



# Metodologia para identificação e caracterização de erros em mapas de produtividade<sup>1</sup>



Leonardo A. Angeli Menegatti<sup>2</sup> & José P. Molin<sup>3</sup>

<sup>1</sup> Extraído da Dissertação de Mestrado do primeiro autor, na ESALQ/USP, com apoio financeiro da FAPESP

<sup>2</sup> Depto. de Eng. Rural, ESALQ/USP. Av Pádua Dias 11, CP 09, Piracicaba, SP. E-mail: laameneg@hotmail.com (Foto)

<sup>3</sup> Depto. de Eng. Rural, ESALQ/USP. E-mail: jpmolin@esalq.usp.br

Protocolo 17 - 19/2/2002 - Aprovado em 9/5/2003

**Resumo:** O mapa de produtividade é considerado, por muitos autores, o início da Agricultura de Precisão. A tecnologia para a coleta automatizada de pontos de produtividade proporciona intensa densidade amostral com a inevitável coleta de pontos que contêm erros na medição da produtividade e que podem interferir na análise e na tomada de decisão baseada no mapa. Este trabalho tem por objetivo desenvolver uma metodologia para identificar e caracterizar erros em mapas de produtividade, e aplicá-la para seis diferentes monitores de produtividade comerciais. Nos dados analisados foram encontrados erros de posicionamento, pontos com produtividade nula ou extrema, ausência de medida de umidade dos grãos, largura de plataforma incorretamente registrada pelo operador, dados inconsistentes durante o intervalo referente ao tempo de enchimento e de esvaziamento da colhedora nos inícios e finais de passadas e, também, dados discrepantes. Todos esses erros foram devidamente caracterizados e forneceram bases para a formulação de uma rotina de filtragem de dados brutos.

**Palavras-chave:** agricultura de precisão, filtragem, monitor de produtividade

## Methodology for identification and characterization of errors in yield maps

**Abstract:** For many researchers, yield mapping is the start of Precision Farming. Yield monitors measure the yield and collect data in a high frequency allowing the collection of erroneous data that may interfere in the decision-making based on yield maps. The objective of this work was to develop a methodology to identify and characterize errors in yield maps and apply the methodology to a set of data obtained from six different commercial yield monitors. With the analyzed data, errors were found and characterized as positioning errors, null or extremely high yield points with no measurements of grain moisture, swath width errors due to wrong interpretation of the effective cutting width by the operator, time errors and disagreeable data. All these errors were characterized and furnished the basis for the formulation of a routine for filtering the new yield data.

**Key words:** precision agriculture, filter, yield monitors

## INTRODUÇÃO

Agricultura de Precisão é um sistema de gerenciamento agrícola baseado na variação espacial de propriedades do solo e da planta encontrada nos talhões, visando à otimização do lucro, sustentabilidade e proteção do ambiente. Trata-se de diferentes tecnologias que são levadas ao campo para permitir um sistema de manejo que considere a variabilidade espacial da produção. De acordo com Cambardella & Karlen (1999) a Agricultura de Precisão contém três componentes primários: o sistema de posicionamento global (GPS) que fornece a posição onde o equipamento está localizado, mecanismos para controle de aplicação de nutrientes, defensivos agrícolas, água ou outros insumos em tempo real, e um banco de dados que fornece a

informação necessária para desenvolver as relações causa efeito e as respostas à aplicação de insumos a várias condições especificamente localizadas. Segundo os mesmos autores, as tecnologias associadas aos componentes eletrônicos serão bastante avançadas se comparadas com o desenvolvimento na interpretação das bases de dados, necessárias para se obter os benefícios econômicos e ambientais que a Agricultura de Precisão (AP) pode fornecer.

O mapeamento da produtividade é considerado, por muitos pesquisadores, como sendo uma parte essencial da agricultura de precisão (Makepeace, 1996). Tylén et al. (1997) consideraram que o mapa de produtividade é um dos métodos mais corretos para estimar a heterogeneidade de um campo. Molin (2000) reforça que essa informação é o ponto de partida tido por

usuários e pesquisadores para se visualizar a variabilidade espacial das lavouras, por materializar a resposta da cultura. Um mapa de produtividade evidencia regiões com alta e baixa produtividade, mas não explica a causa de tal variação, que pode ser por doenças, baixo pH, estresse hídrico e outras; se a causa específica não é determinada, nenhuma resposta ao gerenciamento pode ser obtida (Lark & Stafford, 1997).

A coleta automática de dados permite que uma grande quantidade de dados seja obtida com custo relativamente baixo. Os mapas de produtividade são elaborados a partir de dados coletados automaticamente e existe um considerável número de erros sistemáticos introduzidos, sendo que apenas alguns desses erros são eliminados por determinados programas computacionais (Moore, 1998). Os dados brutos devem ser cuidadosamente analisados antes da elaboração de um mapa de produtividade (Missotten et al., 1996).

Thylén et al. (1997) identificaram diversos erros em mapas de produtividade e os classificaram como relacionados às características físicas da colhedora, ao sistema de medição da produtividade, relacionados a variações dentro do talhão, ao operador e, ainda, devido ao procedimento de obtenção do mapa. Moore (1998) trabalhando com dados de seis anos consecutivos de mapeamento de produtividade, identificou os seguintes erros: suavização, calibração de volume, largura incorreta da plataforma e tempo de enchimento da colhedora. Blackmore & Marshall (1996) listaram várias fontes de erros em mapas de produtividade e incluíram aqueles oriundos da retrilha de grãos e das perdas de colheita.

O tempo de enchimento da colhedora é representado pela distância necessária para que o sistema de monitoramento da produtividade fique sensível às variações naturais da produtividade, após o início de gravação dos dados (Moore, 1998). O autor considerou o monitor sensível às variações naturais, quando o valor de produtividade atingiu 90% da produtividade máxima local.

