próximo
artigo

Anais XIV Simpósio Brasileiro de Sensoriamento Remoto, Natal, Brasil, 25-30 abril 2009, INPE, p. 1465-1472.

Comparison of MODIS NDVI Time Series filtering by Wavelets and Fourier analysis to Generate Vegetation Signatures.

Thomaz Chaves de A. Oliveira [1,]
Luciano Teixeira de Oliveira [1]
Luis Marcelo Tavares de Carvalho [1]
Adriana Zanella Martinhago [1]
Sávio Gouvêa de Freitas [1]

[1] Departamento de Engenharia Florestal - Universidade Federal de Lavras - UFLA
Caixa Postal 3037 – 37.200-000 - Lavras - MG, Brasil
thomazchaves@gmail.com

Abstract. Temporal vegetation signatures (i. e., vegetation indices as functions of time) generated using the MODIS instrument poses many challenges, primarily due to signal to noise-related issues Bruce et al. (2006). This study investigates which methods best generate smoothed curves of vegetation signatures on MODIS NDVI time series. The filtering techniques compared were the HANTS algorithm, Verhoef (1996), which is based on Fourier analyses and Wavelet temporal algorithm which uses the wavelet analysis to generate the smoothed curves. The smoothed data were used as input data vectors for vegetation classification by means of Artificial neural networks. Statistics of the classifications reveal that the Wavelet filtering algorithm outperforms the original time series and the HANTS fft derived algorithms in all cases in all the classification algorithms.

Keywords: remote sensing, image processing, time series, wavelets analysis, NDVI, MODIS, Fourier.

1. Introduction

Since the Stockholm convention in 1972, the humanity has gained consciousness that men's actions can promote high alterations on the Earth's climate, social-economics and environment. Back then, the uncertainty of the magnitude of the changes as well as the unacknowledgment of the Earth's natural systems that oppose change caused polemics and uncertainty. This lead to the creation of the EOS (Earth Observing System), leaded by NASA, which has the objective of studying the Earth's global changes, its processes and to promote its continuous observation, their sensors were designed to operate for a long period of time. TERRA was the name given to the first platform launched by the EOS, by which marked the development of science of remote sensing by incorporating various sensors that collect diverse data types. The MODIS (Moderate-resolution Imaging Spectroradiometer) is the most important sensor aboard the TERRA platform. Its concept has its origin in various predecessors by which the most important is the AVHRR (Advanced Very High Resolution Radiometer) aboard the NOAA (National Oceanic and Atmospheric Administration), from 1978 untill 1986 Soares et al. (2007).

The AVHRR sensor was originally designed for meteorological applications, and has only two spectral bands (red and near-infrared) that can be used to generate the spectral indices of vegetation in vigor. The new generation MODIS sensor has a number of advantages over AVHRR, including more spectral bands that can be used for vegetation analysis YU et al. (2004).

MODIS Vegetation Indices products are appropriate for vegetation dynamics studies and characterization. MODIS-VI are found to be sensitive to multi-temporal (seasonal) vegetation variations and to be correlated with LAI across a range of canopy structure types species and lifeforms, land cover variations. The MODIS NVDI and EVI VI demonstrate a good range and sensitivity for monitoring and assessing spatial and temporal variations in vegetation amount and condition. The seasonal profiles outperform in sensitivity and fidelity the equivalent AVHRR-NDVI profiles particularly in atmosphere with water vapor contents. Huete et al. (2002).

2-Methods

This work's objective is to compare methodologies used to generate vegetation signatures from MODIS NDVI time series. The two methodologies in comparison are wavelets analysis and Fourier Analysis. They both generate smoothed out curves of vegetation signature.

2.1 Vegetation Signatures

According to Zhang et al. (2003), field-based ecological studies have demonstrated that vegetation phenology tends to follow relatively well defined temporal patterns. For example, in deciduous vegetation and many crops, leaf emergence tends to be followed by a period of rapid growth, followed by a relatively stable period of maximum leaf area. Different types of vegetation have different temporal growth patterns (i.e., different growth and senescence rates) Bruce et al. (2006).

