

ANÁLISE ESTATÍSTICA DO USO DA MÉDIA MENSAL DO NDVI NO MONITORAMENTO DA CANA-DE-AÇÚCAR NO ESTADO DE SÃO PAULO

MARIA EDUARDA MENEGHETTI MOREIRA^{1,2}, MANOEL REGIS LIMA VERDE LEAL¹, MICHELLE CRISTINA ARAUJO PICOLI¹, DANIEL GARBELLINI DUFT¹

¹ CNPEM/CTBE - Laboratório Nacional de Ciência e Tecnologia do Bioetanol (Rua Giuseppe Máximo Scalfaro, 10.000 – Campinas/SP, Brasil)

² Faculdade de Engenharia Agrícola (FEAGRI) - UNICAMP (Av. Cândido Rondon, 501 - Barão Geraldo 13083-875 - Campinas/SP)

Apresentado no
XLIV Congresso Brasileiro de Engenharia Agrícola - CONBEA 2015
13 a 17 de setembro de 2015 - São Pedro - SP, Brasil

RESUMO: O Boletim de Monitoramento de Cana-de-Açúcar no Estado de São Paulo é um produto do Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE) que faz uma análise mensal da produção de cana-de-açúcar, considerando o índice de vegetação por diferença normalizada (NDVI), provindo dos satélites AQUA e TERRA, e a precipitação mensal das mesorregiões produtoras do estado de São Paulo. O NDVI é um dos índices de vegetação mais utilizados para avaliação de culturas agrícolas devido à sua correlação com a biomassa e evapotranspiração da planta. O objetivo deste trabalho foi avaliar a metodologia utilizada para a elaboração do Boletim, para isto, foi realizada uma análise estatística da viabilidade do uso da média mensal das imagens NDVI como valor representativo do período. É utilizada a média do NDVI para igualar a escala dos dados mensais de precipitação. Essa análise foi realizada para a cultura da cana nas 12 mesorregiões canavieiras do estado. Para isso foram realizados testes de normalidade e ao rejeitar a normalidade dos dados foi realizado o teste não-paramétrico de Wilcoxon. Como resultado, foi observado que, com 5% de significância, é possível utilizar a média mensal das imagens NDVI para indicar o desenvolvimento da cultura da cana no estado de São Paulo.

PALAVRAS-CHAVE: NDVI, cana-de-açúcar, monitoramento.

STATISTICAL ANALYSIS OF THE USE OF NDVI MONTHLY AVERAGE IN THE MONITORING OF SUGARCANE IN THE STATE OF SÃO PAULO

ABSTRACT: The Sugarcane Monitoring Bulletin in São Paulo State is a product of Brazilian Bioethanol Science and Technology Laboratory (CTBE) that analyzes monthly the sugarcane production, considering the normalized difference vegetation index (NDVI), collected from AQUA and TERRA satellites, and monthly precipitation. NDVI is one of the most used vegetation index for crops evaluation due to its correlation with yield and evapotranspiration of the plants. The objective of this study was to evaluate the methodology used to produce the Bulletin, for this, a statistical

analysis of the viability for the use of NDVI images of the month average as the representative value for the period was performed. The NDVI average is used to match the scale with precipitation, which is monthly represented. This analysis was performed for the sugarcane crop in 12 sugarcane regions of the state. For this, normality tests were performed and the normality of the sample was rejected, after, was performed non-parametric Wilcoxon test. As a result, it was observed that, with 5% significance, it can be used the NDVI images average the month to indicate the development of sugarcane crop in São Paulo State.

KEYWORDS: NDVI, sugarcane, monitoring.

INTRODUÇÃO

Introduzida no período colonial, a cana-de-açúcar se transformou em uma das principais culturas da economia brasileira. O Brasil é o maior produtor de cana, além de liderar a produção mundial de açúcar e etanol (MAPA, 2015). O Brasil possui disponibilidade de terras cultiváveis para o plantio da cana em quase todo país, sem competição com outros alimentos, com tecnologia de produção e a estrutura para distribuição (Embrapa, 2009) e, São Paulo representa 60% da área de cana plantada no país (IBGE, 2015).

