The use of spatial-temporal analysis for noise reduction in MODIS NDVI time series data

Julio Cesar de Oliveira[1,2], José Carlos Neves Epiphanio[1], Camilo Daleles Rennó[1]

[1]Instituto Nacional de Pesquisas Espaciais (INPE), Divisão de Sensoriamento Remoto, Av. dos Astronautas, 1758, Caixa Postal 515, 12245-970, São José dos Campos/SP, Brazil. oliveirajc@ufv.br, epiphanio@dsr.inpe.br, camilo@dpi.inpe.br

[2] Universidade Federal de Viçosa (UFV), Departamento de Engenharia Civil – Setor de Engenharia de Agrimensura, Campus UFV, 36570-000, Viçosa, MG, Brazil.

Abstract

Time series of satellite data can be employed for mapping the development of vegetation in space and time. However, noise induced by cloud contamination and atmospheric variability affects data quality. Science Datasets is an integral part of the MODIS Land production chain that focuses on evaluating and documenting the scientific quality of products. This study aims at the reconstruction of time series of MODIS NDVI data based on the reliability of the science data sets and on a spatial-temporal analysis of the low quality pixels. The MOD13Q1 product was analyzed over a period of one year. After identifying the pixel with the lowest guarantee of quality, it is estimated by regression analysis among neighboring pixels classified as high-quality. The combination of the per-pixel quality and spatial-temporal information is a promising method for reconstructing high-quality MODIS NDVI time series.

Keywords: Time Series, MODIS NDVI, Reliability, Spatial-temporal analysis.

1. Introduction

Many researchers have demonstrated the usefulness of time series MODIS vegetation index data for monitoring inter-annual vegetation changes, deriving biophysical land surface variables and modeling terrestrial ecosystems (Sakamoto et al., 2010).

However, time series of the vegetation index may contain disturbances caused by cloud cover, seasonal snow, atmospheric variability, bi-directional effects and instrument problems (Gu *et al.*, 2009). This noise may give false indications of change and modify the true nature of a temporal profile. Thus, it must be minimized before using the data for any research (Huete *et al.*, 1999).

In literature, there is a wide variety of strategies to reduce noise and construct high-quality NDVI time series data sets. These strategies consist basically of filters, which can be grouped into two types: noise reduction in the frequency domain and noise reduction in the temporal domain. Their efficiencies are conditioned by the

analyst experience to define window sizes, degrees of polynomials and thresholds for elimination of noise.

In this research, we propose a methodology for noise reduction in NDVI time series based on spatial and temporal correlation of the pixel, combined with science datasets from Terra's Moderate Resolution Imaging Spectroradiometer.

The science datasets contain the quality indicators for each grid cell, and focus on evaluating and documenting of the scientific quality of product performance. Thus, the result of the quality analysis over science datasets indicates if the particular pixel of the input dataset is valid.

The principle of contextual analysis assumes that the nearer pixels are more related each other than distant pixels, so that a pixel with low reliability can be recalculated based on its neighbors.

2. Quality Assessment

The MODIS Land (MODLAND) Science Team (ST) has developed protocols to evaluate the performance of the MODLAND products through quality assessment and validation activities (Roy *et al.*, 2002). This quality information, named Quality Assessment (QA), is an integral part of the MODIS Land production chain and its results are formally stored as product metadata and as per-pixel information.

The QA values provide for users a database about atmospheric conditions (presence of cloud, ozone, dust, and other aerosols), sun-sensor-surface viewing geometries and data processing algorithms. Roy et al. (2002) emphasize that the QA values available in MODIS product contain information that contribute to a correct interpretation of the values present in a NDVI time series.

The MODIS VI product suite contains a set of output parameters known as science datasets (SDS) for each composited or selected pixel, which include the following: (I) composited NDVI and EVI, (II) NDVI/EVI output quality, (III) reflectances in band 1, 2, 3, and 7 (620–670 nm, 841–876 nm, 459–479 nm and 2,105– 2,155 nm, respectively), (IV) solar zenith angle, sensor zenith view angle, and relative azimuth, (V) day of year, and (VI) pixel reliability (Huete *et al.*, 2011).