Vegetation dynamics indicate important short and long-term ecological process. Continuous temporal observations of land surface parameters using satellite reveal seasonal and inter-annual developments. Vegetation indices have been extensively applied to characterize the state and dynamics of vegetation, in particular multiple NDVI datasets of the Advanced Very High Resolution Radiometer (AVHRR) instrument used during the last 25 years. Coldiz et al. (2007) Jensen (2000) Different vegetation types exhibit distinctive seasonal patterns on NDVI variation Yu et al. (2004).

A vegetation Index should maximize sensitivity to plant biophysical parameters; normalize or model external such as sun angle, viewing angle, and the atmosphere for consistent spatial and temporal comparisons; normalize internal effects such as canopy background variations; a vegetation index may preferably coupled with a measurable biophysical parameter such as biomes, LAI, or APAR. Jensen (2000) MODIS Vegetation Index products correlate and respond positively to increases in PAI and LAI across a range of canopy structure types and species life forms. Huete et al. (2002)

Spatial and temporal variability in VIs arise from several vegetation related properties, including LAI, canopy structure/architecture, species composition, land cover type, leaf optics, canopy crown cover, understory vegetation, and green leaf biomass Huete et al. (2002).

2.2 Study Site

Due to the widespread occurrence of deciduous forests in the state of Minas Gerais and due to the state's large extent, 586.528 km², this study was conducted in a smaller area of interest, near the towns of Matias Cardoso and Manga, that has following coordinates: (14º40'37,06" S , 44º02'41,28"W),(14º51'59,33"S, 43º47'13,16"W). The Datum that was used was WGS-84, and the projection was Albers Equal Area. This area was primarily chosen because of acknowledged occurrence of deciduous forest. The area was also chosen so that the chosen methodology could be spreaded to other areas in future work. The NDVI time series was derived from the MOD13 product, which has a spatial resolution of 250m, and has a 16-day compositing period. The original data, downloaded from the MODIS FTP was reprojected using MRT(MODIS Reproduction Tool) A set of temporal images from the Landsat TM sensor were also used as auxiliary data. These images were collected from summer and winter dates to each area and were used so the exact locations of the deciduous forests and other types of vegetation were know previously and methodology validation. The original data set was resampled to 12 values to each year. The database included the years of 2003, 2004 and 2005.

This work utilizes the MOD13 NDVI 16 day composite product available for download from the MODIS website via FTP protocol. The original downloaded HDF files were reprojected using the MRT (Modis Reprojection Tool).

To extract pertinent features from time signatures for potential target applications, the signals must first be denoised. Authors have investigated automated methods for denoising, including straightforward methods such as median filterg and moving-average filtering, as well as more advanced methods such as wavelet denoising Bruce et al. (2006). Curve fitting parameterization using logistic functions have also succeeded in generating time signatures of MODIS VI Zahng et al. (2003).

2.3 Fourier Analysis:

The Fourier Analyses or Harmonic analyses have been used traditionally to solve some differential equations and also some partial equations in Math and Physics. Its main objective is to approximate a function in time domain by a function, which contains a linear combination of harmonics (sinusoids) Morettin and Toloi (2006). The most basic property of the sinusoids that makes them suitable for the analysis of time series is their simple behavior under a change in time scale Bloomfield (1976). Equation -1 Fourier transform to the frequency domain, Equation 2 - Inverse Fourier transform.

Fourier analysis have traditionally used for denoising and curve fitting in MODIS VI data sets Colditz et al. (2007) Bruce et al. (2006) Yu et al. (2004) Wang et al. (2004). If the original data is discrete rather than continuous, the discrete rather than continuous, the discrete Fourier transform (DFT), which requires regular spacing on samples within the temporal domains, should be applied Wang et al. (2004).

