

DESENVOLVIMENTO DE UM ALGORITMO PARA REDUÇÃO DE ERROS EM MAPAS DE PRODUTIVIDADE OBTIDOS EM AGRICULTURA DE PRECISÃO

¹L.M.Gimenez, ²J.P. Molin (orientador): Departamento de Engenharia Rural – ESALQ/USP

Os mapas de produtividade gerados durante a colheita são constituídos por um grande número de pontos georreferenciados sendo expressos nestes pontos os valores de produtividade da área que representam e podem ser interpolados para a geração de uma superfície. Nem sempre estes pontos refletem a produtividade de forma correta, em função de erros que ocorrem durante o processo de coleta dos dados. O número e os tipos de erros variam em função de uma série de fatores como o tipo de monitor de colheita, colhedora, tamanho e formato dos talhões, qualidade do sinal DGPS e outros. A ocorrência de erros nos dados utilizados na geração dos mapas levam a interpretações errôneas da variabilidade presente no campo, reduzindo a qualidade do mapa gerado. A eliminação dos erros pode ser realizada através da utilização de planilha eletrônica. Nesse trabalho foram propostas e utilizadas funções de lógica na composição de um algoritmo para realizar a eliminação dos erros presentes nos dados coletados, detectando e eliminando os erros mais importantes e mais freqüentes. O algoritmo criado foi testado em dados gerados durante duas colheitas, tendo ocorrido substancial redução no volume de erros e elevação da qualidade dos mapas de produtividade.

¹Bolsista de Iniciação Científica do CNPq, ²Prof. Dr. do Departamento de Engenharia Rural, ESALQ/USP

PALAVRAS CHAVE: Agricultura de Precisão, Mapa de Produtividade, algoritmo

INTRODUÇÃO

A geração de mapas de rendimento ou produtividade já é vista como prática bastante comum em países onde a tecnologia na agricultura está mais evoluída, onde o monitoramento da produtividade é considerado o primeiro passo na implantação da Agricultura de Precisão. No Brasil em algumas áreas onde a nova tecnologia vem sendo estudada já se tem mapas de produtividade há alguns anos.

Os mapas gerados durante a colheita, pela própria colhedora equipada com sensores e um sistema para posicionamento em campo (GPS), são constituídos de um grande número de pontos onde se tem a produtividade georeferenciada. Nem sempre estes pontos refletem a produtividade adequadamente sendo comum a ocorrência de erros na coleta dos dados.

O número e os tipos de erros variam em função de uma série de fatores como: monitor de colheita, o sistema utilizado para o

posicionamento, a colhedora, o tamanho e o formato dos talhões entre outros.

A presença de erros nos dados utilizados para a confecção dos mapas de produtividade é bastante prejudicial à qualidade do mapa gerado podendo inclusive levar a interpretações errôneas mesmo que seu volume em relação ao total de dados coletados seja pequeno. É bastante importante portanto remover os erros dos mapas.

Este trabalho tem por objetivo gerar um filtro para reduzir os erros em mapas de produtividade a partir da eliminação dos dados discrepantes dentre aqueles coletados durante a colheita.

MATERIAL E MÉTODOS

Foram utilizados dados coletados durante a colheita de duas áreas com diferentes culturas, colhedoras e monitores de colheita.

Os monitores de colheita, juntamente com o sistema de posicionamento, geram um arquivo texto contendo os dados de posicionamento (latitude e longitude), e de quantidade de grão que entra na máquina além de outros fatores como a largura de plataforma, umidade do grão e possíveis fatores encontrados no campo e que sejam georeferenciados durante a colheita pelo operador da máquina.

A figura 1 representa o tipo de informação gerada pelos monitores de colheita, nas primeiras colunas (latitude e longitude) tem-se as coordenadas que permitem o posicionamento no campo, na terceira coluna encontram-se os dados de quantidade de grão que passam pelo sensor no momento em que a máquina se encontra naquela posição, na última coluna encontram-se os dados de umidade do grão.