O Boletim de Monitoramento de Cana-de-Açúcar no Estado de São Paulo é um produto do Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE) que faz uma análise mensal da produção de cana nas mesorregiões produtoras do estado. A análise é baseada no índice de vegetação por diferença normalizada (NDVI), produto provindo do sensor MODIS, a bordo dos satélites AQUA e TERRA, e nos valores de precipitação mensal das mesorregiões analisadas. O NDVI é um dos índices de vegetação mais utilizados para avaliação de culturas agrícolas devido à sua correlação com a produtividade e evapotranspiração da planta (Huete et al., 2002).

O Boletim está na sua sétima edição, mas ainda é veiculado somente para um grupo de avaliadores, especialistas no mercado e pesquisadores, além disso, também é disponibilizado no site do CTBE. O produto tem ganho visibilidade no setor e tornou-se referência mensal de diversos sites de notícias agrícolas (por exemplo: Jornal da Cana, UDOP, entre outros).

Esse trabalho teve como objetivo avaliar a possibilidade em utilizar a média mensal das imagens NDVI, obtidas por satélite, ou invés dos valores absolutos dessas imagens, para que fosse possível igualar aos dados de precipitação, que estão na escala mensal para cada uma das mesorregiões. Para isso, foi necessário, através de recursos estatísticos, comparar os valores da média mensal de NDVI com os valores absolutos de NDVI.

Além de igualar as escalas dos dados (NDVI e precipitação), a média mensal das imagens permitirá que se tenha menos oscilações na curva de NDVI da safra, como mostra a Figura 1.

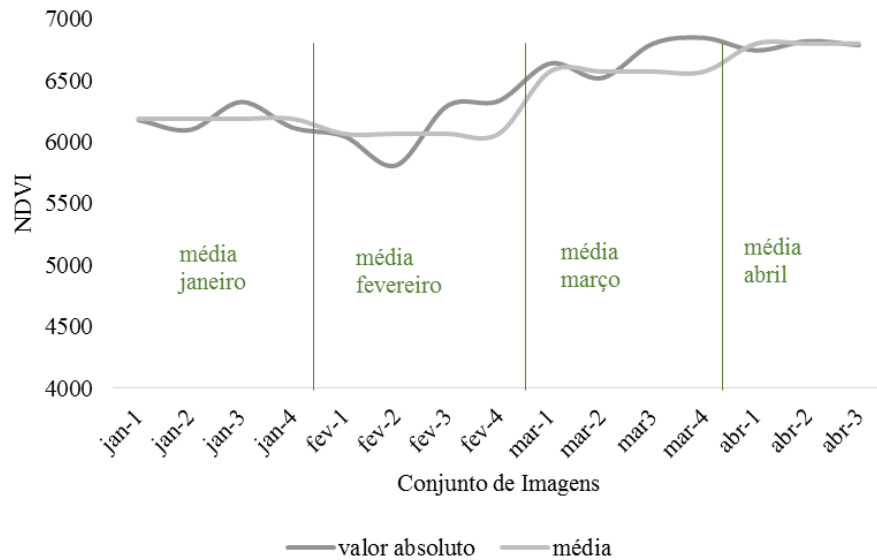


Figura 1. Gráfico com valores absolutos de NDVI (4 valores na composição dos meses de janeiro, fevereiro e março e, apenas 3 valores para compor o mês de abril) e média mensal de NDVI para a mesorregião de Presidente Prudente.

MATERIAIS E MÉTODOS

Mesorregiões

Os índices de vegetação por diferença normalizada (NDVI) foram obtidos para cada mesorregião produtora de cana-de-açúcar no estado de São Paulo – Brasil. As mesorregiões são doze: Araçatuba, Araraquara, Assis, Bauru, Campinas, Itapetininga, Marília, Macro Metropolitana Paulista, Piracicaba, Presidente Prudente, Ribeirão Preto, São José do Rio Preto (Figura 2).