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-2j\pi ft} dt \tag{1}$$

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{2j\pi ft} dt \tag{2}$$

There is a drawback in this approach, since the NDVI images are composite images of different dates. The pixels have different acquiring dates that lead unequal time spacing. However, the next generation Hants algorithm Harmonic Analysis of Time Series (HANTS) was developed to deal with time series of irregularly spaced observations and to identify and remove cloud contaminated observations Verhuef (1996), Roerink et al. (2000).

This algorithm considers only the most significantcant frequencies expected to be present in the time profiles (determined, for instance, from a preceding FFT analysis), and applies a least squares curve fitting procedure based on harmonic components (sines and cosines). Verhoef (1996), Roerink, et al. (2000). For each frequency the amplitude and phase of the cosine function is determined during an iterative procedure. Input data points that have a large positive or negative deviation from the current curve are removed by assigning a weight of zero to them. After recalculation of the coefficients on the basis of the remaining points, the procedure is repeated until the maximum error is acceptable or the number of remaining points has become too small Roerink et al. (2000).

Many different phenological indicators have been defined in various satellite-based studies. The advantage of the HANTS algorithm is that the output consists of a completely smoothed NDVI profile which is convenient for calculating derivatives. (DE WIT 2005) The calculations of derivates are very important so you can estimate the start of growing season and senescence dates Sakamoto et al. (2005).

The version of HANTS used was implemented in IDL by DE WIT, 2005 and is under the GNU General Public License.

The data set of temporal NDVi images, which contained a 36 monthly sampled time series was subdued as input to the HANTS algorithm. The output is a similar 36 month time series but smoothed and containing only the annual, 6 months and 3 months frequencies of the signal.

2.4 Wavelet transform:

Fourier series are ideal for analyzing periodic signals, since harmonics modes used in the expansions are themselves periodic. By contrast the Fourier integral transform is far less natural tool because it uses periodic functions to expand nonperiodic signals. Two possible substitutes are the windowed Fourier transform (WFT) and the wavelet transform. The windowed Fourier transform can however be an inefficient tool to analyzing regular time behavior that is either very rapid or very slow relative to T (size of the analyzing window). The Wavelet transform solves both of these problems by replacing modulation with scaling to achieve frequency localization. WFT can be a very inefficient tool when very short time intervals are of interest and a similar situation occurs when very long and smooth features of the signal are to be reproduced by the WFT. Kaiser (1994).

The function $f(x)$ is transformed in the wavelet transform as follows:

$$Wf(a,b) = \int_{-\infty}^{+\infty} f(x) \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) dx \tag{2}$$

Where a is a scaling parameter, b is a shift parameter and ψ implies a mother wavelet. In discrete form, parameters a and b are given as follows:

$(a,b) = (2^j, 2^j k)$, where ***j*** and ***k*** are integers.

A wavelet is a small, localized in time or space and here satisfies the orthogonal condition. Since a wavelet has compact support, which means that its value becomes 0 outside a certain interval of time, the time components of time-series can be maintained during the wavelet transformation. Sakamoto et al. (2005)

Wavelets transforms are quite complex subjects and its way beyond the scope of this work to try and explain its full theory. It would riquerie whole book library on the subject to cover all the theory and its application on digital signals.

Previous works reveal that the wavelet transform is a powerful tool for denoising data sets and for curve fitting procedures in NDVI time series Sakamoto et al. (2005), Galford et al. (2007), Bruce et al. (2006).

For the present work, we used the methodology by proposed by Carvalho (2001).