LONGITUDE	LATITUDE	FLOW	TIME	DISTANCE	SWATH	MOISTURE
.....
-49,933493	-24,857763	349,27	7119068	594	4775	31,40
-49,933490	-24,857411	353,80	7119069	594	4775	31,40
-49,933490	-24,857058	385,55	7119070	594	4775	31,40
-49,933490	-24,856705	385,55	7119071	620	4775	31,40
-49,933487	-24,856352	381,02	7119072	645	4775	31,40
-49,933487	-24,855999	471,74	7119073	645	4775	31,40
-49,933487	-24,855788	430,91	7119074	645	4775	31,40
.....

Figura 1 – Exemplo de um arquivo gerado pelo monitor de colheita.

Estes dados são importados para um SIG (Sistema de Informações Geográficas), voltado à agricultura de precisão onde um algoritmo converte as informações em produtividade no ponto considerado. Na figura 2 temos o mesmo conjunto de dados após o cálculo da produtividade.

LONGITUDE	LATITUDE	FLOW	TIME	DISTANCE	SWATH	MOISTURE	DRY_YIELD
-49,933493	-24,857763	349,27	7119068	594	4775	31,40	11,1978
-49,933490	-24,857411	353,80	7119069	594	4775	31,40	11,3430
-49,933490	-24,857058	385,55	7119070	594	4775	31,40	12,3609
-49,933490	-24,856705	385,55	7119071	620	4775	31,40	11,8426
-49,933487	-24,856352	381,02	7119072	645	4775	31,40	11,2498
-49,933487	-24,855999	471,74	7119073	645	4775	31,40	13,9283
-49,933487	-24,855788	430,91	7119074	645	4775	31,40	12,7228

Figura 2 – Arquivo após o cálculo da produtividade, expressa como “dry yield”.

Após o cálculo, os dados podem ser utilizados para gerar o mapa de produtividade. Este mapa normalmente apresenta erros, uma vez que o SIG somente converte os dados gerados pelo monitor de colheita em campo. Cada linha do arquivo representa um ponto com uma série de atributos, sendo a produtividade um deles. Os erros nos valores de produtividade se expressam como valores muito altos, baixos, nulos ou até mesmo pela falta de valores.

A eliminação dos dados discrepantes foi realizada com a utilização de uma planilha eletrônica e suas funções de lógica, os dados foram importados para a planilha sendo filtrados por meio do algoritmo criado.

O algoritmo trabalha os dados de produtividade um a um comparando-os à média de todos os dados, ao valor de um dado coletado antes do dado avaliado e também à distância em metros entre o ponto estudado e o ponto imediatamente anterior, o valor do ponto estudado é então classificado pelo algoritmo como VERDADEIRO, quando está dentro das exigências impostas pelo algoritmo ou FALSO quando foge às exigências impostas.

Quando o valor do ponto é considerado VERDADEIRO, é atribuído a ele o valor ZERO uma vez que o mesmo deve ser descartado; quando o mesmo é considerado FALSO, seu valor é MANTIDO.

Desta forma, é possível eliminar pontos que estejam com valores 2, 3, 4,,n vezes superiores e inferiores à média de todos os pontos, corrigindo possíveis falhas grosseiras do sensor de produtividade e de largura de plataforma inadequada. Através da comparação do dado avaliado com um dado coletado cerca de 15, 16, 17,....., n pontos atrás e também da distância entre os pontos coletados, é possível eliminar boa parte dos erros presentes nos finais dos talhões, onde são realizadas manobras e também erros no tempo de enchimento da máquina.

A eliminação de pontos com valores maiores ou menores que a média é realizada pela condição:

$$SE(OU(X > n * MÉDIA); (X < MÉDIA / n))$$

Sendo o valor do ponto X avaliado considerado:

VERDADEIRO quando o mesmo é n vezes maior ou menor que a MÉDIA de todos os dados, recebendo o valor zero.