Na produção dos mapas são coletados dados em grande quantidade para a geração de isolinhas de produtividade, a partir dos pontos amostrais. De acordo com Haneklaus et al. (2000) a interpolação dos pontos de produtividade numa grade regular é necessária para que diferentes informações possam ser comparadas. Em Agricultura de Precisão, alguns dos métodos mais utilizados para interpolação de dados são os métodos do inverso da distância, vizinho mais próximo e krigagem. No método do inverso da distância, os pontos mais longínquos têm menor efeito no valor derivado para o nó da grade, que é proporcional ao inverso da distância entre cada ponto e o nó da grade. De acordo com Wollenhaupt (1997) este método não requer modelagem ou estimativas, e é computacionalmente rápido. O método do vizinho mais próximo é bastante utilizado para conversão de um conjunto de dados obtidos em espaços regulares em uma grade, ou quando os dados obtidos em campo se encaixam numa grade, com poucos vazios ou pontos perdidos. Quando usado com um pequeno raio de atuação e busca, este método representa muito bem os dados brutos; no entanto, quanto maior o raio, maior a deterioração do coeficiente de determinação e pode chegar ao extremo de não ocorrer interpolação entre pontos (Isaaks, 1989). Ainda segundo o

mesmo autor, a krigagem é um dos mais flexíveis e complexos métodos de interpolação e extrapolação de valores para quase todos os tipos de dados, porém com grande quantidade de dados pode tornar-se bastante demorado; sua utilização requer o tratamento preliminar dos dados através da geração de um modelo de variabilidade espacial, denominado semivariograma.

O objetivo deste trabalho é desenvolver uma metodologia para identificar e caracterizar erros em mapas de produtividade, e aplicá-la para dados de seis diferentes monitores de produtividade comerciais.

## MATERIAL E MÉTODOS

Os dados de produtividade utilizados para a execução deste estudo provêm de seis monitores disponíveis no mercado: RDS® Ceres 2, RDS® Pro Series 8000, AFS®, FieldStar®, GreenStar® e New Holland®, trabalhando em lavouras com área de 12,6 a 42,2 ha. A Tabela 1 sintetiza as características dos dados obtidos em cada uma das áreas. A densidade de coleta de dados foi diferente para cada monitor, dependendo da configuração do programa interno de coleta, variando de 214 a 1674 pontos ha<sup>-1</sup>. Cada monitor de produtividade tem suas próprias características construtivas (Tabela 2), tanto com relação aos sensores envolvidos como nas diversas formas de gravação dos dados, em código aberto ou fechado, ou em arquivo texto. Para os arquivos em código fechado há a necessidade de conversão dos arquivos para texto, através do programa específico que acompanha cada um dos monitores.

Tabela 1. Síntese dos dados obtidos com os vários monitores e referentes a cada uma das áreas

Monitores	Culturas	Áreas (ha)	Número de Pontos Coletados	Densidades de Pontos (pontos ha <sup>-1</sup> )
RDS® Ceres 2	Milho	22,0	12022	546
RDS® Pro Series 8000	Soja	22,0	9955	452
GreenStar®	Milho	17,7	29631	1674
FieldStar®	Soja	78,0	19309	247
AFS®	Soja	42,2	9047	214
New Holland®	Trigo	12,6	8356	663

O sistema FieldStar® não possuía sensor de umidade e, portanto, tal informação não era gravada no arquivo de produtividade. O arquivo gerado pelo sistema também apresentava a largura da plataforma utilizada para a geração da informação. Os monitores AFS® e New Holland® são, na sua essência, os mesmos produzidos pela empresa AgLeader, sendo que o monitor é personalizado para o fabricante de máquinas que o utiliza em seus produtos. As diferenças de comportamento entre ambos são função da dinâmica interativa entre o monitor e a colhedora e dos fatores de configuração do monitor. O monitor GreenStar® é o único em que os dados são gravados em código fechado e que devem ser exportados pelo programa que acompanha o sistema. Os demais arquivos gerados em códigos podem ser abertos por programas especializados, além dos programas que acompanham os monitores.

A metodologia desenvolvida considera erros que podem ser identificados no mapa de produtividade, sendo a ocorrência

Tabela 2. Resumo das características técnicas dos monitores de produtividade comerciais utilizados no estudo

Monitor	Fabricante	Colhedora (marca/modelo)	Sensor de Fluxo de Grãos		Tipo de Arquivo
			Princípio de Leitura	Tipo	
RDS <sup>®</sup> Ceres 2	RDS Technology	NH TC 57	Volumétrico	Ótico	Texto
RDS <sup>®</sup> Pro Series 8000	RDS Technology	NH TC 57	Volumétrico	Ótico	Texto
GreenStar <sup>®</sup>	John Deere	JD 9510	Deteção de massa	Placa de impacto	Código fechado
FieldStar <sup>®</sup>	AGCO	MF 34	Deteção de massa	Radiométrico	Código aberto
AFS <sup>®</sup>	AgLeader <sup>®</sup>	Case 2166	Deteção de massa	Placa de impacto	Código aberto
New Holland <sup>®</sup>	AgLeader <sup>®</sup>	NH TC57	Deteção de massa	Placa de impacto	Código aberto

de tais pontos relatada em forma de porcentagem sobre o total de pontos do arquivo, após a sua produção. Para cada conjunto, o primeiro passo foi uma análise visual do mapa de produtividade num Sistema de Informações Geográficas (SIG) em busca de pontos localizados fora dos talhões (erros grosseiros de posicionamento); em seguida, os dados foram analisados para se identificar produtividade nula. Também foram identificados erros de leitura do sensor de umidade, representados por valor nulo de umidade ou mínimo aceito pelo sistema, que representa ausência de sinal proveniente do sensor.

