10º **ENTEC – Encontro de Tecnologia: 28 de novembro a 3 de dezembro de 2016**

A LINGUAGEM "R" COMO INSTRUMENTO DE APOIO À AGRICULTURA

Germano Rodrigues Ferreira[1]; Luciano Rodrigo Ferreira[2]
[1, 2] Universidade de Uberaba - UNIUBE
germanorf@gmail.com; luciano.ferreira@uniube.br

Resumo
A linguagem "S" foi utilizada por vários anos, entretanto foi superada pela linguagem "R", devido à recursos mais sofisticados que vieram de encontro às necessidades do mercado. Neste caso, destaca-se a função revolucionária fornecida pela linguagem "R" que possui um ambiente operacional dotado de praticidade, excelente para execução de cálculos, além da produção de gráficos que são capazes de fornecer a devida precisão e qualidade diante do gerenciamento de dados estatísticos. A praticidade que é proporcionada pela linguagem "R", contribuiu plenamente para produção de softwares, e tem se tornado muito popular não só no mercado profissional, mas também entre os meios acadêmicos, que fazem uso constante, em suas diversas atividades de trabalho. Um dos fatores que faz com o que a linguagem "R" seja bem acessível, é em razão de ser oferecida gratuitamente para a população; além de possuir um grande volume de pacotes diferenciados, sendo cada um destinado a uma função específica. A relação de áreas em que os softwares produzidos em linguagem "R" podem atuar com todo o potencial é bem extensa, e conforme analisado em um exemplo real no estudo de caso, prova ser eficiente também no segmento da agricultura, onde por meio da krigagem, é eficaz no que diz respeito a cálculos geoestatísticos.

Palavras-chave: Software. Cálculos. Gráficos. Acadêmicos. Geoestatísticos.

1 Introdução

No âmbito atual da informática ao se direcionar de maneira específica ao segmento de softwares, nota-se que há uma ampla produção de programas destinados as mais diversas funcionalidades, as quais são requeridas pelas organizações comerciais, empresariais, pelos profissionais liberais, dentre outros.

A problemática quanto a um determinado software, envolve normalmente diversos fatores, tais como o seu alto custo, a dificuldade de acesso, certas restrições no seu desempenho, sua não aplicação em determinado sistema operacional, dentre outros.

Há algum tempo surgiu no mercado nacional, a linguagem denominada "R", sendo esta destinada ao uso em vários aspectos, tais como a computação que abrange estatística e gráficos de forma geral. São também destaques desta linguagem R, os fatores relacionados diretamente a manipulação de dados e a visualização de elementos.

No estudo de caso é apresentado um exemplo real de utilização eficiente de um software específico que foi produzido por intermédio da linguagem R, sendo este executado continuamente no segmento de geoestatística. Foi analisado um exemplo de realização do processo de krigagem, que visa determinar o nível de acidez do solo de uma área localizada no município de Perdizes, no Estado de Minas Gerais; tendo-se como objetivo escolher o melhor tipo de cultura para determinada região. Justifica-se fazer uma análise plena do que está envolvido na linguagem "R", em razão de demonstrar

que há sim à disposição da população, um recurso tecnológico tão eficiente, fato este que será provado, mediante o estudo das suas características, sua real importância, aplicabilidade, especificações e suas vantagens.

O artigo que se inicia tem como objetivo de comprovar que a utilização da linguagem “R” é plenamente viável e resulta em benefícios aos mais variados segmentos, devendo esta ser cada vez mais indicada e adotada na sociedade moderna.

2 Materiais e Métodos

O foco principal aborda uma apresentação plena da linguagem “R”, demonstrando alguns conceitos presentes na literatura, o ambiente operacional, os pacotes que podem ser utilizados, as diversas funções, alguns recursos que são disponibilizados, a aplicabilidade dos softwares produzidos em linguagem “R”, além das vantagens e acessibilidade da mesma.

Em razão do exemplo de uso de um software em linguagem “R”, foram observados alguns fatores sobre a geoestatística e a krigagem.

Dentre os materiais que foram utilizados na revisão bibliográfica, estão os livros e artigos científicos disponíveis na internet. Na produção do estudo de caso, foi usado um software produzido em linguagem “R”, o qual foi fundamental para que fosse executada uma análise laboratorial de amostras de solo, e consequentemente fossem elaborados gráficos com informações precisas, sobre o nível de pH.

Os resultados foram analisados por profissionais plenamente capacitados, e o conteúdo, expressos através de gráficos e tabelas.

3 Resultados

Os resultados práticos que comprovam a eficácia dos softwares produzidos em linguagem “R”, ficaram evidentes no exemplo real apresentado no estudo de caso.

