

PROCESSAMENTO DE DADOS DE PRODUTIVIDADE DE SOJA POR TRÊS DIFERENTES MÉTODOS

MARCO ANTONIO ZANELLA¹, DANIEL MARÇAL DE QUEIROZ², DOMINGOS SÁRVIO MAGALHÃES VALENTE³

¹ Eng. Agrícola, Mestrando em Engenharia Agrícola, UFV, Viçosa – Minas Gerais, (31) 3899-3463, marco.zanella@ufv.br

² Eng. Agrícola, Prof. Doutor, Depto. Engenharia Agrícola, UFV, Viçosa – Minas Gerais (31) 3899-1882, queiroz@ufv.br

³ Eng. Agrícola, Prof. Doutor, Depto. Engenharia Agrícola, UFV, Viçosa – Minas Gerais (31) 3899-1876, valente@ufv.br

Apresentado no
XLVI Congresso Brasileiro de Engenharia Agrícola - CONBEA 2017
30 de julho a 03 de agosto de 2017 - Maceió - AL, Brasil

RESUMO: Mapas de produtividade são excelentes indicadores da variabilidade espacial podendo ser considerados como o primeiro passo para a implantação da agricultura de precisão. Contudo, dados de produtividade gerados pelas colhedoras, geralmente possuem erros sistemáticos que, se não removidos, podem levar a decisões errôneas de manejo. Este trabalho teve por objetivo comparar três métodos de filtragem dos dados de produtividade e identificar suas respectivas influências sob os mapas de produtividade gerados. Foi utilizado o programa de computador Yield Editor que é disponibilizado pela USDA, o programa de computador Map Filter que foi desenvolvido pelo Laboratório de Agricultura de Precisão (LAP) da ESALQ e por fim utilizou-se uma rotina de filtragem baseado nas estatísticas descritivas dos dados de produtividade utilizando-se o programa de computador QGIS. Os dados utilizados são referentes à safra de soja 2015/2016 no município de Iepê-SP e foram obtidos de monitores de produtividade comercial de duas colhedoras. Os resultados foram submetidos a análise de estatística descritiva, os programas de computadores Yield Editor e Map Filter apresentaram decréscimo no coeficiente de variação e obteve-se melhores resultados qualitativos elevando a qualidade final do mapa gerado. Já na filtragem realizada no QGIS observou-se um menor decremento no coeficiente de variação da produtividade.

PALAVRAS-CHAVE: agricultura de precisão, erros em mapa de produtividade, filtragem de dados de produtividade.

YIELD DATA PROCESSING OF SOYBEAN BY THREE DIFFERENT METHODS

ABSTRACT: Yield maps are excellent indicators of spatial variability and can be considered as the first step in the implementation of precision agriculture. However, yield data generated by the combine harvester usually have systematic errors that, if not removed, can lead to erroneous management decisions. The objective of this work was to compare three methods of filtering the yield data and identify their respective influences on the yield maps generated. We used the Yield Editor software that is available from USDA, the Map Filter software developed by Precision Agriculture Laboratory (LAP) - ESALQ and a filtering routine based on the descriptive statistics of the data using the QGIS software. The data used to analyze the three methods of yield map generation was from the 2015/2016 soybean crop cultivated in the Iepê county, Sao Paulo state and were obtained from commercial yield monitors of two combine harvesters. The results were submitted to descriptive statistics analysis, the Yield Editor and Map Filter softwares showed a decrease in the coefficient of variation and better qualitative results were obtained, enhancing the final quality of the generated map. The filtering performed using the QGIS software showed a smaller decrease in the coefficient of variation of yield.

KEYWORDS: precision agriculture, yield map errors, filtering yield data.