Figura 2. Mesorregiões produtoras de cana-de-açúcar no estado de São Paulo.

NDVI

A NASA lançou, em dezembro de 1999, o satélite Terra, que vem coletando dados sobre as mudanças climáticas do planeta. O satélite é composto de 5 sensores com o objetivo de analisar as interações entre a atmosfera, terras, oceanos e energia de radiação da Terra (NASA¹, 2015). Em 2002, o satélite Aqua foi lançado com o objetivo de coletar informações sobre o ciclo hídrico do planeta (NASA², 2015) por meio de 6 sensores. Ambos, os satélites Aqua e Terra possuem em comum, o sensor MODIS (*Moderate-resolution Imaging Spectroradiometer*).

O sensor MODIS permite medir o crescimento da vegetação em escala global, fornece, segundo Weier e Herring (2010), alta resolução temporal; ou seja, gera imagens de um determinado pixel de uma região com uma repetição diária de coleta. Os mapas de vegetação provenientes das informações coletadas pelo satélite apresentam a densidade de crescimento da planta para todo o planeta. Os autores destacam que a medida resultante do sensor mais utilizada é a do índice de vegetação por diferença normalizada (NDVI).

O índice é obtido através da medição dos comprimentos e intensidades das ondas de luz infravermelho e vermelho visível do sol refletidas pela superfície da Terra, mais especificamente pelas plantas, de volta ao espaço. Weier e Herring (2010) explicam que a clorofila, pigmento das plantas, absorve a luz visível no processo de fotossíntese, enquanto a estrutura das células das folhas reflete o infravermelho.

Segundo Tucker (1979), Jordan (1969) foi o primeiro pesquisador a utilizar o método da razão infravermelho/vermelho para estimar biomassa ou índice de área foliar em uma floresta tropical. Mais tarde, o autor destaca que Cowell (1973, 1974) conclui que a razão de infravermelho/vermelho era efetiva em normalizar o efeito da variação da refletância do solo na estimativa de biomassa.

Contudo, foi Rouse et al. (1973, 1974 apud Tucker, 1979) que desenvolveu a medida de índice de vegetação. Para os autores, a localização e as variações do ciclo representavam uma grande fonte de erro. Dessa forma, utilizar a razão da diferença entre canais MSS7 e MSS5 do satélite LANDSAT pela soma dos mesmos, possibilitaria a obtenção de um índice de vegetação. Uma das conclusões do estudo foi de que inferências fenológicas poderiam ser colhidas para certas culturas de modo permitir a monitoramento desses tipos de vegetação.

A comparação entre a banda do vermelho e do infravermelho permite então medir a quantidade de vegetação no pixel. Segundo os autores, a obtenção do NDVI dá-se pela Equação 1:

$$NDVI = \frac{\text{radiação do Infravermelho} - \text{radiação do vermelho}}{\text{radiação do Infravermelho} + \text{radiação do vermelho}} \quad (1)$$

E sempre resulta em números que variam de -1 a +1. Contudo, a ausência de vegetação resulta em valores próximos de zero, então a escala considerada para avaliar NDVI é proposta de 0 a 1.

Gunnula et al. (2011) em pesquisa realizada na Austrália, avaliaram a relação entre o NDVI da cana-de-açúcar, com os padrões de precipitação. A pesquisa revela que nem a média da precipitação anual, nem a média no NDVI obtido para a cana apresentaram relação com a produção da área analisada, contudo o NDVI foi estabelecido como medida eficiente de avaliação da cultura, durante o período de chuvas, e está associado com a biomassa da planta e maturidade da cana. Os autores afirmam que variações temporais do NDVI estão linearmente ou log-linearmente relacionadas com a precipitação, e essa relação varia de acordo com o tipo de vegetação

e o solo em que esta se encontra. Ainda, discute que uma relação entre NDVI e produção final pode ser obtida quando utiliza-se uma resolução espacial maior do que a do sensor MODIS. Apesar da baixa resolução, o sensor permite medidas diárias e cobre grandes áreas.