Erros de largura de corte na plataforma da colhedora devido à interpretação errônea por parte do operador podem ser identificados no arquivo proveniente dos monitores. Tais erros se manifestam, possivelmente, quando a largura parcial efetiva da plataforma é diferente da parcial informada ao sistema e gravada no arquivo. Para sua identificação, buscaram-se pontos no arquivo cuja largura de plataforma era diferente daquela informada como teórica.

As coordenadas geográficas dos pontos para cada arquivo de dados foram convertidas para unidade métrica na projeção Universal Transversa de Mercator (UTM) e, através do teorema de Pitágoras, foi calculada a distância entre pontos consecutivos. Com base na distância, foi possível identificar pontos coletados no mesmo lugar, com distância nula. A ocorrência de distância nula foi considerada erro devido à própria natureza destrutiva da coleta de dados de produtividade. Com base nas coordenadas métricas de cada ponto, foi proposto o índice de deslocamento nas direções norte e leste (ID(N) ID(L)) (Eqs. 1 e 2), que informam o sentido e a direção de deslocamento da colhedora. A Tabela 3 mostra a interpretação do índice com relação à direção e sentido de deslocamento da colhedora.

$$ID(N) = X(i) - X(i + 1) \quad (1)$$

$$ID(L) = Y(j) - Y(j + 1) \quad (2)$$

em que:

- ID(N) - índice de deslocamento na direção Norte-Sul
- ID(L) - índice de deslocamento na direção Leste-Oeste
- X - coordenada Norte da colhedora
- i - ponto na direção N-S e sentido considerado
- Y - coordenada Leste da colhedora
- j - ponto na direção L-O e sentido considerado

O sinal do índice de deslocamento foi utilizado como indicador de início e fim de passada de colheita; sempre que ocorria mudança de sinal, nova passada era iniciada. Tendo-se conhecimento dos inícios e fins de passada, procedeu-se à

Tabela 3. Interpretação do índice de deslocamento da colhedora (ID)

	Valor	Direção	Sentido
ID(N)	Negativo	N-S	Sul->Norte
	Positivo	N-S	Norte->Sul
ID(L)	Positivo	L-O	Leste->Oeste
	Negativo	L-O	Oeste->Leste

caracterização dos erros provenientes da coleta de dados durante o tempo em que a colhedora iniciou ou concluiu uma passada, ou seja, intervalo de enchimento e de esvaziamento da colhedora. Para a caracterização do intervalo de enchimento, foram escolhidas cinco passadas de colheita aleatórias em cada área e as informações de cada ponto foram colocadas de forma seqüencial, sendo que a distância acumulada foi calculada com base na distância entre pontos. Calculou-se, também, a média da produtividade das passadas escolhidas e os pontos foram localizados em gráficos de dispersão da produtividade, em função da distância acumulada a partir do início da passada. O intervalo de enchimento foi considerado aquele referente à distância necessária para que o valor da produtividade atingisse 90% da produtividade máxima local, na média de 5 passadas. As passadas utilizadas para a determinação do intervalo de enchimento foram marcadas no mapa, para análise posterior.

Procedeu-se, então, a interpolação dos dados brutos, com o intuito de se verificar o efeito dos erros no mapa interpolado. O interpolador utilizado foi o método do inverso da distância ao quadrado, com tamanho de célula de 5 m. Com base nas passadas marcadas para determinação das características do intervalo de enchimento, os valores das quadrículas interpoladas correspondentes à mesma posição foram resgatados e comparados com os valores dos pontos e, também, com a média do talhão. As informações dos valores dos pontos, das quadrículas e da média do talhão, foram dispostas num gráfico de dispersão com a distância a partir da borda do talhão no eixo X e produtividade no eixo Y.

Enfim, conduziu-se análise estatística exploratória dos dados, com cálculo da média, mediana, quartil superior e inferior e variância. Com base nos quartis superior e inferior calcularam-se os limites superior e inferior para os valores de produtividade, de acordo com metodologia sugerida por Tukey (1977) para a determinação de limites para dados discrepantes. As Eqs. 3 e 4 mostram a forma de cálculo dos limites superior e inferior.

$$LS = QS + 1,5 AI \quad (3)$$

$$LI = QI - 1,5 AI \quad (4)$$

em que:

- LS - limite superior
- LI - limite inferior
- QS - quartil superior
- QI - quartil inferior
- AI - amplitude interquartilica

Foram construídos histogramas de distribuição da frequência de ocorrência de produtividade. Os limites estatísticos superior e inferior foram analisados juntamente com a distribuição de frequência, para aceitação de tais limites na caracterização de dados discrepantes no conjunto de dados.

## RESULTADOS E DISCUSSÃO

A Tabela 4 resume as características dos erros de posicionamento, valores nulos de produtividade, largura de plataforma de corte diferente da total, distância nula entre pontos e valor nulo para a umidade dos grãos.

A identificação de erros do sensor de umidade pode ser falha, pois identifica apenas casos em que não houve leitura de umidade mas, sim, do sensor de fluxo, tendo como resultado um ponto com umidade nula no arquivo. Não são identificados erros sistemáticos ou aleatórios na leitura do sensor, mas apenas falhas na leitura. No arquivo de dados obtido com o monitor de produtividade RDS® Ceres 2, a constatação de valores nulos de umidade em 9,50% dos pontos, está associada à ocorrência de valores de produtividade de 9999,9 t ha<sup>-1</sup>, que, por sua vez, está associada às leituras no sensor de fluxo com a plataforma da colhedora levantada, situação que poderia ser interpretada como produção em área nula. A produtividade é o resultado da derivação da produção em relação à área e, sendo esta nula, a produtividade tende ao infinito, resultado interpretado pelo monitor como 9999,99 t ha<sup>-1</sup>. Pode haver outros motivos para a ocorrência de tais dados, visto que o algoritmo de leitura do monitor é inacessível.