INTRODUÇÃO: A produção de grãos no país tem um papel de destaque no agronegócio brasileiro e mundial. A safra de 2016/2017 deve ter um incremento de 15,3% em relação à safra anterior, atingindo uma produção total de 215,3 milhões de toneladas (CONAB, 2017). Esta produção torna o Brasil um dos principais países produtores de grãos e com potencial para expansão. A produtividade de grãos brasileira tem aumentado de forma muito mais acelerada que a expansão da área plantada (CONAB, 2017). Sendo que este incremento na produtividade provavelmente ocorreu devido ao melhoramento genético, avanços na mecanização agrícola, desenvolvimento de pesquisa e novas tecnologias entre outros fatores (SENTELHAS et al., 2015). A agricultura de precisão requer o pleno conhecimento para que seja realizada a tomada de decisões, podendo ser parte de uma ação importante no aumento da produtividade das culturas (BUTTAFUOCO et al., 2016). Contudo, o grande desafio no país está associado com a disseminação de tecnologias, bem como o uso adequado que pode proporcionar um aumento na produtividade (SENTELHAS et al. 2015). Os Mapas de produtividade podem ser considerados como o primeiro passo em direção a adoção da agricultura de precisão (ROSS; MORRIS; JOHANNSEN, 2008), no entanto ao serem coletados diretamente dos monitores das colhedoras podem apresentar erros sistemáticos, os quais necessitam ser identificados e/ou removidos. Para isso, existem diversas metodologias para remoção de erros de dados de produtividade. Estas variam em interface, recursos e complexidade dos processos. Sudduth, Drummond e Myers, (2012) desenvolveram a versão 2.0 do programa computacional Yield Editor. Este possui seleção automatizada de vários valores para a filtragem e uma melhor visualização dos dados em relação a versão 1.0. O *software* livre Map Filter (SPEKKEN; ANSEMI; MOLIN, 2013) foi desenvolvido pelo Laboratório de Agricultura de Precisão da ESALQ-USP para a análise e remoção de dados inconsistentes em mapas de alta densidade. Já o programa QGIS é um sistema de informação geográfica (SIG) de livre acesso (*open source*), o qual permite o processamento de dados e elaboração de mapas temáticos. Com base nesse contexto, o presente trabalho teve como objetivo comparar três métodos de filtragem de dados de produtividade e identificar suas respectivas influências sob os mapas de produtividade gerados.

MATERIAL E MÉTODOS: Os dados de produtividade foram adquiridos de monitores instalados em duas máquinas colhedoras, e são referentes as safras de 2015/2016 para a cultura de soja. A área onde os dados foram coletados é uma propriedade rural produtora de grãos que está localizada no município de Iepê ao sudoeste do estado de São Paulo, a cerca de 530 quilômetros da capital São Paulo. Posicionada entre as coordenadas: 22°39'51.89" S de latitude e 51°2'14.07" O de longitude. Foram utilizados três programas de computador de domínio público Yield Editor 2.0, Map Filter e QGIS. O programa de computador disponibilizado pela USDA, Yield Editor 2.0 é o que apresenta mais opções para a filtragem dos dados de produtividade. O programa Map Filter é destinado a análise e remoção de dados inconsistentes em mapas de alta densidade. Enquanto que o QGIS, é um SIG, em que é possível analisar e editar informações espaciais de diversos formatos de arquivos. Dentre as várias funções disponíveis no Yield Editor 2.0, foram utilizados os filtros de atraso de início de passada e de fluxo de grãos, de máxima e mínima velocidade e de máxima e mínima produtividade. Os parâmetros dos filtros foram estabelecidos conforme as recomendações de Sudduth e Drummond (2007), sendo alguns ajustados automaticamente. Os dados de entrada no programa foram: produtividade, fluxo de grãos, distância entre pontos, umidade, largura da passada, velocidade, número da passada, intervalo de leitura e status de colheita. O programa utiliza o sistema de unidades britânico, sendo necessário a transformação de dados de

produtividade, fluxo de grãos, velocidade e largura da passada. No Map Filter utilizou-se um raio de busca de mesmo tamanho da largura da plataforma da colhedora, 12 metros e 20% para o percentual máximo de variação aceito para a média. Neste caso, se o valor do coeficiente de variação (CV) ou a média encontrada para este ponto for maior que o máximo estabelecido, um peso era adicionado a todos estes, conforme descrito por Spekken, Anselmi e Molin (2013). Ao término deste processo, os pontos são classificados em adequados, inconsistentes e inadequados. Foram removidos os dados classificados pelo programa como inconsistentes e inapropriados, sendo considerados *outliers*. Para a filtragem dos dados no QGIS, foi realizada a remoção de dados que são impossíveis de ocorrer, aqueles com valor próximo a zero ou produtividade muito elevada. Os limites inferior e superior foram de 900 e 5100 kg/ha, respectivamente. Posteriormente, os dados foram submetidos a uma rotina de filtragem baseada na estatística descritiva, utilizando a média e o desvio padrão (DP) do conjunto de dados. Assim, foram considerados *outliers* os pontos que ficaram situados fora dos limites $média \pm 2,5 DP$. Para avaliar a influência de cada método, foi realizada uma análise da estatística descritiva. Os parâmetros analisados foram a média, valor mínimo, valor máximo, variância, desvio padrão, coeficiente de variação e as medidas de posição e dispersão (coeficiente de assimetria e curtose). Foram elaborados mapas temáticos com os dados de produtividade filtrados e aqueles removidos. Por fim, os dados filtrados foram interpolados para avaliação qualitativa dos mapas.