Pixel reliability is useful in post processing analysis and summarizes the QA status of the product. This parameter is a simple decimal number that ranks the product into five categories (NoData – Good – Marginal – Snow/Ice – Cloudy), where users can consult this information instead of working with the QA layer (Didan and Huete, 2006).

The need to analyze the reliability pixel values together with the temporal profile of vegetation index is illustrated in Figure 1, where 2-neighbor pixels present different reliability and NDVI values (from the MOD13Q1 - MODIS/Terra Vegetation Indices - 16-Day - 250m) for the same composite date, land use and land cover.

The difference in NDVI value among neighboring pixels can be related only to the vegetation spectral response, but the QA values indicate that the conditions for acquisition and processing of the data were not the same.

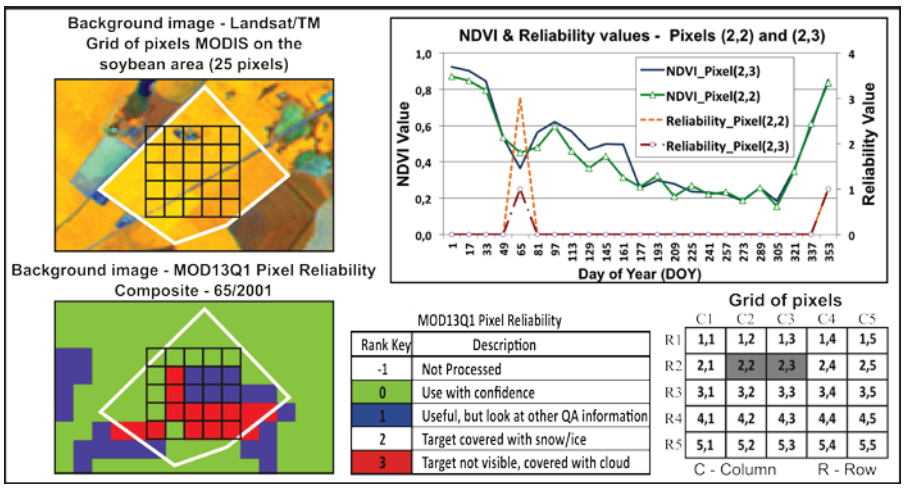

Figure 1: NDVI and Reliability Temporal Profile for 2-neighbors pixels.

3. Methodology

We used 23 images of the product MOD13Q1 over 2001 (Table 1), and selected a subset with 25 pixels as an example. An algorithm R was made to analyze in the dataset and to reconstruct the time series.

Table 1: MOD13Q1 product time series analyzed.

Product	MOD13Q1 (MODIS/Terra Vegetation Indices 16-Day 250m) tile h13v11
Dates (Composite of 16 days over 2001)	001; 017; 033; 049; 065; 081; 097; 113; 129; 145; 161; 177; 193; 209; 225; 241; 257; 273; 289; 305; 321; 337; 353

Figure 2 shows the science dataset (SDS) values of the product MOD13Q1 (composite date 65) for the 25 pixels selected in this study.

The first step of the methodology is based on an analysis of the reliability values available in each image of the product MOD13Q1, where the goal will be the identification of the pixel (R, C)i (R-row, C-column, for date i) with the lowest guarantee of quality. The value of NDVI for the pixel identified with low quality on date i (pixel(R, C)i observed) is recalculated after a linear regression analysis in time for each neighbor pixel. The best regression (or highest R^2) among 8-neighbor pixels is used to estimate the new value of NDVI (pixel(R,C)i estimated). The values of NDVI identified in time with low quality are not used in the regression analysis. In other words, regression analysis is performed only with dates in which the pixel is classified with good reliability in SDS.