Nos dados obtidos com o monitor RDS® Pro Series 8000, a ocorrência de produtividades nulas está associada, de certa forma, à ocorrência de umidade nula, mas a totalidade dos pontos que possuem largura de plataforma nula possui também produtividade com mesmo valor. Nos dados obtidos com o monitor FieldStar®, a análise inicial identificou pequena quantidade de erros, com apenas algumas ocorrências de distâncias nulas entre pontos.

A existência de distância nula entre pontos nos dados obtidos pelo monitor GreenStar® está associada também à produtividade nula, ou simplesmente ausente, e a valores nulos de umidade. Parte dos pontos que apresentaram distância nula

entre si possuía também produtividades improváveis e muito elevadas. Os erros apresentados pelos dados obtidos com o monitor New Holland® mostraram-se independentes, sendo que o erro de maior ocorrência foi o de posicionamento, representado por um grupo de pontos localizados fora do talhão. A mesma ausência de relação entre os erros é mostrada pelos dados obtidos com o monitor AFS®. Os erros devidos à imprecisão do GPS não são identificáveis por esta metodologia.

Especificamente com relação ao monitor GreenStar®, a alta frequência de ocorrência de pontos com distância nula é devida, provavelmente, à alta taxa de coleta de pontos, associada a pequenos erros de posicionamento e, também, à baixa velocidade da operação de colheita. Muitos pontos são coletados num intervalo curto de tempo, provavelmente insuficiente para que o GPS detecte a mudança de posição da colhedora e informe ao sistema a nova posição diferente da anterior. Por outro lado, como a velocidade da colheita é baixa, a distância percorrida pela colhedora dentro do intervalo de coleta de dados também é pequena e está dentro do círculo de erro provável da maioria dos receptores de GPS utilizados em agricultura de precisão, possibilitando que a nova posição seja erroneamente equivalente à anterior.

O resultado da aplicação do índice de deslocamento em um dos arquivos de produtividade pode ser visto na Figura 1. Tal índice permite a identificação exata do percurso da colhedora, o início e final de cada passada, porém pode ser aplicado apenas

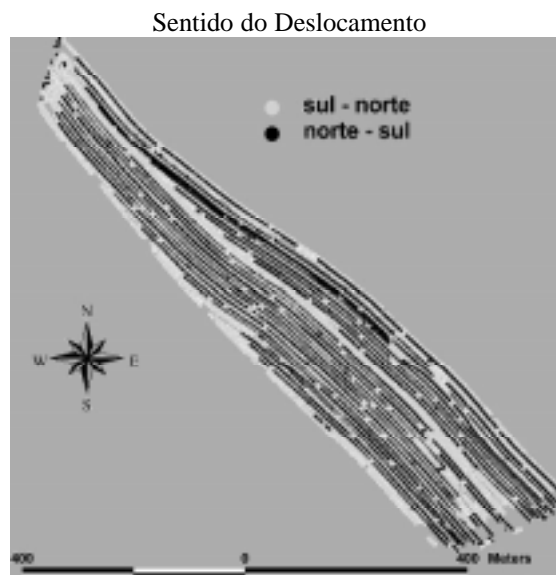


Figura 1. Resultado da aplicação do método de determinação do sentido de deslocamento nos dados obtidos com o monitor RDS® Ceres 2

Tabela 4. Frequência de ocorrência de erros encontrados nos arquivos de produtividade de acordo com a metodologia utilizada

Tipos de Erro	Monitor em que os Arquivos Foram Gerados					
	RDS® Ceres 2	RDS® Pro Series 8000	Field Star®	Green Star®	New Holland®	AFS®
Frequência de Erro (%)						
Erros de posicionamento	0,00	0,07	0,00	0,00	6,30	0,00
Produtividade nula	0,33	4,60	0,00	2,34	1,83	0,34
Largura de plataforma	4,10	10,80	ND*	5,60	4,30	1,29
Distância nula entre pontos	0,00	0,00	0,08	11,48	0,07	0,00
Umidade nula	9,50	9,30	ND*	3,80	1,40	1,06

\* ND - não disponível no arquivo

quando há uma única direção predominante de colheita, seja Norte-Sul, Leste-Oeste ou direções intermediárias. Quando as passadas de colheita possuem mais de um sentido como, por exemplo, em curvas acentuadas, a metodologia se mostra limitada na identificação do sentido de deslocamento.

A Figura 2 mostra a produtividade das cinco passadas da colhedora, escolhidas aleatoriamente dentro dos dados obtidos com cada um dos monitores, em função da distância acumulada a partir da borda do talhão. Também apresenta a produtividade média das cinco passadas, em função da distância acumulada, a partir da qual foi determinado o intervalo de enchimento.

Os dados obtidos com o monitor RDS® Ceres 2 apresentaram intervalo de enchimento correspondente a 19,58 m, enquanto os dados obtidos com o monitor RDS® Pro Series 8000 apresentaram intervalo de enchimento de 44,00 m, sendo que foram necessários 20 m de percurso para que os valores lidos pelo monitor apresentassem 20% da produtividade máxima local. Embora ambos utilizem o mesmo sensor de fluxo de grãos, o comportamento das linhas de produtividade em função da distância acumulada é bastante diferenciado, com tendência suave no caso dos dados obtidos com o monitor RDS® Ceres 2 e com tendência oscilante para os dados obtidos com o monitor RDS® Pro Series 8000. Tais diferenças são função do programa interno de coleta de dados, que insere maior suavização nos dados do primeiro com relação ao segundo. Os dados obtidos com o monitor GreenStar® apresentaram intervalo de enchimento de 14,11 m, com a particularidade da alta frequência de coleta de pontos. Para os dados obtidos com o monitor New Holland®, o intervalo de enchimento foi equivalente a 13,53 m, de acordo com o critério adotado. Porém, considerando-se que com 89% da produtividade máxima (ao invés dos 90% propostos) o monitor está sensível às variações naturais da cultura, o intervalo de enchimento cai para 6,20 m. Os dados obtidos com o monitor AFS® não apresentaram erros devido ao intervalo de enchimento da colhedora e os dados obtidos com o monitor FieldStar® apresentaram distância de 17,52 m referente ao intervalo de enchimento, apesar de não ter seu início com produtividade nula. Para os dados coletados com esse monitor, as passadas da colhedora analisadas apresentaram pontos bastante uniformes, com grande suavização dentro de uma mesma passada.