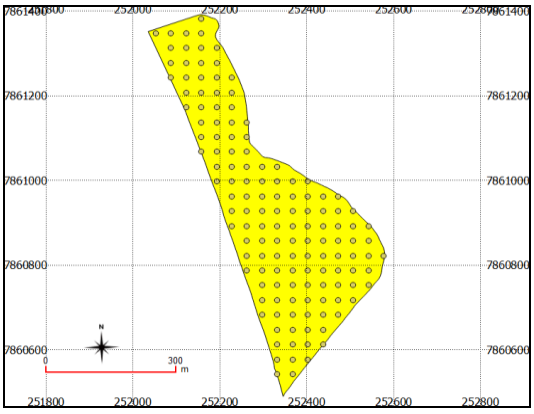
O estudo teve como objetivo verificar a qualidade do solo, por intermédio de amostragem e aplicação da krigagem. Ressalta-se que o resultado e as imagens apresentadas, refletem a eficácia obtida no procedimento realizado.

Em relação à área disponibilizada para os estudos, a mesma corresponde a 17 hectares de uma área de testes pertencente a uma fazenda, situada na zona rural da cidade de Perdizes, no Estado de Minas Gerais.

Foi realizada a espacialização das amostras em toda a extensão da área, com o objetivo de que ficassem distribuídas uniformemente evitando ao máximo as distorções. Da maneira citada é possível obter a verificação dos atributos desse solo.

Conforme consta nas informações que estão presentes na Figura 1 a seguir, foram determinados os resultados das análises geoestatísticas executadas em um total de 139 pontos amostrais.

Figura 1: Mapa da área com as coordenadas geográficas e representação dos pontos amostrados

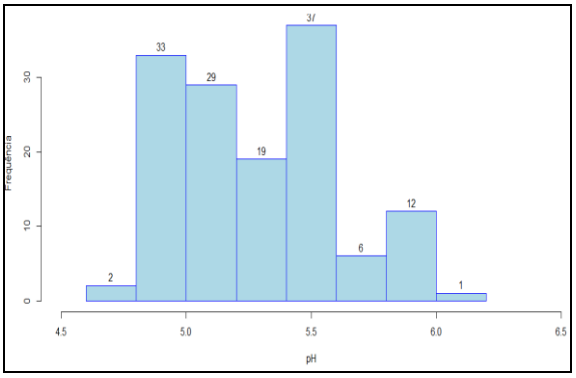


A análise exploratória iniciou através do gráfico *boxplot*, e através deste percebeu-se que não houve ocorrências

de valores atípicos, ou seja, não há amostras fora dos padrões normais.

Na sequência executou-se o histograma, que constitui em um método eficaz de ilustrar as frequências aferidas das amostras. Na Figura 2 é evidente que o pico da frequência ocorre quando o pH atinge 5,5, situação esta que indica que o solo tem um alto índice de acidez.

Figura 2: Histograma mostrando os níveis de pH



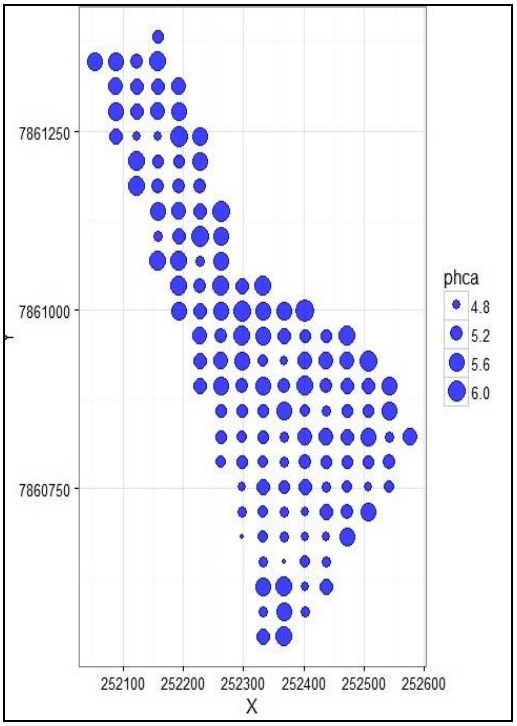
Na Tabela 1, a seguir, constam os valores referentes as estatísticas descritivas que foram obtidas através da análise laboratorial.

Tabela 1 - Valores das estatísticas descritivas.

Estatísticas	Valores
Valor Mínimo	4,700
Valor Máximo	6,100
Média	5,288
Mediana	5,300
1° Quartil	5,050
3° Quartil	5,500
Variância	0,102
Desvio Padrão	0,320
Coeficiente de Variação	6,045

A primeira etapa da análise foi realizada, constando a exibição dos níveis de pH (Figura 3) que foram determinados no estudo, em conformidade com a imagem que foi produzida.