RESULTADOS E DISCUSSÃO: Os histogramas da distribuição de frequência dos dados de produtividade dos dados obtidos diretamente do monitor da colhedora sem processamento (dados brutos), e dos dados filtrados pelos três métodos estão apresentados na Figura 1. Os parâmetros da estatística descritiva destes dados de produtividade estão apresentados na Tabela 1.

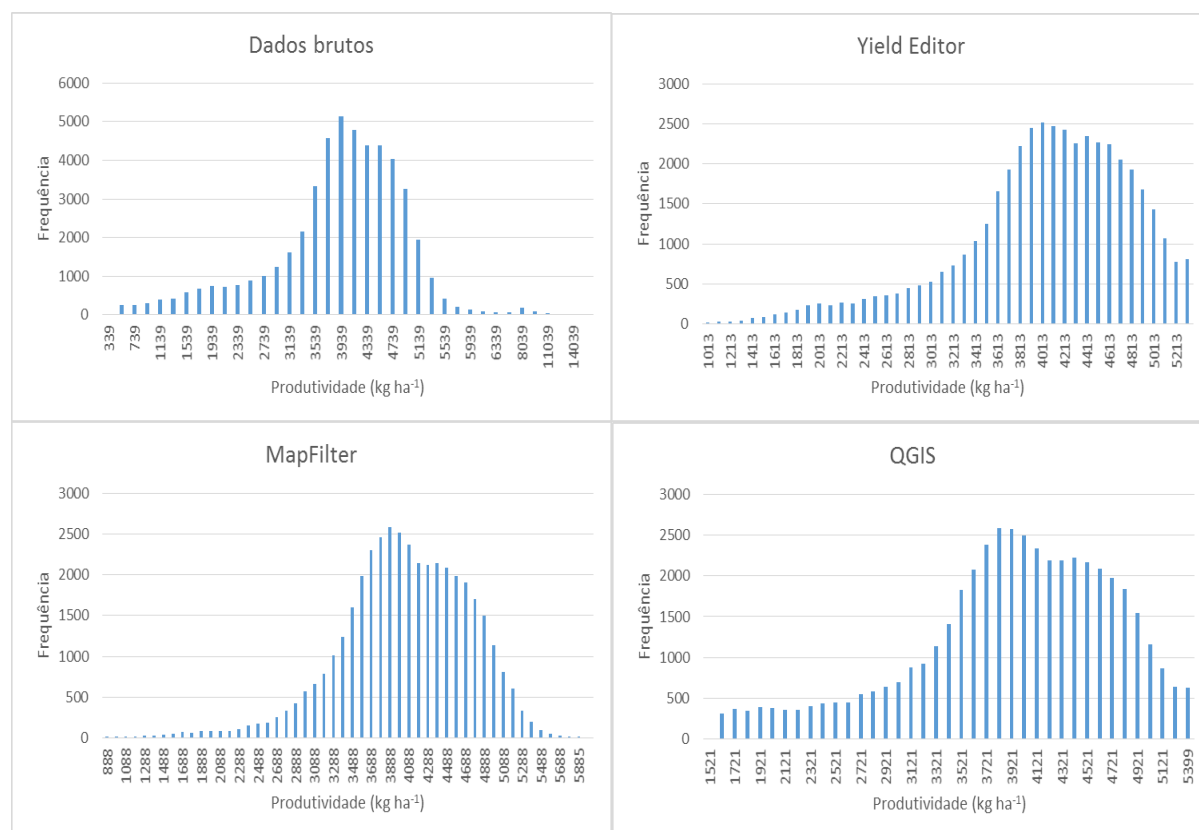


FIGURA 1. Histogramas da distribuição de frequência dos dados de produtividade brutos e dos dados filtrados por três métodos.