O efeito do intervalo de enchimento obtido no mapa após a interpolação, pode ser observado na Figura 3. De maneira geral, todos os monitores que apresentaram erros de intervalo de enchimento, o transferiram para o mapa de dados interpolados. Por outro lado, o interpolador produziu um efeito suavizador nos dados, aproximando a região sob efeito do intervalo de enchimento, da média do talhão.

O interpolador produziu efeitos diferentes nos dados dos monitores, aproximando mais os valores em relação à média para os dados obtidos com o monitor RDS® Ceres 2, em relação aos outros. O monitor AFS®, por não apresentar erros de intervalo de enchimento, apresentou resultados muito semelhantes para os dados interpolados e para a média das passadas. Mesmo com intervalo de enchimento de 44 m, o interpolador promoveu grande suavização nos dados obtidos pelo monitor RDS® Pro Series 8000, diminuindo a diferença existente entre a média do talhão e a média das passadas (Figura 3).

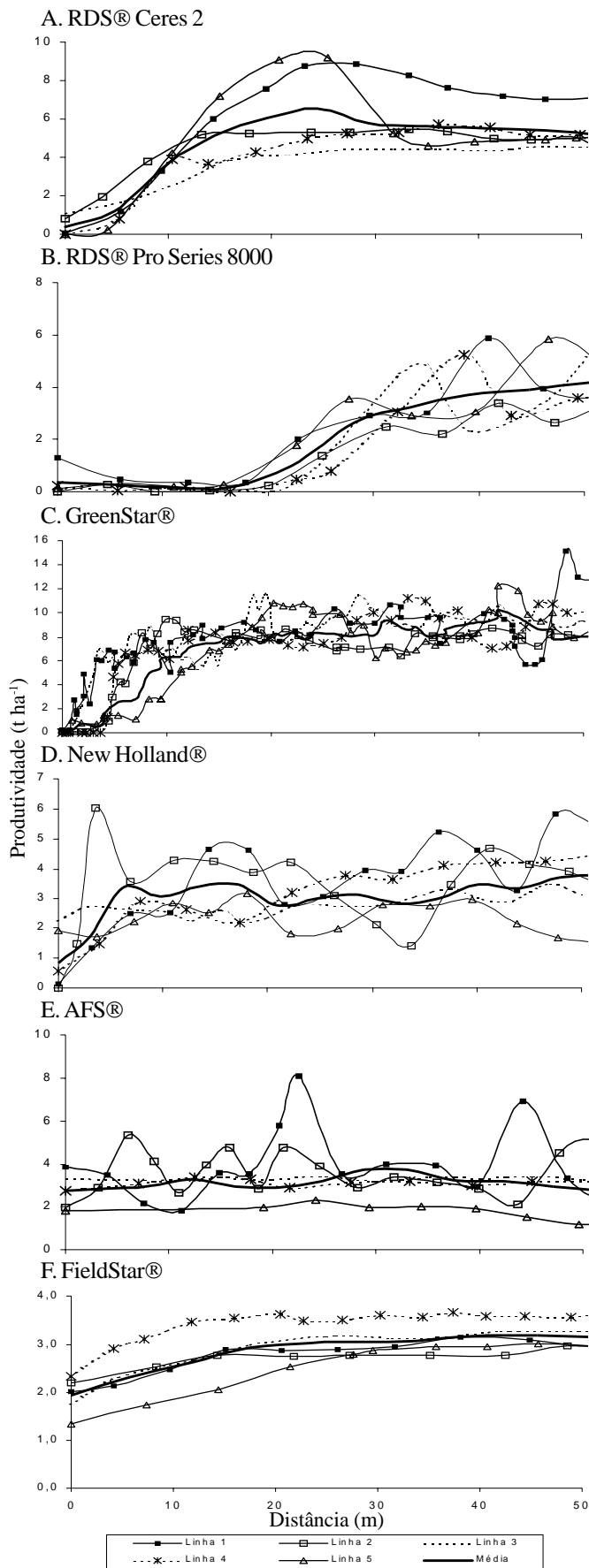


Figura 2. Efeito do intervalo de enchimento na coleta de dados de produtividade para cada arquivo analisado proveniente dos diferentes monitores