Outras possíveis variáveis que afetam essa relação são: topografia, variedade da cultura e tipo de solo. Gunnula et al. (2011) concluíram que o NDVI proveniente do sensor MODIS é uma provável medida a ser utilizada na previsão de produção pré-colheita, e uma vez que a época de chuvas acaba, o NDVI sofre uma pequena queda em seus valores e o peso da cana aumenta.

Os NDVIs utilizados para a produção do Boletim provêm do sensor MODIS (*Moderate-resolution Imaging Spectroradiometer*) e são obtidos através dos produtos MOD13Q1 e MYD13Q1, dos satélites TERRA e AQUA respectivamente. Com os valores de radiação obtidos da banda do vermelho (visível) e do infravermelho próximo é possível calcular o valor de NDVI para cada pixel (250 x 250 metros) para todos os dias, porém é realizada a seleção dos maiores valores de NDVI em um período de 16 dias e para compor esses dois produtos, por isso, os produtos MOD13Q1 e MYD13Q1 são disponibilizados a cada 16 dias. É importante ressaltar que a disponibilização dos produtos MOD13Q1 e MYD13Q1 ocorre em um intervalo de 8 dias, portanto em um mês há aproximadamente de 3 a 4 imagens NDVI.

Para cada uma das imagens NDVI (MOD13Q1 ou MYD13Q1), fornecidas pelo sensor MODIS, são calculadas as médias para cada mesorregião, e então têm-se um único valor de NDVI para cada mesorregião por imagem.

Para calcular a média mensal das mesorregiões é realizado primeiramente o passo anterior, onde é calculada a média do NDVI de cada imagem para cada mesorregião. Em seguida, é feito o levantamento de quais são as imagens MOD13Q1 e MYD13Q1 que compõe aquele mês (de 3 a 4 imagens) e então é calculada a média do NDVI para aquele mês para cada uma das mesorregiões (Figura 3).

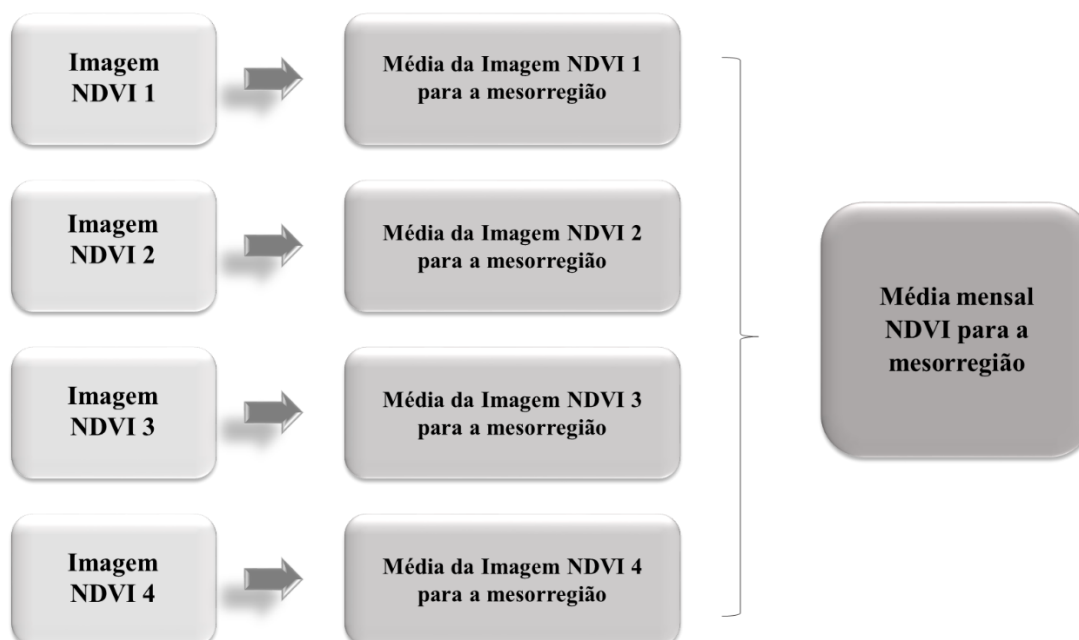


Figura 3. Fluxograma do processo para cálculo da média mensal para cada mesorregião, supondo que em um mês é tem-se quatro imagens NDVI da mesma região.

