

INPE-5468-NTC/307

**OS RECENTES AVANÇOS TECNOLÓGICOS DO PROGRAMA DE PREVISÃO
DE SAFRA DOS ESTADOS UNIDOS**

Sherry Chou Chen

**INPE
SÃO JOSÉ DOS CAMPOS
Novembro de 1992**

SUMÁRIO

	<u>Pág.</u>
1 <u>INTRODUÇÃO</u>	1
2 <u>PESQUISAS ATUAIS DO NASS</u>	1
3 <u>OS SISTEMAS DE COMPUTADOR</u>	4
4 <u>O PROJETO "CASS" - (Computer Aided Stratification and Sampling Project)</u>	5
BIBLIOGRAFIA.....	11

OS RECENTES AVANÇOS TECNOLÓGICOS DO PROGRAMA DE PREVISÃO DE SAFRA DOS ESTADOS UNIDOS

1 INTRODUÇÃO

A previsão de safras agrícolas é uma tarefa prioritária do governo, tendo em vista sua importância para a economia do País. Nas últimas duas décadas, o INPE tem desempenhado papel importante nas áreas de pesquisa e de disseminação da tecnologia de Sensoriamento Remoto com a finalidade de aprimorar a qualidade de estimativas agrícolas.

Como o uso de técnicas computacionais é indispensável quando se trata de grandes volumes de dados, o acompanhamento dos últimos avanços tecnológicos, tanto em hardware como em software, é crucial para produzir estimativas mais precisas em tempo hábil.

Este relatório apresenta as mais recentes atividades de pesquisa de United States Department of Agriculture (USDA) para o programa de previsão de safras, que conheci durante visita técnica realizada em agosto, 1992. As experiências da USDA, especialmente do projeto CASS, servirão como indicações para o projeto SIAGRO, da cooperação INPE/IBGE.

O Serviço Nacional de Estatística Agrícola (National Agricultural Statistics Service-NASS), do USDA (United State Department of Agriculture), tem a responsabilidade de fornecer estimativas sobre qualquer aspecto da agricultura dos EUA. Com esta função, a revisão constante nos procedimentos de coleta de dados e da metodologia de estimativa é uma tarefa fundamental do NASS para acompanhar os avanços tecnológicos. A Divisão de Levantamento do NASS é responsável por pesquisa e desenvolvimento de metodologias estatísticas e por aplicações de tecnologias avançadas com o objetivo de aprimorar os programas operacionais atuais e futuros da agência.

2 PESQUISAS ATUAIS DO NASS

As novas tecnologias, tais como os novos computadores, os métodos de coleta de dados e as tecnologias de processamento de dados são avaliados

constantemente pelos dois setores da Divisão do USDA: o Setor de Sensoriamento Remoto (SSR), que administra as atividades da agência relativas o uso de sensoriamento remoto para previsão de safras e o Setor de Pesquisa e Tecnologia (SPT), que é responsável pela pesquisa e pelo desenvolvimento relativos à implementação de novas tecnologias. Alguns tópicos de pesquisas atuais do SPT são relatados a seguir:

- a) A pesquisa da edição interativa de dados vai avaliar e implementar o uso de software interativo no PS-LAN's para detectar os dados suspeitos para revisão. O software de edição interativa Blaise, do Bureau Central de Estatísticas da Holanda, foi instalado em todos os escritórios do Statistical Survey Office (SSO) e em sua administração central.
- b) Os recentes avanços tecnológicos na área de computadores portáteis permitem uma nova abordagem na coleta de dados e os projetos de pesquisa que utilizam esta abordagem são denominados estudos de Mobile Computerized Data Entry and Edit (McDEE). Os estudos de McDEE incluem a entrevista da pessoa com o computador (CAPI-Computer Assisted Personal Interview) e a coleta de dados "in loco" onde a entrevista é dispensada. Os dados coletados são transmitidos via linha telefônica (via "modem"). Esta abordagem substitui o antigo método de levantamento onde os dados eram registrados com lápis em papel. McDEE tem o potencial de coletar dados de modo mais rápido, mais barato, com menos erros e incorpora vários fatores tecnológicos na área de coleta de dados. Como os dados já são coletados em forma digital, eles podem ser enviados diretamente para escritório via "modem", minimizando os possíveis erros causados no processo de transcrição manual. Um exemplo prático é o uso de um "laptop" pelo enumerador para registrar e emitir o recibo de venda de gado.
- c) Outro projeto usa a tecnologia de sistemas de informações geográficas (SIG) para analisar os dados georeferenciados. O NASS tem a responsabilidade de formar um banco de dados de SIG como parte do Programa de Qualidade da Água da USDA e o ARCINFO foi escolhido como o software. O uso de receptores do Global Positioning System (GPS) para referenciar a latitude e a longitude de dados coletados está sendo investigado.