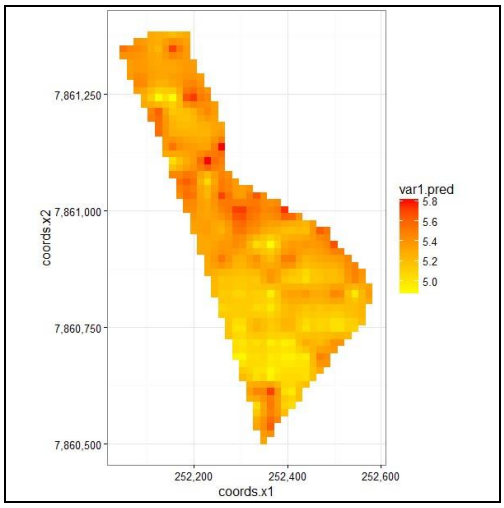
Figura 3: Escala de pH



Após ter sido aplicado a correção do variograma, foram determinados todos os pontos com medição que possuem informações das análises realizadas em laboratório e os pontos que iriam ser estimados.

Após a validação do semivariograma, foi realizado o processo de krigagem ordinária como forma de executar a interpolação, que está presente no pacote denominado de "*gstat*". Através da krigagem foram produzidos os índices estimados, para toda extensão da área além dos pontos que já possuíam a informação, também foram expostos os valores estimados do pH para os pontos interpolados conforme pode ser observado na Figura 4.

Figura 4: Pontos e índices produzidos através da krigagem



Foi possível notar que o índice de acidez encontrado foi de 5,5 e este nível é considerado como alto, não sendo indicado para alguns tipos de cultura (EMBRAPA, 2015).

O índice de 5,5 de pH permite que sejam cultivadas culturas específicas, tais como a batata, que neste caso é capaz de suportar um pH de até 6,0. Nota-se que o percentual de pH identificado através do processo krigagem, é de suma importância, pois de acordo com o índice encontrado, pode-se determinar a cultura mais adequada a área pesquisada (EMBRAPA, 2015).

No estudo de caso foi mostrado na prática, que é possível determinar com precisão a variação espacial do pH do solo, utilizando-se de recursos tais como a função semivariograma e a krigagem. O processo utilizando-se *scripts* produzidos na plataforma da linguagem "R", se mostra eficiente tanto na geração de figuras, gráficos e dados numéricos.

4 Discussão

Verificou-se de acordo com o que foi mencionado por Alvim (2001), que a linguagem "R" corresponde a um ambiente computacional destinado a realizar cálculos estatísticos, e também criar gráficos.

Uma síntese dos pontos positivos de um software em linguagem "R", que é essencial para se manipular as informações, os cálculos, e visualizar os gráficos produzidos por este, é citada da seguinte maneira:

- uma facilidade efetiva para manipulação e armazenagem de dados;
- um conjunto de operadores para cálculos sobre quadros de dados, em particular as matrizes;
- uma grande e coerente coleção integrada de ferramentas intermediárias para análise de dados;
- facilidades gráficas para análise de dados e visualização na tela ou impressa (VIEIRA, 2009, p. 1).

Segundo Barbetta (2005), em razão de os softwares produzidos em linguagem "R", terem inúmeros recursos eficientes, estes são popularmente usados por acadêmicos; o que é impulsionado também em vista de possuir código livre.

Nota-se que a linguagem R pode ser muito utilizada devido ao fato de funcionar perfeitamente em grande parte dos sistemas operacionais existentes na atualidade, tais como o *Macintosh*, o *Linux*, *Windows* e outros. Esta característica evidencia a flexibilidade desta linguagem (ALVIM, 2001).

De acordo com Vieira (2009), ao se falar em "ambiente da linguagem R" tem-se a pretensão de afirmar que se trata de um sistema, onde para a sua edição aplicou-se um devido planejamento, além de este possuir ampla coerência, ou seja, sua estrutura vai muito além de um simples conjunto de "recursos" que nem ao menos permitem a sua flexibilização.

Baseado no que foi citado por Vieira (2009), pode-se concluir que a linguagem “R” tem como diferencial, um formato de documentação exclusivo, o qual através deste pode-se obter uma documentação de fácil compreensão, ocorrendo esta de forma online ou por meio impresso.

Em se tratando dos diversos pacotes que podem ser inseridos em um software em linguagem “R”, estes evidenciam a sua real importância, visto que desta forma, a sua ampla funcionalidade só tende a ampliar, em razão das constantes pesquisas visando o seu aprimoramento tecnológico.

Além dos milhares de pacotes oficiais, existe também a disposição para *download*, diversos pacotes que são considerados como “não sendo oficiais” (VIEIRA, 2009).