FALSO quando o mesmo se encontra no intervalo compreendido entre 3 vezes o valor da MÉDIA e o valor da MÉDIA dividido por n, sendo mantido então o seu próprio valor.

A eliminação de pontos com valores discrepantes, principalmente devido a problemas com coleta de dados durante as manobras e também com o tempo de enchimento da máquina é realizada com a seguinte condição:

$SE(OU((X_n > (X_{n-15}) * n); (X_n < (X_{n-15}) / n)))$

Sendo o valor do ponto X_n avaliado considerado:

VERDADEIRO quando o mesmo é n vezes maior ou menor que o valor do ponto X_{n-15} , recebendo o valor zero.

FALSO quando X_n considerado se encontra entre $n * (X_{n-15})$ e $(X_{n-15}) / n$, sendo mantido o seu valor.

De modo complementar, a distância entre os pontos coletados pode também ser utilizada para reduzir os erros de manobras e de falhas no sistema de posicionamento com a seguinte condição:

$SE(OU(DISTÂNCIA < Y_m); (DISTÂNCIA > Y_m))$

Onde:

DISTÂNCIA: é o valor calculado, em metros, da distância entre o ponto X_n e o ponto X_{n-1} a partir das coordenadas dos pontos, previamente convertidas para coordenadas UTM.

Y_{min} e $Y_{máx}$: são os valores limites admitidos para a menor e a maior distância entre pontos, de acordo com a colhedora e monitor de colheita utilizados.

Os valores de distância entre a coleta dos pontos são bastante regulares para uma mesma colhedora, colhendo o mesmo tipo de grão o que permite obter para cada colhedora valores bastante seguros de Y_{min} e $Y_{máx}$.

Após a junção das condições acima estabelecidas chega-se ao algoritmo:

$SE(OU(X > n * MÉDIA); (X < MÉDIA / n); (X_n > (X_{n-15}) * n); (X_n < (X_{n-15}) / n)); (DISTÂNCIA < Y_m); (DISTÂNCIA > Y_m))$

Onde o valor do ponto avaliado será:

ZERO, caso o resultado da expressão seja VERDADEIRO
MANTIDO, caso o resultado da expressão seja FALSO

Após passar pelo algoritmo, os dados de coordenadas e valor do ponto considerados pelo algoritmo são copiados para um novo arquivo,

onde são eliminados os pontos com valor igual a ZERO, podendo depois disso serem importados novamente pelo SIG para a geração dos mapas.

RESULTADOS E DISCUSSÃO

Foram estudados dados de duas áreas colhidas com diferentes culturas, máquinas e monitores de colheita. Houve significativa melhora na qualidade dos dados sendo reduzidos a Variância e o Desvio Padrão dos dados nas duas áreas. O número de dados eliminados pelo filtro foi bastante diferente em termos proporcionais, a qualidade da informação pode ter sido em parte prejudicada pela eliminação de um grande número de pontos entretanto a densidade de pontos é bastante elevada mesmo após a remoção a passagem do filtro.

Mapa de Produtividade em Pontos - Soja 1999

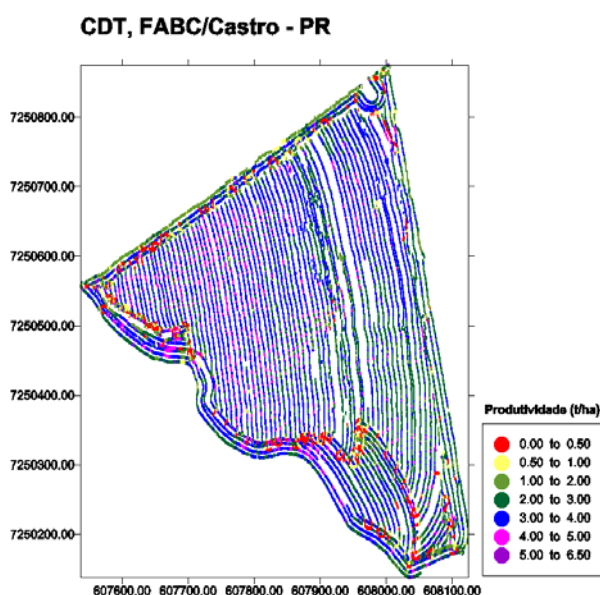


Figura 3 – Mapa gerado com dados antes da aplicação do algoritmo.