A partir dos resultados da estatística descritiva foi possível observar que os dados brutos apresentaram um coeficiente de assimetria positivo de 0,05. Enquanto que, para os dados filtrados, os coeficientes calculados são negativos e apresentaram valores de -0,90, -0,60 e -0,72. A distribuição dos dados se afasta da normalidade, sendo caracterizada como bimodal. Os valores do coeficiente curtose demonstram que os dados brutos apresentam uma distribuição leptocúrtica (5,27), enquanto os dados filtrados apresentaram distribuição platicúrtica. Os valores do coeficiente de curtose foram 0,86, 0,89 e 0,21, respectivamente. Como os dados foram provenientes de duas colhedoras, é possível que este seja o motivo da distribuição bimodal observada nos histogramas. Em cada método de filtragem foi observado um aumento do valor médio de produtividade em relação aos dados não filtrados, sendo os aumentos de 4,63% para o Yield Editor e Map Filter e 1,36% para a filtragem no QGIS. O desvio padrão reduziu em todos os métodos, resultando na redução do coeficiente de variação que apresentou valores de 19,47%, 16,95% e 21,05%, respectivamente. Assim, tanto os valores de CV quanto os valores de coeficiente de assimetria estão de acordo com os valores encontrados por Joernsgaard e Halmoe (2003) para os dados brutos, sendo que a distribuição dos valores de produtividade observados pelos autores foram geralmente não normal. Para os dados filtrados a média de produtividade tende a aumentar, enquanto o DP e CV tende a diminuir em relação aos dados não filtrados (SUDDUTH; DRUMMOND, 2007; SUN et al., 2013).

TABELA 1. Estatística descritiva dos dados de produtividade obtidos diretamente da colhedora (dados brutos) e dos dados filtrados por cada um dos métodos.
Descriptive statistics of the yield data obtained directly from the combine harvester (raw data) and data filtered by each of the methods.

| Parâmetros | Dados brutos | Dados filtrados | | |
|----------------------------------|--------------|-----------------|------------|----------|
| | | Yield Editor | Map Filter | QGIS |
| Média (kg ha ⁻¹) | 3820 | 3997 | 3997 | 3872 |
| Mediana (kg ha ⁻¹) | 3941 | 4687 | 4014 | 3959 |
| Mínimo (kg ha ⁻¹) | 339 | 1013 | 888 | 1521 |
| Máximo (kg ha ⁻¹) | 13430 | 5379 | 5885 | 5399 |
| Intervalo (kg ha ⁻¹) | 13091 | 4365 | 4997 | 3878 |
| Variância | 1216271,6 | 606024,6 | 458911,8 | 664903,6 |
| Desvio padrão | 1102,8 | 778,5 | 677,4 | 815,4 |
| Curtose | 5,27 | 0,86 | 0,89 | 0,21 |
| Assimetria | 0,05 | -0,90 | -0,60 | -0,72 |
| Coeficiente de Variação (%) | 28,86 | 19,47 | 16,95 | 21,05 |
| Número de pontos | 50231 | 43734 | 41214 | 46865 |

A porcentagem de pontos de dados removidos foi de 12,9%, 17,9% e 6,7%, respectivamente. Isto impactou diretamente na variação dos parâmetros observados, que foi maior nos métodos que removeram maior quantidade de pontos. As menores variações dos parâmetros estudados em relação aos dados brutos, foram observadas no método utilizado no QGIS. Os dados filtrados e removidos foram plotados para avaliação visual. Na Figura 1 são apresentados os mapas temáticos resultantes dos métodos de filtragem, mostrando a distribuição dos pontos filtrados e dos removidos. Foi possível observar que muitos pontos removidos estão localizados nas bordas do talhão, indicando a ocorrência de erros no início e fim das passadas da colhedora. O programa de computador Yield Editor possui um filtro específico para este

tipo de erro. Os filtros eliminam pontos com valores de produtividade baixas e não confiáveis, pois são obtidos à medida que a colhedora se enche no início e depois esvazia-se no final de cada passada (SUDDUTH; DRUMMOND; MYERS, 2012). Observa-se que os outros métodos, também retiram muitos pontos nas bordas do talhão.