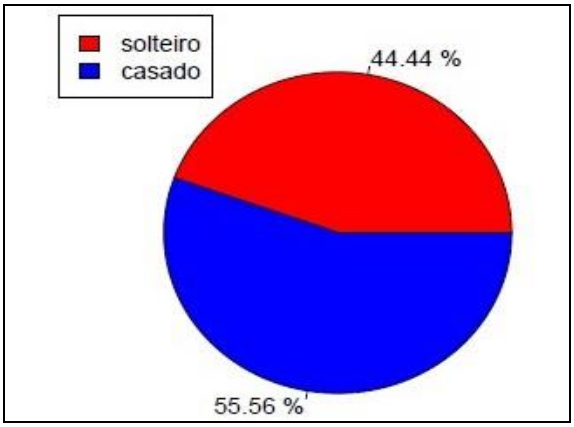
Em conformidade com Landeiro (2011), os softwares elaborados com a linguagem “R”, mediante a diversidade de funções, apresenta inúmeras possibilidades de execução. Tendo como objetivo facilitar o trabalho por parte do usuário, foi desenvolvido um sistema de ajuda, para que possa ser encontrada e verificada a função ideal para um determinado fim.

Pode-se concluir tendo como base os autores pesquisados, que a relação ampla de funções referentes a linguagem “R”, a facilidade no processo de busca e na utilização das mesmas, contribuem para que o trabalho do usuário ocorra de maneira prática.

Em razão do que foi mencionado por Amaral (2009), pode ser relatado que os softwares produzidos em linguagem “R”, são providos de vários recursos que contribuem para o aprimoramento visual das imagens, como no caso do “*boxplot*” que tem como função apresentar de forma gráfica os dados relacionados a distribuição das variáveis.

Em síntese, ao se utilizar todos os recursos gráficos, pode-se obter os mais variados tipos de gráficos, independentemente do seu formato e com máxima qualidade como apresentado na Figura 5.

Figura 5: Exemplo de gráfico de pizza, produzido mediante um software em linguagem “R”



Todos os recursos e funcionalidades do software em linguagem “R” já mencionados e analisados, justificam a sua grande aplicação na sociedade, incluindo os segmentos acadêmico, comercial, financeiro, industrial, na agricultura, na mineração, dentre outros.

Ao se mencionar que os softwares em linguagem “R” são realmente eficazes no segmento da agricultura, esta afirmação pôde ser averiguada no estudo de caso, onde foi realizada uma análise de amostras de solo, visando verificar o índice de pH presente no mesmo; tendo-se como resultado a definição do tipo de cultura mais apropriado para cultivo. De fato, os recursos gráficos dentre outros pertinentes a linguagem “R”, foram fundamentais para o sucesso da análise do estudo de caso.

As informações contidas no trabalho contribuem para que haja uma maior aceitação dos softwares em línguagem “R”, não somente no setor da agricultura, mas também nos mais diversos segmentos mencionados.

Ao final pode-se dizer que o objetivo do trabalho artigo foi deveras alcançado, ficando comprovada a eficácia dos softwares produzidos em linguagem "R", quando estes são utilizados no setor da agricultura.

5 Conclusão

Por todos os aspectos analisados no artigo, nota-se que a escolha da utilização da linguagem R e seu software para a elaboração de uma análise estatística pode trazer grandes diferenças no produto final, ou seja, em mapas, gráficos, números que orientam em tantas decisões importantes.

O software da linguagem R mostra-se bastante proveitoso e em alguns casos plenamente superior à muitas ferramentas existentes no mercado voltadas para tarefas específicas e que tenho um custo considerável.

Conclui-se mencionando que a linguagem "R" é uma poderosa ferramenta, além de ser robusta, e estar em crescente utilização nas mais diversas áreas de conhecimento, incluindo o setor agropecuário.

Ressalta-se, ainda, que o estudo apresentado tem o intuito de demonstrar a aplicação da linguagem em um cenário real, e ressalva que a análise da linguagem "R" poderia ainda ser melhorada quantitativamente com a utilização de outros pacotes e implementação de novos códigos.

Referências

ALVIM, A. **Linguagem R**. Universidade Federal do Pernambuco. 2001.

AMARAL, M. R. dos S. do; CESARIO, C. de V. **Apostila do minicurso: Software R**. 2009.

BARBETTA, P. **Linguagem R: R-Gui e R-Commander**. 2005.

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. **Sistema de produção da batata: correção do solo**. 2015.

LANDEIRO, V.L. **Introdução ao uso do Programa R**. Instituto Nacional de Pesquisas da Amazônia. Programa de Pós Graduação em Ecologia. 2011.

VIEIRA, T.; CERQUEIRA JÚNIOR, P.; SANTOS, F.; MENDES, D. **Curso de introdução ao R**. 2009.