In remote sensing outliers caused by clouds and shadows (noise) appear as peaks if narrow bandwidth in temporal spectrum. They appear similarly in the spatial domain, but with variable bandwidth. If we consider the presence of clouds and shadows as signal response against a “noisy” background, a framework for their detection can forth based on noise modeling in transformed space. The discrete wavelet transform was implemented with the ‘à trois’ algorithm with a linear spline as the wavelet prototype. It produces a vector of wavelet coefficients *d* at each scale j, with j=0,……,J. The original function $f(t)$ was then expressed as the sum of all wavelets scales and the smoothed version a_j. The input signal was decomposed using one scale, two scales and three scales. The resulting different data sets were used as input data sets for image classification, described as followed.

$$f(t) = a_j(t) + \sum_{i=1}^{J} d_j(t) \tag{3}$$

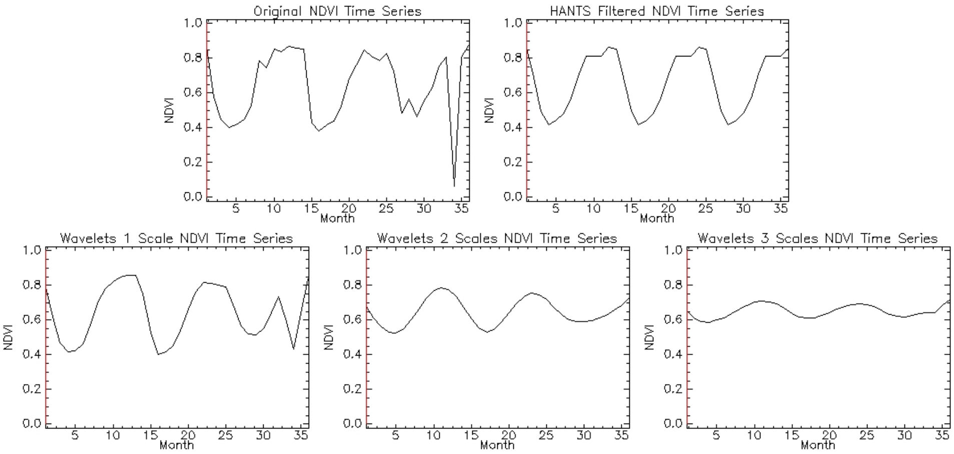

Figure 1 - Results of time series filtering

2.5 Image Classification

The main objective of this work is to compare different filtering techniques and their output vegetation signature (figure 1) for time series of NDVI. One way to accomplish this is to use the output smoothed time series as input vectors to automated image classification procedes. For Moreira (2003) an automatic image identification and classification can be understood as the analyses and the manipulation of images through computational techniques, with the goal of extracting information regarding an object of the real world.

2.6 Neural Networks

Humans and other animals process information with *neural networks*. These are formed from *trillions* of neurons (nerve cells) exchanging brief electrical pulses called action potentials. Computer algorithms that mimic these biological structures are formally called artificial neural networks to distinguish them from the squishy things inside of animals. Smith (1998). Even though some researchers do not recognize the neural networks as being the general natural solution surrounding the problems of recognizing patterns on processed signals, it can be noticed that a well trained network is capable of classifying highly complex data. The use of neural networks in pattern recognition and classification has grown in the last years in the field of remote sensing Kanellopolous (1997). For comparison reasons other two different types of networks were used. An unsupervised neural networks called Self Organizing Maps (SOM) Kohonen (1990), were used to classify the vegetation. Unsupervised learning does not need input samples for pattern recognition. A Fuzzy ArtMap is a clustering algorithm that operates on vectors with fuzzy analog input patterns (real numbers between 0 and 1) and incorporates an incremental learning approach which allows it to learn continuously without forgetting previous learned patterns. For Mather (1999) the use of soft (fuzzy) classification paradigms with neural networks is adequate when we want to avoid errors of classification due to ambiguity of the generated classes. This type of network considered in the present study was based on the ART (Adaptative Resonance Theory), which exhibits a high degree of stability in order to preserve significant past learning, but remain enough adaptable to incorporate new information whenever it might appear.