Contudo foi observado que algumas imagens, apesar de começarem nos meses em que foram atribuídas como valores para serem considerados no cálculo da média, representam, de modo mais significativo, o mês seguinte, como é o caso da imagem 2 do satélite MYD13Q1, a qual tem início de captação em 24 de janeiro e encerramento em 09 de fevereiro. Apesar da imagem ser atribuída ao mês de janeiro devido ao seu início, ela é mais representativa do mês de fevereiro.

Dessa forma, foi elaborada uma forma de cálculo das médias das imagens dos satélites de acordo com o mês de produção do Boletim para o cálculo das médias mensais (Tabela 1).

Tabela 1. Calendário das imagens consideradas no mês para a produção do Boletim.

Mês	Imagens		Imagens (dias)	
	MOD13	MYD13	MOD13	MYD13
Janeiro	1,2	23,1	Jan 01-16/Jan 17-01	Fev Dez 26-11 Jan/ Jan 08-24
Fevereiro	3,4	2,3	Fev 02-17/ Fev 18-05	Mar Jan 24-09 Fev/ Fev 09-25
Março	5,6	4,5	Mar 06-21/Mar 22-06	Abr Feb 25-13 Mar/ Mar13-29
Abril	7,8	6,7	Abr 07-22/Abr 23-08	Mai Mar 29-14 Abr/Abr 14-30
Mai	9	8,9	Mai 09-24	Abr 30-16 Mai/ Mai 16-01
Junho	10,11	10,11	Mai 25-09 Jun/ Jun 10-25	Jun 01-17/ Jun 17-03
Julho	12,13	12,13	Jun 26-11 Jul/ Jul 12-27	Jul 03-19/ Jul 19-04
Agosto	14,15	14,15	Jul 28-12 Ago/ Ago 13-28	Ago 04-20/ Ago 20-05
Setembro	16,17	16,17		Set 05-21/ Set 21-07
Outubro	18,19	18	Set 30-15 Out/ Out 16-31	Out 07-23
Novembro	20,21	19,2	Nov 01-16/ Nov 17-02	Dez Out 23-08 Nov/ Nov 08-24
Dezembro	22,23	21,22	Dez 03-18/ Dez 19-03	Jan Nov 24-10 Dez/ Dez 10-26

Estadística

Para analisar se era viável utilizar os valores de média mensal, ao invés dos valores absolutos de NDVI foram realizados testes estatísticos, primeiro de normalidade (para verificar a distribuição dos dados) e depois o teste de hipótese de igualdade para dados pareados (para verificar se havia igualdade entre as médias mensais e os valores absolutos de NDVI).

Teste de Normalidade

Testes de normalidade avaliam se uma amostra de dados de um experimento é modelada por uma distribuição normal. O formato da distribuição de um conjunto de dados pode ser, dessa forma, um dos objetivos da análise estatística; por isso então, a necessidade de certificar se ocorre a normalidade da distribuição ou não. Para avaliar a normalidade dos dados de NDVI (valores absolutos e as médias mensais), foram realizados os testes Anderson-Darling, Shapiro-Wilk e Kolmogorov-Smirnov.

O teste de Anderson-Darling foi proposto em 1952, este teste é usualmente utilizado quando o tamanho da amostra é inferior a 25. Anderson Darling é um teste estatístico que pertence à classe quadrática de estatísticas baseadas na função de distribuição empírica (FDE), pois trabalha com as diferenças quadráticas entre a distribuição empírica e a hipotética.