- d) Para acessar com eficiência os dados de propriedades coletados no programa de levantamento, uma biblioteca de SAS é estabelecida. Os dados são transportados de computador "mainframe" e arquivados "in-house" nas fitas magnéticas acessíveis com estação de trabalho de SUN. Os usuários podem recuperar os dados e usá-los com Sun SAS ou PC SAS.
- e) Outros projetos que estão nos estágios iniciais de pesquisa são: a avaliação do uso de software PC-WAVE como ferramenta para mostrar os gráficos estatísticos avançados; a utilização conjunta de sistemas operacionais de DOS e Unix na mesma rede para as aplicações de dados e a avaliação da tecnologia de varredura eletrônica para arquivar os relatórios históricos.

Os itens de pesquisas do Setor de Sensoriamento Remoto (SSR) incluem: a utilização de dados de TM-Landsat para melhorar as estimativas de área de culturas na região do Delta do Rio Mississippi, e o uso de dados de NOAA-AVHRR para monitorar condições de culturas.

- a) O projeto do Delta do Rio Mississippi, em 1991, implementou um estimador de regressão para avaliar as áreas de cultura estaduais e municipais de modo operacional usando os dados digitais TM-Landsat. Este projeto foi o primeiro teste do NASS usando o grande volume de dados TM-Landsat. Os dados digitais de TM-Landsat, na primavera e no fim de verão, foram usados para uma análise multitemporal. Na aplicação anterior de sensoriamento remoto na previsão de safras foi usado um computador "mainframe" alugado e volumes grandes de produtos de TM-Landsat em papel cujos dados foram ajustados manualmente sobre mesa de luz. O atual sistema usa microcomputador PCs ligados a um MicroVAX 3500 via rede. Os recentes melhoramentos do software "PEDITOR" permitem os usuários analisar os dados de satélite no ambiente computarizado dispensando a utilização das informações registradas em papel. Além de estimativas numéricas, os mapas coloridos de uso da terra gerados a partir dos dados de satélite podem ser produzidos para cada município incluído no projeto de 1991, usando uma impressora termal Tektronix 4693DX.

- b) Um projeto de cooperação entre o NASS e o Laboratório de Sensoriamento Remoto do Agricultural Research Service (ARS) foi firmado com o objetivo de estudar o uso de dados de satélite meteorológico NOAA-AVHRR para monitorar as condições das culturas. Até hoje, nenhuma técnica de sensoriamento remoto foi capaz de fornecer indicações diretas sobre rendimento de cultura com aceitável precisão e em tempo hábil. Entretanto, progressos significativos têm sido feitos na área de monitoramento das condições das culturas. Esta área engloba tópicos tais como; comparação entre o crescimento das culturas do ano corrente com o(s) do(s) ano(s) anterior(es), comparação dos crescimentos das culturas, dentro de um determinado ano, entre Estados ou municípios e o monitoramento da seca. O índice de vegetação diferença normalizada (NDVI), calculado com os dados de AVHRR-NOAA e fornecido bissemanalmente pelo Centro de dados do EROS, também será avaliado. As pesquisas iniciais sobre monitoramento das condições das culturas serão concentradas: (a) na avaliação de saídas do NDVI da impressora colorida, para formar um banco de dados históricos de NDVI e (b) no uso do índice NDVI em modelos de rendimento de trigo de verão. Uma estação de trabalho, DEC VAXStation, foi comprada para este projeto; ela será utilizada com a última versão do software LAS (Land Analysis System) desenvolvido pelo United States Geological Survey (USGS).

3 OS SISTEMAS DE COMPUTADOR

Para cumprir as funções requisitadas, várias tecnologias de microcomputador são adaptadas no USDA. Na análise de grandes volumes de dados de sensoriamento remoto, o "VAXCLUSTER", com um MicroVAX 3500 e uma VAXStation 3100, é usado. Outras aplicações são realizadas num sistema UNIX que utiliza um servidor SUN 4/380 com estações de trabalho de SPARC e SUN IPC (em modo PC ou em rede). Os servidores podem usar as fitas magnéticas de 6250 bpi ou Exabyte além de outros "drives" e periféricos externos. Na análise rotineira o PC/386 ou de maior potencial e a estação de trabalho em rede com o servidores de SUN e VAX são usados. Todos os servidores, estações de trabalho e PCs são ligados com a rede ETHERNET utilizando o servidor "Network File", DECNET e os protocolos TCP/IP. Os equipamentos periféricos incluem monitores coloridos de

alta resolução, varreduras, câmeras de vídeo e mesas digitalizadoras. Outros computadores tais como "laptop" e "notebook" também são utilizados (GRID Pads e Zenith Supersports).