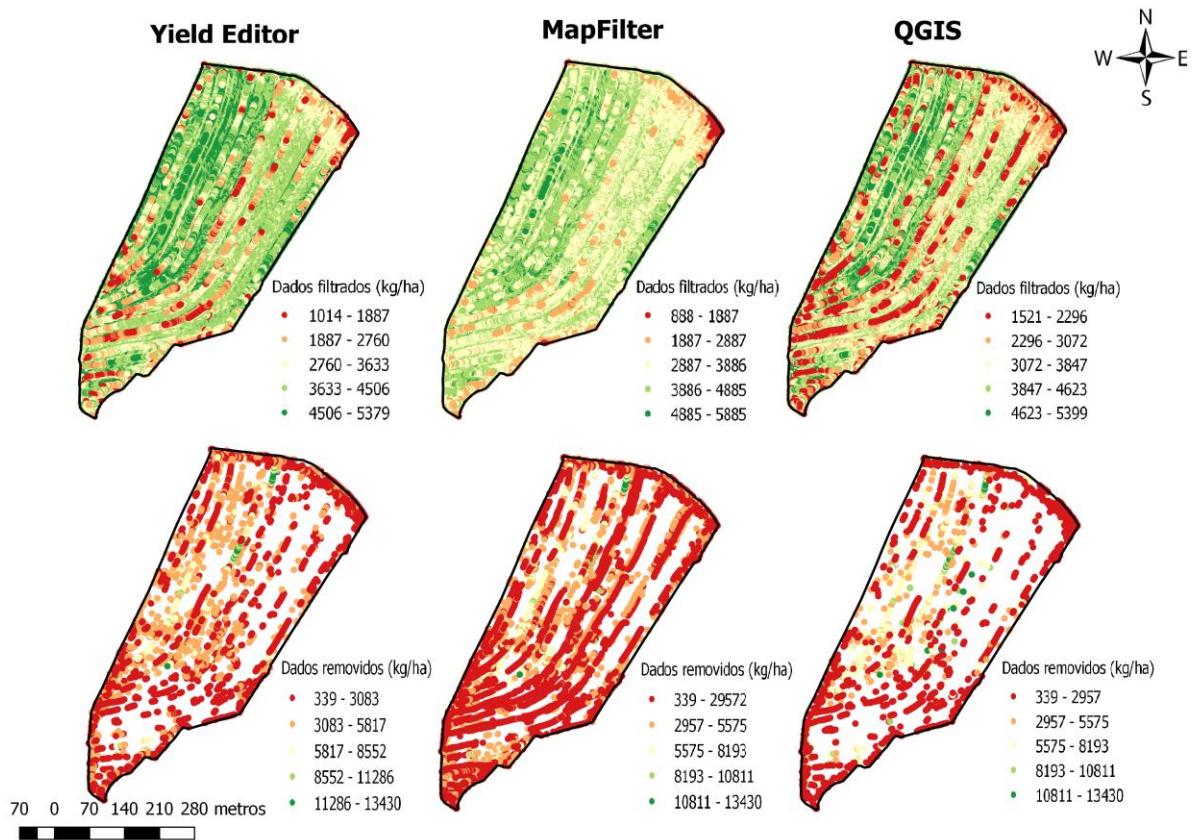


FIGURA 2. Mapas temáticos resultantes de três métodos de filtragem de produtividade e mapas com os dados removidos por cada um dos métodos.

Utilizando os dados de produtividade filtrados da Figura 2, foi realizada a interpolação dos pontos pelo método matemático do “inverso da distância na segunda potência”. Os resultados da interpolação são apresentados na Figura 3.