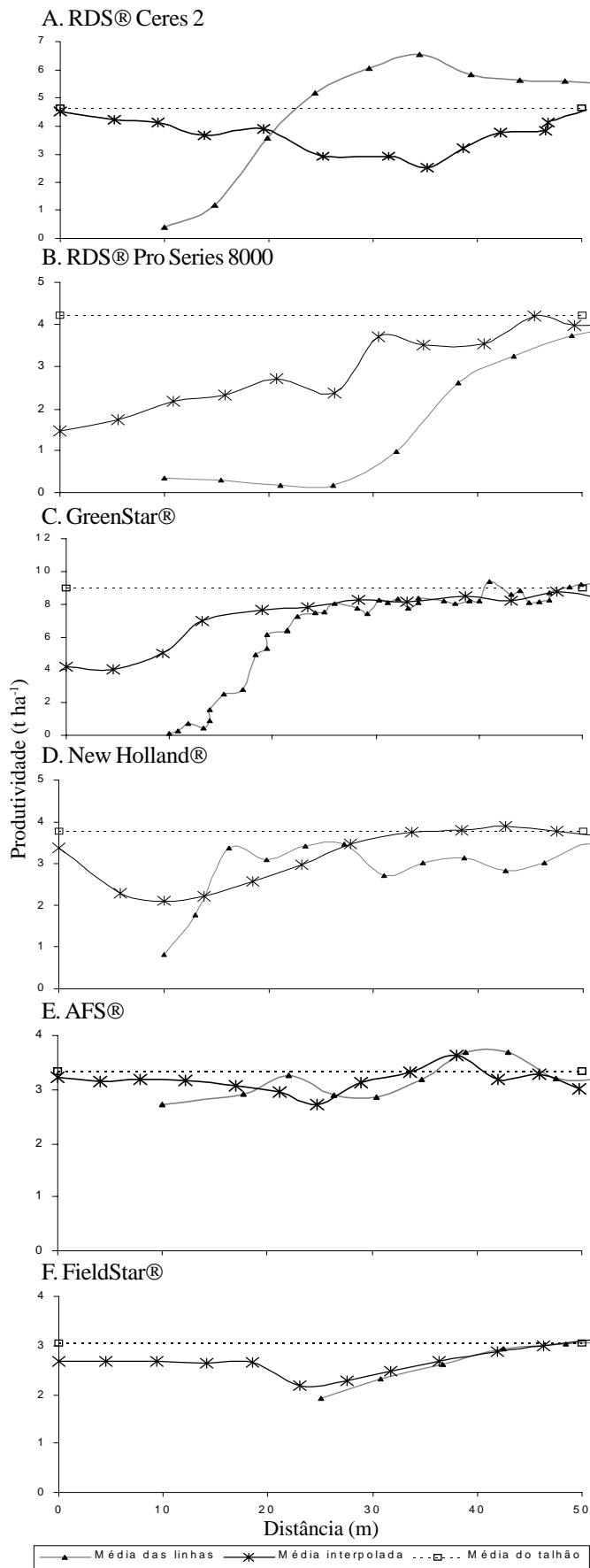


Figura 3. Efeito do intervalo de enchimento no mapa de dados interpolados para os arquivos analisados obtidos pelos diferentes monitores de produtividade

Não há nada que assegure a normalidade da distribuição de frequência dos mapas de produtividade e, tampouco, deve ser assim. O histograma de distribuição pode mostrar informações importantes no entendimento do mapa de produtividade. O histograma apresentado na Figura 4, que representa os dados obtidos com o monitor AFS®, mostra claramente dois padrões de distribuição, com alta ocorrência de alta produtividade para soja (entre 3,5 e 4,0 t ha<sup>-1</sup>) e uma segunda parte do talhão, com produtividade menor e alta frequência de ocorrência (2,5 t ha<sup>-1</sup>). Valores nulos de produtividade certamente são erros de medição.

Os pontos com baixa produtividade e baixa ocorrência contêm, provavelmente, os erros de intervalo de enchimento, produtividade nula, erros de medição do sensor de umidade, mas também podem conter pontos que representam fielmente a produtividade da cultura. Visualmente, pode-se verificar nos histogramas da Figura 4 onde ocorre a mudança de comportamento da frequência de ocorrência dos pontos de produtividade, mas a aplicação da metodologia proposta por Tukey (1977) forneceu base para a classificação dos dados supostamente errôneos como discrepantes. A Tabela 5 mostra a análise estatística descritiva e a análise de dados discrepantes.

Pela análise de dados discrepantes, o limite inferior estatístico para os dados obtidos com o monitor RDS® Ceres 2 foi calculado em 3,03 t ha<sup>-1</sup> e coincide, no histograma, com a região onde ocorre pequena ocorrência de pontos de produtividade, resultando em 406 pontos abaixo do limite inferior. O limite superior foi de 6,51 t ha<sup>-1</sup>, resultando em 1138 pontos considerados discrepantes, coincidindo com a diminuição na ocorrência de pontos de produtividade. Observa-se grande discrepância entre o valor da média e da mediana, em função da alta frequência de ocorrência de valores 9999,9 t ha<sup>-1</sup>; já para os dados obtidos com o monitor RDS® Pro Series 8000, o limite inferior foi negativo, resultando em que nenhum ponto foi selecionado abaixo desse limite. Este resultado foi obtido devido à alta ocorrência de valores nulos de produtividade. Apesar da mediana ser uma medida de posição mais robusta que a média, alta frequência de ocorrência pode afetar os resultados (Tukey, 1977). No processo de filtragem, por ocasião do cálculo dos limites, os dados já terão passado por filtros que eliminarão a interferência de produtividade nula no cálculo dos limites. Com relação ao limite superior, de 8,36 t ha<sup>-1</sup>, há 286 pontos maiores, que foram considerados discrepantes.