4 O PROJETO "CASS" - (Computer Aided Stratification and Sampling Project)

O projeto CASS já está operacional devido ao esforço de pesquisa dos últimos três anos. As estimativas de área cultivada são baseadas na enumeração de cada cultura plantada nos segmentos que foram selecionados do painel amostral de área de cada Estado.

O sistema CASS incorpora os dados digitais de TM-Landsat e Digital Line Graph (DLG) do USGS. Os dados TM-Landsat, na escala de 1:100.000, são usados para delinear o uso da terra de acordo com os estratos definidos no estudo. Na definição de limites dos estratos, os dados de DLG (1:100.000) são sobrepostos nas imagens TM usando um plano gráfico. As seguintes etapas são definidas no projeto CASS:

a) Visualização de Dados de Satélite

Três das sete bandas do TM são usadas com resolução plena (i.e., 30m) para visualização de imagem colorida. A imagem mostrada no monitor pode ser ampliada várias vezes e fornece a visão mais recente da área a ser estratificada. O uso deste produto digital é equivalente ao uso de imagem de satélite em papel colorido e de fotografia aérea P&B de alta altitude.

b) Realce de Imagem

O monitor usa 8 bits de resolução gráfica para cada banda produzindo uma imagem colorida nítida, e o usuário pode manipular o nível de cinza e o contraste para realçar a perceptibilidade da imagem. A imagem realçada é gravada para uso posterior. Outras funções permitem o usuário visualizar uma única banda ou seu histograma e realçar a imagem com a transformação linear, "piecewise" ou equiprovável. A imagem gravada para o uso posterior é aquela que distinguir melhor as culturas e seus limites.

c) Registro com os Dados DLG

Os dados DLG são usados como uma referência para definir os limites das Unidades Primárias de Amostragem (UPAs) e dos segmentos. No momento, somente as informações de estradas e hidrografia são incorporadas no DLG enquanto os limites políticos e administrativos vão ser disponíveis em breve. Os arquivos DLG, em fita, são regravados em disquete, por município. Para registrar o DLG com a imagem, vários pontos na imagem e seus correspondentes no DLG foram selecionados para estabelecer uma equação de regressão de mínimo quadrado para ajustar os pontos restantes. Os pontos usados e a equação foram armazenados em um arquivo de registro e este arquivo também pode ser usado para determinar as coordenadas de longitude e latitude de qualquer ponto na imagem.

d) Determinação do Limite de UPA

Interpretando a imagem colorida mostrada no monitor o(a) analista estima a percentagem da área cultivada e rotula o estrato apropriado para a área (em nível de UPA) que esta sendo analisada. Nesta etapa há uma grande diferença entre o sistema CASS e o procedimento antigo onde os estratos numa imagem inteira foram delineados e depois divididos em UPAs. Após a definição do estrato apropriado, o limite da UPA é delineado de acordo com os limites permanentes (estrada, rio etc.) visíveis na imagem ou no DLG e seu tamanho deve seguir a especificação de cada estrato. Uma vez que o limite da UPA é fechado, sua área é imediatamente calculada para verificar se o tamanho do polígono é aceitável de acordo com o critério pré-estabelecido do estrato. Outras opções disponíveis são: dividir, englobar, "undo" o polígono; listar as UPAs já criadas; verificar sua numeração e a área além de verificar se há polígonos sobrepostos ou área não rotulada.

e) Divisão das UPAs Selecionadas em Segmentos

Após a estratificação do Estado inteiro, a área de cada estrato é calculada e um programa de amostragem é aplicado para selecionar as UPAs que devem ser divididas em segmentos (i.e., unidade secundária de amostragem). No processo da divisão de UPA em segmentos de tamanhos semelhantes, várias funções utilizadas na fase de estratificação e divisão de UPA citadas em (d) foram usadas. Os mesmos procedimentos de controle de qualidade da etapa (d) também são executados. Uma vez que a UPA foi

dividida em segmentos, um deles é sorteado aleatoriamente com o comando da seleção de segmento. Esta seleção é a segunda fase de amostragem (i.e., a primeira é a seleção das UPAs).

