da Agricultura,  
Pecuária e Abastecimento