Já o teste de Kolmogorov-Smirnov pertence à classe de estatísticas baseadas na FDE, pois trabalha com a maior diferença entre a distribuição empírica e a hipotética.

O teste de normalidade Shapiro-Wilk (1965, segundo os autores), é um procedimento estatístico aplicado para testar a normalidade de uma amostra completa. O teste é obtido através da divisão da raiz de uma combinação linear apropriada da ordem estatística da amostra, pela simetria estimada de variação. O teste apresenta vantagens como: poder ser utilizado com um teste de hipóteses compostas, ser fácil de calcular, exigindo apenas a tabela de coeficientes lineares e, por fim, é bastante sensível e utilizável em uma gama de alternativas, até mesmo amostras menores (SHAPIRO e WILK, 1965). Contudo os autores destacam que seu uso é entendido como um suplemento dos testes de probabilidade normal, e não como um substituto.

Para realização dos testes foi utilizado o software Minitab 1.6.

Teste de Wilcoxon

Após o cálculo dos testes e verificada a não normalidade dos dados, aplicou-se o teste de hipótese de igualdade para dados pareados Wilcoxon (teste não paramétrico), para analisar se a utilização da média mensal do NDVI proposta era viável. Segundo Wilcoxon (1945), a comparação entre dois tratamentos geralmente se distribui em duas categorias: dados para os dois tratamentos que não são pareados; ou dados pareados em que a comparação destes leva para uma série de diferenças, positivas ou negativas.

O autor destaca que a possibilidade de utilizar testes de ranking que fazem uso dos próprios dados numéricos do experimento, ao invés de um determinado valor atribuído, permite obter uma rápida ideia aproximada da significância das diferenças entre tratamentos.

O método compara dois tratamentos através do resultado da magnitude da diferença, considerando o sinal, positivo ou negativo. Para comparação de dados pareados, o ranking é estabelecido sem considerar o sinal, e posteriormente é atribuído o sinal para as diferenças negativas. O teste trabalha com a hipótese de que a amostra apresenta diferenças positivas e negativas distribuídas normalmente em zero (WILCOXON, 1945).

Para realização dos testes foi utilizado o software Minitab 1.6.

RESULTADOS E DISCUSSÃO

Para os testes estatísticos de normalidade foi calculado o p-valor dos valores absolutos (NDVI) e da média mensal (IM) para cada método: Anderson-Darling, Shapiro-Wilk e Kolmogorov-Smirnov. Para os três testes, todas as mesorregiões produtoras de cana-de-açúcar apresentaram p-valores inferiores ao nível de significância adotado (5%), ou seja, p-valor < 0,05 em todos os casos, como demonstra a Tabela 2.

Tabela 2. Teste de normalidade (nível de significância 5%).

Mesorregião	Teste de normalidade	Anderson-Darling	Shapiro-Wilk	Komolgorov
Assis	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Araçatuba	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Araraquara	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Bauru	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Campinas	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Itapetininga	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
MMP	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Marília	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Piracicaba	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Presidente Prudente	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
Ribeirão Preto	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01
São José do Rio Preto	p-valor NDVI	<0,005	<0,01	<0,01
	p-valor Media IM	<0,005	<0,01	<0,01

Com o resultado de p-valores menores que a significância estabelecida para todos os três métodos, rejeita-se a hipótese de normalidade, ou seja, os dados não possuem distribuição normal. Então, para avaliar a igualdade dos dados, utiliza-se o teste não-paramétrico de Wilcoxon (Tabela 3), pois não exige distribuição normal dos dados.

Tabela 3. Teste de Wilcoxon (nível de significância 5%).