A média e a mediana, para os dados obtidos com o monitor GreenStar®, também apresentaram distanciamento, sendo a média de 9,00 e a mediana de 7,52 t ha<sup>-1</sup>. O limite inferior para valores discrepantes foi de 2,43 t ha<sup>-1</sup>, ocorrendo 2216 pontos menores que o limite estabelecido. O limite inferior calculado coincide com a alteração no padrão da distribuição. Abaixo deste limite, estão todos os pontos cujos valores de produtividade são nulos e, também, os pontos cuja umidade é nula. O limite superior para os dados foi de 12,36 t ha<sup>-1</sup>, com a ocorrência de 581 pontos acima do limite e foram considerados

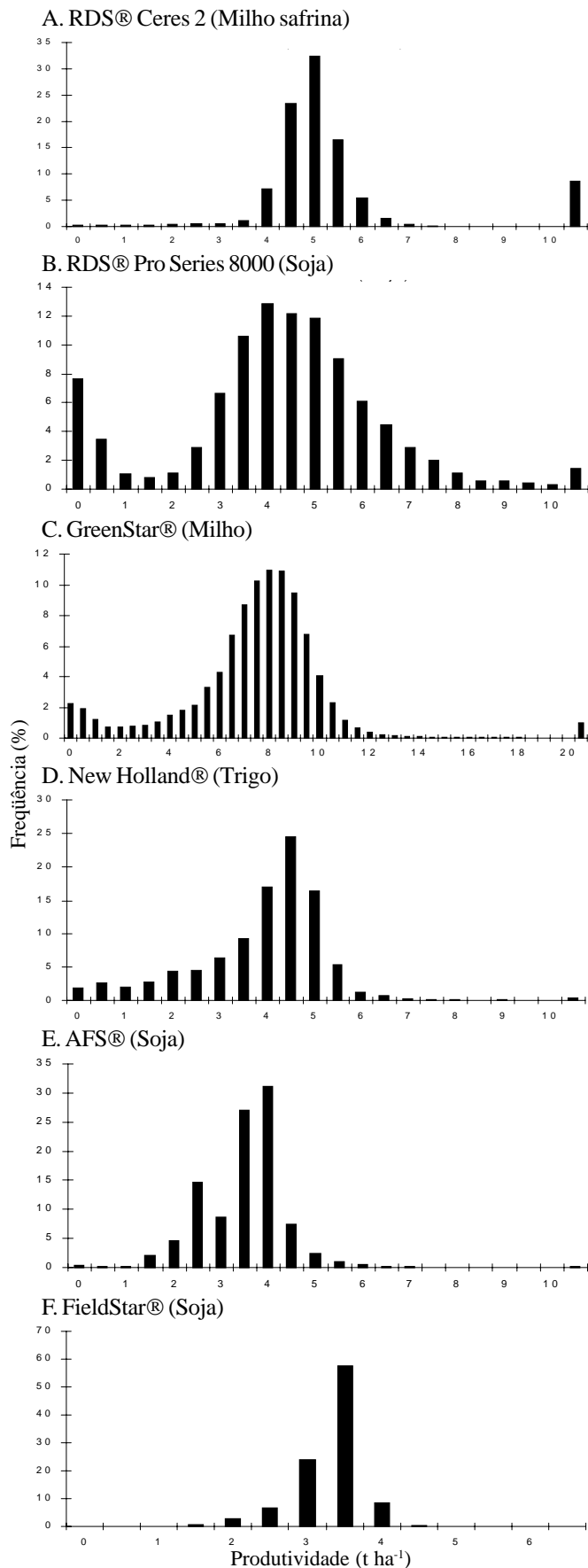


Figura 4. Histogramas de distribuição da frequência de ocorrência de produtividade para os arquivos analisados obtidos pelos diferentes monitores de produtividade

discrepantes. Os dados obtidos com o monitor New Holland® apresentaram relativa simetria, com média bastante próxima à mediana. O limite superior para classificação de dados como discrepantes foi calculado com produtividade igual a 6,67 t ha<sup>-1</sup>, havendo ocorrência de 94 pontos com valores superiores ao limite. No outro extremo da distribuição, há 491 pontos menores que o limite inferior, de 0,87 t ha<sup>-1</sup>. A localização dos limites superior e inferior no gráfico de distribuição de frequência de produtividades coincide com os pontos onde visualmente se pode verificar alteração no padrão de distribuição, como se constata com o limite inferior, que considera discrepantes todos os dados referentes à distribuição nas classes de 0 a 1 t ha<sup>-1</sup>. No entanto, podem ocorrer pontos reais com produtividade baixa que estarão sendo classificados como discrepantes por esta análise. Por outro lado, pode-se considerar que a quantidade de pontos reais classificados como discrepantes será pequena, devido à própria metodologia utilizada.

Com média de 3,30 e mediana de 3,40 t ha<sup>-1</sup>, os dados obtidos com o monitor AFS® apresentaram seus limites superior e inferior calculados em 5,25 e 1,21 t ha<sup>-1</sup> (Tabela 5), resultando em 118 pontos com valores maiores que o limite superior e 98 pontos menores que o limite inferior.

Os dados obtidos com o monitor FieldStar® apresentaram produtividade mínima de 1,30 t ha<sup>-1</sup>, indicando que há, provavelmente, no programa interno de coleta de dados, um filtro que impede a coleta de dados nulos. Os valores da média e mediana estão muito próximos (3,10 e 3,13 t ha<sup>-1</sup> respectivamente) indicando simetria nos dados, confirmada pelo coeficiente de assimetria. O limite inferior calculado para os dados foi de 2,29 t ha<sup>-1</sup> e o limite superior foi de 3,93 t ha<sup>-1</sup>, sendo que a amplitude do limite é de 1,64 t ha<sup>-1</sup>. A pequena amplitude do limite pode ser atribuída à grande suavização promovida pelo programa interno de coleta de dados do monitor ou à pequena variabilidade espacial da produtividade nos dados analisados. Tanto uma como outra hipótese leva à concentração de valores em torno da média, que resultou em amplitude interquartilica de 0,41 t ha<sup>-1</sup>.

Foram encontrados erros ou pontos discrepantes em todos os conjuntos de dados analisados. Os dados obtidos com os monitores RDS® Ceres 2, RDS® Pro Series 8000 e GreenStar® apresentaram maiores quantidades de pontos errôneos, enquanto os dados obtidos com o monitor AFS® indicaram apenas dados considerados discrepantes. Com relação aos dados obtidos com o monitor FieldStar®, o maior problema encontrado foi o intervalo de enchimento. Todos os erros apresentados pelos dados brutos são função da relação dinâmica entre o monitor e a máquina e, também, de parâmetros do monitor. Espera-se que os erros e suas características sejam constantes enquanto não forem alteradas as relações monitor x máquina e os parâmetros do monitor. Quando da alteração das relações ou parâmetros, as características dos erros provavelmente não serão alteradas, mas se espera que a metodologia seja eficaz em sua caracterização.