Mapa de Produtividade em Pontos - Soja 1999

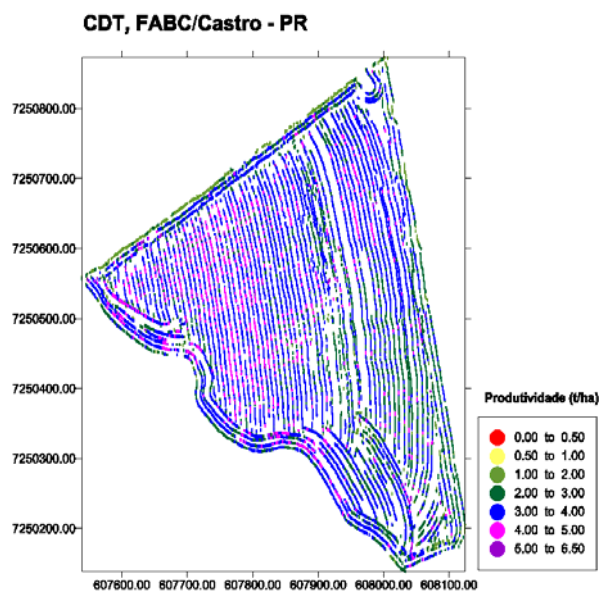


Figura 4 – Mapa gerado com dados após a aplicação do algoritmo.

Os dados são apresentados como uma série de pontos coletados durante a manobra e pelo tempo de enchimento da máquina. Também podem ser considerados erros os pontos em que se passa de uma classe de produtividade para outra distante.

Na figura 2, após a aplicação do filtro pode se notar uma significativa redução nos erros de manobras e tempo de enchimento da máquina, mesmo com uma redução significativa no número de dados, é possível identificar com bastante segurança as zonas de maior e menor produtividade.

Soja 1999 - CDT, FABC/Castro - PR		
	Dados Brutos	Dados Filtrados
Média (t/ha)	2,97	3,17
Variância	1,01	0,41

Desvio Padrão	1,01	0,64
N° de Total de Pontos	27.177	15.811
N° de pontos/ha	1.151	669

Tabela 1 – Comportamento dos dados antes e após a utilização do filtro.

Na tabela 1 é possível perceber que ocorre redução do desvio padrão e da variância, devido à eliminação de pontos com valores discrepantes. A elevação da média se deve à eliminação de um grande número de pontos onde os valores de produtividade são bastante baixos, característica marcante dos erros em manobras e tempo de enchimento.

Mapa de Produtividade em Pontos - Milho 1998

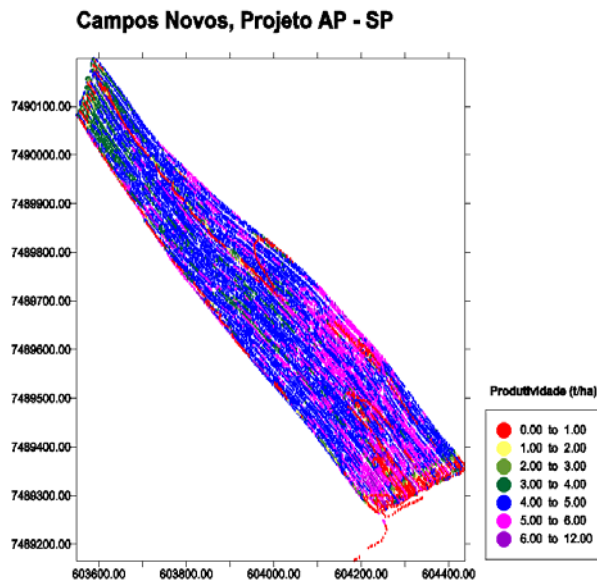


Figura 5 - Mapa gerado com dados antes da aplicação do filtro.