	C1	C2	C3	C4	C5
	NDVI				
R1	0,34	0,34	0,35	0,54	0,41
R2	0,31	0,45	0,37	0,35	0,35
R3	0,23	0,45	0,44	0,43	0,35
R4	0,20	0,32	0,32	0,51	0,53
R5	0,43	0,26	0,51	0,51	0,59

	C1	C2	C3	C4	C5
	Julian Day of Year				
R1	80	80	80	80	80
R2	80	76	73	73	73
R3	80	76	73	73	73
R4	80	73	74	76	73
R5	73	80	76	76	73

	Sun Zenith Angle				
R1	30,31	30,31	30,31	30,31	30,31
R2	30,32	25,21	30,01	30,01	30,01
R3	30,32	25,21	30,01	30,01	30,01
R4	30,32	30,02	22,82	25,21	30,01
R5	30,02	30,32	25,21	25,21	30,01

	View Zenith Angle				
R1	18,59	18,59	18,68	18,68	18,68
R2	18,59	55,39	4,62	4,62	4,62
R3	18,59	55,39	4,62	4,62	4,62
R4	18,59	4,53	65,31	55,39	4,62
R5	4,53	18,69	55,39	55,39	4,62

	Reliability				
R1	0	0	0	0	0
R2	0	3	1	1	1
R3	0	3	1	1	1
R4	0	3	3	3	3
R5	3	0	3	3	3

	Quality Assessment (QA)				
R1	2116	2116	2116	2116	2116
R2	2116	3158	34837	34837	34837
R3	2116	3158	34837	34837	34837
R4	2116	36070	34842	3158	35930
R5	36070	2116	3158	3158	35930

C - Column
R -Row

MOD13Q1 Vegetation Index Usefulness

Value	Description
2116	High Quality
3158	Decreasing quality (mixed clouds)
34837	Intermediate quality
34842	Intermediate quality
35930	Below intermediate quality (mixed clouds)
36070	Below intermediate quality (mixed clouds)

Figure 2: Values of NDVI, DOY, VZA, SZA, Reliability and QA for pixels analyzed.

4. Results

Figure 3 shows the reliability pixel values for images of 2001, and the pixels with value equal to 0 (zero) were considered as the highest guarantee of quality (Didan and Huete, 2006). The application of this rule to dataset allowed the identification of pixels with low quality in the images DOY065, DOY273 and DOY353. Particularly, the pixel (2,2) DOY065 was selected to demonstrate the results of this study.

Linear regression analysis was performed only for 5-neighboring pixels (2,2), since the other neighbors have low or intermediate quality pixels on DOY065. The best regression analysis was obtained between NDVI values of the pixel (2,2) and pixel (3,1) from 21 images MOD13Q1 (all images, except composite date 065 and 353, which have low quality assessment).

Figure 4 presents the estimated and observed NDVI values for the pixel (2,2) DOY065. The values presented for the dates subsequent or previous to DOY065 refer to NDVI values observed in the image, because these dates have good reliability.

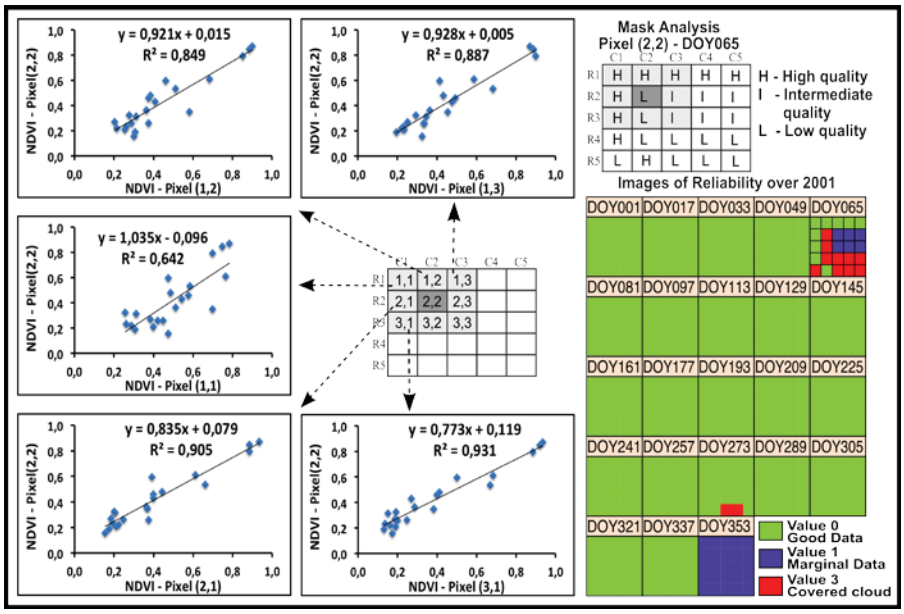

Figure 3: Regression analysis for Pixel (2,2) – DOY065.