Tabela 5. Análise estatística descritiva e de dados discrepantes para os arquivos de dados estudados

Parâmetro	Monitor em que os Arquivos Foram Gerados					
	RDS® Ceres 2	RDS® Pro Series 8000	GreenStar®	New Holland®	AFS®	Field Star®
	Cultura					
	Milho safrinha	Soja	Milho	Trigo	Soja	Soja
Pontos válidos	12022	9953	28712	8356	9047	20821
Média (t ha <sup>-1</sup> )	852,7	4,2	9,0	3,7	3,3	3,1
Mediana (t. ha <sup>-1</sup> )	4,73	4,11	7,52	3,99	3,40	3,13
Mínimo (t ha <sup>-1</sup> )	0	0	0	0	0	1,3
Maximo (t ha <sup>-1</sup> )	9999,9	174,3	2162,6	147,9	485,2	4,4
Quartil inferior (t ha <sup>-1</sup> )	4,33	3,07	6,15	3,05	2,72	2,90
Quartil superior (t ha <sup>-1</sup> )	5,20	5,19	8,64	4,50	3,73	3,31
Amplitude interquartilica	0,87	2,11	2,48	1,45	1,01	0,41
Variância	7757874,94	11,84	1398,09	7,29	29,90	0,18
Desvio-Padrão (t ha <sup>-1</sup> )	2785,30	3,44	37,39	2,70	5,47	0,43
Límite Inferior (t ha <sup>-1</sup> )	3,03	-0,10	2,43	0,87	5,25	2,29
Límite Superior (t ha <sup>-1</sup> )	6,51	8,36	12,36	6,67	1,21	3,93
Valores > LS*	1138	286	581	94	118	100
Valores < LI**	406	0	2216	491	98	1337

\* LS é o limite superior

\*\* LI é o limite inferior

## CONCLUSÕES

1. A classificação dos erros possíveis de serem identificados em pós-processamento dos dados levanta aspectos importantes da caracterização dos erros, fornecendo bases para que, por processo de filtragem, os pontos considerados errôneos possam ser removidos para elevar o nível qualitativo da informação contida em mapas de produtividade, qualquer que seja o monitor utilizado para a coleta de dados.

2. A metodologia desenvolvida para identificação de início e fim de passada de colheita é bastante confiável, permitindo a identificação de dados inconsistentes referentes ao tempo de enchimento e de esvaziamento da colhedora nos inícios e finais de passadas.

## AGRADECIMENTOS

Os autores agradecem à FAPESP pelo apoio financeiro e à Fundação ABC, AGCO, Case, John Deere e New Holland, pela colaboração na obtenção dos dados.

## LITERATURA CITADA

- Blackmore, B.S.; Marshall, C.J. Yield mapping; errors and algorithms. In: International Conference On Precision Agriculture, 3, Minneapolis, 1996. Proceedings, Madison, WI, 1996, p.403-415.
- Cambardella, C.A.; Karlen, D.L. Spatial analysis of soil fertility parameters. Precision Agriculture, Warwick, v.1, n.1, p. 5-14, 1999.
- Haneklaus, S.; Lilienthal, H.; Schnug, E.; Panten, K.; Haverech, E. Routines for efficient yield mapping. In: International Conference on Precision Agriculture, 5, Madison, Wisconsin, 2000. Proceedings, Madison, 2000. CD-Rom
- Isaaks, E.H.; Srivastava; R.M. Applied geostatistics. 1 ed. Oxford: Oxford University Press, 1989. 561p.
- Lark, R.M.; Stafford, J.V. Exploratory analysis of yield maps of combine crops, In: European Conference On Precision Agriculture, 1, 1997. Precision Agriculture 97, v.1, Warwick, 1997, p.887-894.
- Makepeace, R.J. Benefits and limitations of precision farming, In: Proceedings of the Brighton conference on pests and diseases, v.3, p.1235-1242, 1996.
- Missotten, B.; Strubbe, D. De Baerdemaeker, J. Accuracies of grain and straw yield maps. Agricultural Engineering, St. Joseph, v.9, p.23-26, 1996.
- Molin, J.P. Geração interpretação de mapas de produtividade para agricultura de precisão, In: Borém, A.; Giúdice, M.P. del; Quiroz, D.M. de; Mantovani, E.C.; Ferreira, L.R.; Valle, F.X.R. do; Gomide, R.L. Agricultura de precisão. Viçosa: Universidade Federal de Viçosa, p.237-258, 2000.
- Moore, M. An investigation into the accuracy of yield maps and their subsequent use in crop management, Cranfield: Cranfield University, 1998, 379p. PhD Thesis
- Morgan, M.; Ess, D. The Precision farming guide for agriculturists. Moline USA: Deere & Company, 1997, 117p.
- Thylén, L.; Jurschik, P.; Murphy, D.L.P. Improving the quality of yield data, In: European Conference on Precision Agriculture, 1, 1997. Precision Agriculture '97, v.1, Warwick, 1997, p.743-750.
- Tukey, J.W. Exploratory data analysis. 1 ed. Reading, Massachusetts, 1977. v.1/3.
- Wollenhaupt, N.C.; Mulla, D.J.; Crawford, C.A.G. Soil sampling and interpolation techniques for mapping spatial variability of soil properties, In: The state of site specific management for agriculture, Madison, Wisconsin, p.19-53, 1997.