Um monitor colorido, de alta resolução, da estação de trabalho CASS é usado para visualizar a imagem, enquanto outro monitor de "menu" mostra os software. O teclado é usado para mandar os comandos do menu; o "mouse" é usado para interagir com o monitor de imagem. Outra função das setas é para mudar a cor da imagem, ligá-la; desligá-la ou ampliá-la.

A estação de trabalho UNIX da Hewlett-Packard (HP) é usada para processar e armazenar este grande volume de dados. Esta estação possui a capacidade mínima necessária para a construção do painel amostral de área; i.e., com 3 planos de imagem, 4 planos gráficos sobrepostos e um sistema de coordenadas de 1024x1280. Esta configuração permite a visualização de imagem colorida usando os 3 planos de imagem, enquanto os planos gráficos são usados para mostrar o DLG, as UPAs e o menu. Uma recente modificação permite que o DLG seja mostrado em dois planos; uma cor para as estradas e outra para a hidrografia.

Alguns benefícios na utilização do CASS são apresentados a seguir:

- O dado digital TM-Landsat é a informação mais atualizada do que as fotografias aéreas de alta altitude e tem melhor resolução do que seu produtos analógicos;
- Os trabalhos de "papel e lápis", i.e., transferir as informações da imagem analógica de satélite à fotografia aérea e desta para um mapa topográfico do USGS (1:100.000) e finalmente para um arquivo digital são eliminados tornando o processo mais rápido, preciso e eficiente. No sistema antigo duas semanas de trabalho eram necessárias para preparar o painel amostral de um município. No sistema CASS a mesma tarefa é feita durante 2 a 3 dias;
- No antigo processo a área da UPA é somente calculada após a digitalização manual de seu limite. No sistema atual as áreas de UPAs e segmentos podem ser determinadas logo após o delineamento dos polígonos, facilitando assim os processos de estratificação e a seleção de amostras;

- A construção de painel amostral usando os dados digitais de satélite torna a tarefa de atualização um processo relativamente fácil;
- A seleção de segmentos amostrados é feita através de um "software" que gera os números aleatórios em vez de procurar os números na tabela;
- Fácil revisão de UPAs;
- Fácil exploração do painel amostral de área para uma classe especial;
- Um melhor controle da homogeneidade das UPAs num estrato.

As considerações futuras para o CASS são:

- O teste da utilização de classificação supervisionada e não-supervisionada como um auxílio no processo de estratificação.
- Atualmente as bandas 2, 3 e 4 do TM-Landsat foram usadas na construção de painel amostral. A banda 5 que tem a maior nitidez devido sua maior penetração de nuvens será incluído no futuro.
- A utilização da banda pancromática (i.e., 15m) do futuro Landsat-6 juntamente com as bandas multiespectrais com 30m de resolução será testada.
- A investigação do uso de filtros espaciais para detecção de bordas ou limites.
- A investigação da utilização de interface gráfica de usuário (i.e., substituindo o terminal de menu por "ícones").
- O uso de outras fontes de dados de censo (i.e., "Census Tiger data and Public Law File") para localizar os limites dos blocos de censo no CASS e verificar os números de residências para melhorar a classificação da área urbana. Os mapas topográficos quadrantes também podem ser usados na classificação urbana e para definição de seus limites apropriados.

- A possibilidade de converter os dados dos polígonos num formato compatível com o SIG (Sistema de Informações Geográficas), assim facilitando o transporte dos dados entre SIG e CASS deve ser investigada.

BIBLIOGRAFIA

- U.S. Department of Agriculture. Applications of advanced technology for agricultural statistics. Seminar on integrated statistical information system and related matters, Bratislava, Czech and Slovak Federal Republic, 18-22 May, 1992.
- Cotter, J.J.; Mazur, C., Automating the development of area sampling frames using digital data displayed on a graphhic workstation; NASS Staff Report, USDA - Washington. DC, 1991.
- Ozga, M.; Mason, W.; Craig, M., Peditor - current status and improvements. ISPRS Congress, Washington, DC, Aug. 3-7, 1992.

**MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

INPE-5468-NTC/307

**OS RECENTES AVANÇOS TECNOLÓGICOS DO PROGRAMA DE PREVISÃO DE
SAFRA DOS ESTADOS UNIDOS**

Sherry Chou Chen

**INPE
São José dos Campos
Novembro de 1992**

CDU: 528.711.7:632.164

PALAVRAS-CHAVE: Previsão de Safra; Sensoriamento Remoto.