Mesorregião	Teste de Igualdade	Wilcoxon
Assis	p-valor NDVI vs Média IM	0,916
Araçatuba	p-valor NDVI vs Média IM	0,882
Araraquara	p-valor NDVI vs Média IM	0,998
Bauru	p-valor NDVI vs Média IM	0,927
Campinas	p-valor NDVI vs Média IM	0,752
Itapetininga	p-valor NDVI vs Média IM	0,778
MMP	p-valor NDVI vs Média IM	0,648
Marília	p-valor NDVI vs Média IM	0,988
Piracicaba	p-valor NDVI vs Média IM	0,786
Presidente Prudente	p-valor NDVI vs Média IM	0,904
Ribeirão Preto	p-valor NDVI vs Média IM	0,784
São José do Rio Preto	p-valor NDVI vs Média IM	0,832

Para o teste de Wilcoxon, as médias mensais de NDVI quando comparada com os valores absolutos de NDVI das mesorregiões analisadas apresentaram p-valores maiores que 0,5, nível de significância estabelecido. Dessa forma, não se rejeita a hipótese de que os dados têm igualdade, e que a média mensal IM pode ser utilizada para o cálculo do NDVI da cultura de cana-de-açúcar no estado de São Paulo.

CONCLUSÕES

Como resultado, foi observado que, com 5% de significância, é possível utilizar a média mensal das imagens NDVI para indicar o desenvolvimento da cultura da cana no estado de São Paulo ao invés de utilizar os valores absolutos de NDVI.

Essa média mensal permite igualar as escalas de precipitação e índice de vegetação por diferença normalizada mensais, além de fornecer valores de NDVI com menos oscilações mensais.

REFERÊNCIAS

ANDERSON, T. W.; DARLING, D. A. *Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes*. Annals of Mathematical Statistics, 23, 193–212. 1952.

HUETE, A.; DIDAN, K.; MIURA, T.; RODRIGUEZ, E.; GAO, X.; FERREIRA, L. G. *Overview of the radiometric and biophysical performance of the MODIS vegetation index*. Remote Sensing of Environ, vol. 83, p.195-213. 2002.

Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA). *Zoneamento agroecológico da cana-de-açúcar* / organização Celso Vainer Manzatto ... [et al.]. — Rio de Janeiro: Embrapa Solos, 2009. 55 p. (Documentos/Embrapa Solos, ISSN 1517-2627 ; 110).

Instituto Brasileiro de Geografia e Estatística (IBGE). *Produção Agrícola Municipal (PAM)*.

Disponível em: <<http://www.sidra.ibge.gov.br/>>

Acesso em: [30/04/2015].

GUNNULA, W.; KOSITTRAKUN, M.; RIGHETTI, T. L.; WEERATHAWORN, P.; PRABPAN, M. *Normalized difference vegetation index relationships with rainfall patterns and yield in small plantings of rain-fed sugarcane*. Australian Journal of Crop Science, December 2011, Vol.5 (13), pp.1852-1857

MAPA (Ministério da Agricultura, Pecuária e Abastecimento), 2015. *Cana-de-açúcar*.

Disponível em: <<http://www.agricultura.gov.br/vegetal/culturas/cana-de-acucar>>

Acessado em: [15/04/2015]

NASA¹, 2015. *About Terra*.

Disponível em: <http://www.nasa.gov/mission_pages/terra/index.html>

Acessado em: [16/04/2015]

NASA², 2015. *Aqua Earth - observing satellite mission*.

Disponível em: <<http://aqua.nasa.gov/>>

Acessado em: [27/05/2015]

SHAPIRO, S. S.; WILK, M. B. *An analysis of variance test for normality (complete samples)*. Biometrika, 1 December 1965, Vol.52 (3/4), pp.591-611. Archival Journals (JSTOR).

TUCKER, C. J. *Red and Photographic Infrared Linear Combinations for Monitoring Vegetation*. Remote Sensing of Environment, 1979, Vol.8 (2), pp.127-150.

WEIER, J.; HERRING, D. *Measuring Vegetation (NDVI & EVI)*. NASA, 2010.

Disponível em :<http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_1.php>

Acessado em: [16/04/2015]

WILCOXON, F. *Individual Comparisons by Ranking Methods*. Biometrics Bulletin, 1 December 1945, Vol.1 (6), pp.80-83. Archival Journals (JSTOR).