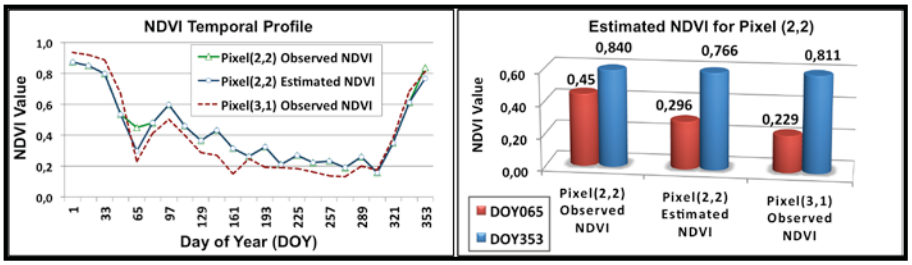

Figure 4: NDVI Temporal Profile and Estimated NDVI values for Pixel (2,2) DOY065.

5. Conclusion

There is a debate on NDVI reconstruction methods, but few studies focus on MODIS NDVI product and the use of ancillary QA data sets (Gu *et al.*, 2009). However, it is highly recommended to evaluate the quality indicators of each dataset, mainly for science quality time series, where this information is product specific and is available for each composite in a separate layer (Justice *et al.*, 2002; Udelhoven, 2011). Roy *et al.* (2002) propose the users of the MODIS products to inspect the science quality metadata, so that the product labeled with the lowest guarantee of quality should be used with caution.

Our work proposes a method for time series data smoothing that uses ancillary QA data sets and spatial and temporal analysis for noise reduction. The preliminary results indicate good potential of the methodology to reduce noise in time MODIS NDVI series data.

Acknowledgments

The authors are thankful to the Graduate Program in Remote Sensing of the National Institute for Space Research (PGSER/INPE) for the financial support for this symposium. They also thank the Federal University of Viçosa and INPE for the support for this research.

References

Didan, K., Huete, A. (2006). *MODIS Vegetation Index Product Series Collection 5 Change Summary.* pp. 17. Retrieved from http://landweb.nascom.nasa.gov/QA_WWW/forPage/MOD13_VI_C5_Changes_Document_06_28_06.pdf.

Gu, J., Li, X., Huang, C., Okin, G. S. (2009). "A simplified data assimilation method for reconstructing time-series MODIS NDVI data". *Advances in Space Research*. Vol. 44 (4): 501-509.

Huete, A., Justice, C. O., van Leeuwen, W. J. D. (1999). *MODIS Vegetation Index (MOD13): Algorithm Theoretical Basis Document*. pp. 120, Retrieved from http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf.

Huete, A., Didan, K., van Leeuwen, W., Miura, T., Glenn, E. (2011). MODIS Vegetation Indices. *In*: Ramachandran, B., Justice, C. O., Abrams, M. J. (Eds.). *Land Remote Sensing and Global Environmental Change: NASA's Earth Observing System and the Science of ASTER and MODIS*. Springer, pp. 579-602.

Justice, C.O., Townshend, J.R.G., Vermote, E.F., Masuoka, E., Wolfe, R.E., Saleous, N., Roy, D.P., Morisette, J.T. (2002). "An overview of MODIS Land data processing and product status". *Remote Sensing of Environment*, Vol. 83: 3-15.

Roy, D.P., Borak, J. S., Devadiga, S., Wolfe, R. E. (2002). "The MODIS land product quality assessment approach". *Remote Sensing of Environment*, Vol. 83: 62-76.

Sakamoto, T., Wardlow, B.D., Gitelson, A.A., Verma, S.B., Suyker, A.E., & Arkebauer, T.J. (2010). "A Two-Step Filtering approach for detecting maize and soybean phenology with time-series MODIS data". *Remote Sensing of Environment*, Vol. 10: 2146-2159.

Udelhoven, T. (2011). "TimeStats: A Software Tool for the Retrieval of Temporal Patterns From Global Satellite Archives". *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 4(2): 310-317.