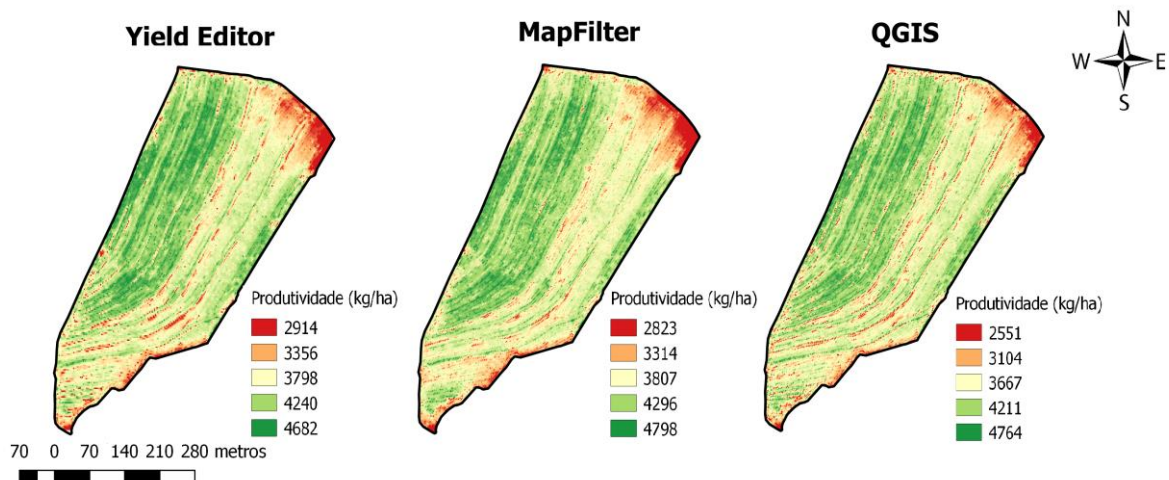


FIGURA 3. Mapas de produtividade interpolados pelo inverso da distância na segunda potência.

Diante dos mapas da Figura 3, foi possível observar as regiões de maior e de menor produtividade. Ao comparar os três métodos nota-se que todos identificaram as regiões de menor produtividade e as regiões de maior produtividade, variando o tamanho das mesmas. É possível observar também que o mapa interpolado pelo Map Filter apresentou uma maior suavização na espacialização dos resultados. Os valores de produtividade nos mapas interpolados variaram de 2914 kg ha⁻¹ a 4632 kg ha⁻¹ para o Yield Editor, de 2823 kg ha⁻¹ a 4798 kg ha⁻¹ para o Map Filter e 2551 kg ha⁻¹ a 4764 kg ha⁻¹ para o QGIS.

CONCLUSÕES: Os dados filtrados utilizando os programas computacionais Yield Editor e Map Filter apresentaram maior decréscimo no coeficiente de variação, enquanto que para os dados filtrados no QGIS a redução foi menor. Deste modo, a interpolação dos dados filtrados apresentou melhores resultados, elevando o nível qualitativo das informações nos mapas de produtividade. Apesar de apresentar resultados menos expressivos quanto a filtragem dos dados, o QGIS possibilitou de maneira precisa processar os dados e gerar os mapas de produtividade.

AGRADECIMENTO: Os autores agradecem a Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG pelo apoio financeiro e a Coordenação de Pessoal de Nível Superior – CAPES pela concessão da bolsa de mestrado.

REFERÊNCIAS

- BUTTAUFUOCO, G. et al. Geostatistical modelling of within-field soil and yield variability for management zones delineation: a case study in a durum wheat field. **Precision Agriculture**, p. 1–22, 2016.
- CONAB: COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da Safra Brasileira de Grãos Monitoramento agrícola- Safra 2016/2017**. [s.l.: s.n.]. Disponível em: <<http://www.conab.gov.br>>.
- JOERNSGAARD, B.; HALMOE, S. Intra-field yield variation over crops and years. **European Journal of Agronomy**, v. 19, n. 1, p. 23–33, 2003.
- ROSS, K. W.; MORRIS, D. K.; JOHANNSEN, C. J. A review of intra-field yield estimation from yield monitor data. **Applied Engineering in Agriculture**, v. 24, n. 3, p. 309–317, 2008.
- SPEKKEN, M.; ANSELMINI, A. A.; MOLIN, J. P. A simple method for filtering spatial data. In: STAFFORD, J. V (Ed.). **Precision agriculture '13**. Wageningen: Wageningen Academic Publishers, 2013. p. 259–266.
- SUDDUTH, K. A.; DRUMMOND, S. T. Yield editor: Software for removing errors from crop yield maps. **Agronomy Journal**, v. 99, n. 6, p. 1471–1482, 2007.
- SUDDUTH, K. A.; DRUMMOND, S. T.; MYERS, D. B. Yield Editor 2.0: Software for Automated Removal of Yield Map Errors, 2012 Dallas, Texas, July 29 - August 1, 2012. St. Joseph, Mich.: **ASABE - American Society of Agricultural and Biological Engineers**, 2012.
- SUN, W. et al. An integrated framework for software to provide yield data cleaning and estimation of an opportunity index for site-specific crop management. **Precision Agriculture**, v. 14, n. 4, p. 376–391, 2013.