Mapa de Produtividade em Pontos - Milho 1998

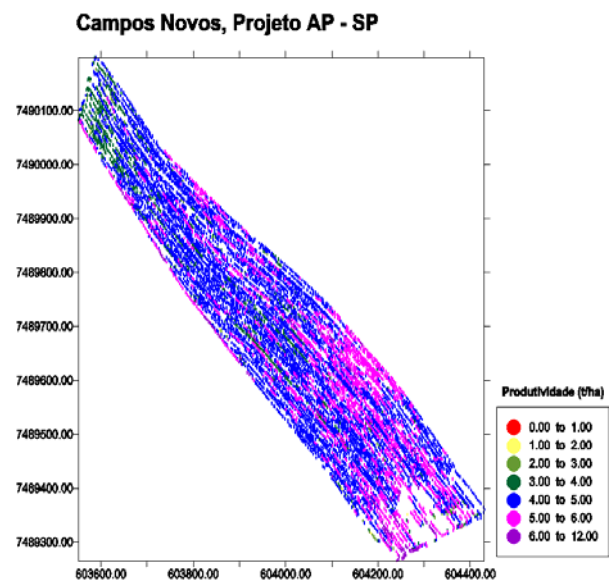


Figura 6 - Mapa gerado com dados após a aplicação do filtro

Na segunda área estudada o comportamento do filtro foi bastante similar havendo entretanto uma menor redução no número de pontos eliminados. Na região inferior da figura é bastante significativo o número de pontos retirados. A observação do mapa produzido com os dados antes da aplicação do filtro permite afirmar, entretanto que os pontos retirados representavam de fato erros, o que pode se notar pela produtividade que variou de 0 – 0,5t/ha a 5 – 6t/ha em um espaço muito reduzido o que não ocorre normalmente. Outro tipo de erro que foi bastante reduzido é o erro de largura de plataforma que é caracterizado pela presença de linhas contínuas com produtividade bastante diferente das linhas próximas.

Milho 1998 - Campos Novos - SP		
	Dados Brutos	Dados Filtrados
Média (t/ha)	4,25	4,68
Variância	2,51	0,36
Desvio Padrão	1,58	0,6
Nº de Total de Pontos	12.022	9.396
Nº de pontos/ha	505	394

Tabela 2 – Comportamento dos dados antes e após a utilização do filtro.

Os valores adotados como parâmetros para o algoritmo variaram para as duas áreas em virtude das diferenças intrínsecas a cada uma sendo necessário obter estes valores para o filtro apresentar resultados satisfatórios.

CONCLUSÃO

A aplicação de filtros para a remoção de pontos onde os valores de produtividade coletados pelo monitor de colheita são discrepantes é importante para a obtenção de um mapa de produtividade com boa qualidade, evitando interpretações errôneas.

A criação de um algoritmo para eliminação dos erros é bastante simples sendo necessário entretanto o conhecimento dos principais tipos de erros presentes nos mapas bem como o tipo de monitor e colhedora utilizada.

A aplicação do algoritmo no grande volume de dados da colheita de uma área é bastante trabalhosa quando se utiliza a planilha eletrônica, além de exigir grande espaço para o armazenamento das extensas planilhas geradas.

O algoritmo criado permitiu remover boa parte dos erros presentes nos dados estudados podendo ser utilizado com a restrição de que sejam fornecidos adequadamente os parâmetros requeridos pelo mesmo.

A criação de um software com este tipo de algoritmo parece o próximo passo na redução dos erros presentes nos dados gerados pelos monitores de colheita e pelo sistema de posicionamento de modo mais eficiente.