2.7. Training and Classification and Accuracy Assessment

The classification proceeded, using training samples, so that the algorithm can "learn" to recognize the patterns. The training phase must happen to each algorithm before it can be used for classification. The multi layer perceptron had the following characteristics which were obtained by testing: 17 layers on the hidden layer, learning rate of 0.01, momentum factor of 0.5, 20 training samples per category, 10 testing samples per category and liner activation function. The network was trained with 10000 iterations. It was used the same training data for both original and denoised time series. The SOM does not need training, however was parameterized as following: maximum learning rate of 1.0, minimum learning rate of 0.5, minimum gaining term of 0.0001, max gaining term of 0.0005, LVQ2 fine tuning rule, the algorithm was ran with 1000 epochs. The Fuzzy ARTMAP had the following parameters for ARTa: choice parameter of 0.01, learning rate of 0.5, vigilance parameter of 0.98. The ARTb had learning rate of 0.5 and vigilance parameter of 1.

In order to compare the results of classification, we need to parameterize the classification among different classifiers. In order to accomplish this, a set of accuracy samples was aquired. With the accuracy samples a confusion matrix was generated, by which the Kappa coefficient was extracted from. By doing this, we can compare statistically the performance of each algorithm and for this data set Mather (1999).

3. Results

Results (Figure 2) show promising results regarding the use of time series for vegetation characterization and classification. For all classification procedures, the wavelet analysis proved to be the best method for the generation of vegetation signatures for vegetation characterization. In all cases the accuracy of the classifications were superior to 70% in accuracy and range up to 90%. This can be noticed in Tables 1, 2 and 3. Table 4 depicts the confusion matrix of best qualification result, which was performed by the multi layer percecptron with wavelet analysis smoothed in one scale vector data set.

Table 1- Multi layer perceptron classification accuracy

Method	**Accuracy (%)**	**Kappa Coefficient**
Original times series	88,7500	0,8714
Wavelets 1 coefficient	**90,0000**	**0,8857**
Wavelets 2 coefficients	86,7674	0,8453
Wavelets 3 coefficients	76,2500	0,7286
Hants (Fourier)	82,5000	0,8000

Table 2- Self Organizing Maps classification accuracy

Method	**Accuracy (%)**	**Kappa Coefficient**
Original time series	79,1045	0,7566
Wavelets 1 coefficient	78,7500	0,7571
Wavelets 2 coefficients	**80,0000**	**0,7714**
Wavelets 3 coefficients	78,7500	0,7571
Hants (Fourier)	78,7500	0,7571

Table 3- Fuzzy ArtMap classification accuracy

Método	**Accuracy (%)**	**Kappa Coefficient**
Original times series	78,7500	0,7571
Wavelets 1 coefficient	77,5000	0,7429
Wavelets 2 coefficients	78,7500	0,7571
Wavelets 3 coefficients	**81,2500**	**0,7857**
Hants (Fourier)	73,7500	0,7000

Table 4 - Confusion matrix of best resulting classification – Wavelets with one scale classified by multi layer perceptron

Classes	Water	Alagadiço	plantation	Semi deciduous	Deciduous	Carrascal	Cerrado	Caatinga	Total
Water	7	0	0	0	0	0	0	0	**7**
Alagadiço	0	9	0	2	0	0	0	0	**11**
plantation	0	0	10	0	0	0	0	0	**10**
Semi deciduous	2	1	0	6	0	0	0	0	**9**
Deciduous	0	0	0	0	10	0	0	0	**10**
Carrascal	0	0	0	2	0	10	0	0	**12**
Cerrado	1	0	0	0	0	0	10	0	**11**
Caatinga	0	0	0	0	0	0	0	10	**10**
Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**80**

The multi layer perceptron performed better than the other algorithms giving a higher overall accuracy. The HANTS algorithm proved no improvements on the original time series of vegetation signatures. Theses results proved similar to Bruce et al (2006), that in all cases the Wavelet analysis at the most suited scale improve the original data set on accuracy but the same does not occur to the Fourier filtered data set, that performs worse that the original data set. Despite its widespread use over the years for time series filtering it has proved inefficient data used as input for vegetation classification. This can be due to fourier analysis not being resolved in time. By this, temporal information can be averaged into the whole of the time series resulting in incorrect vegetation profiles. Some incorrect samples can deteriorate both network training and classification leading to inferior results.

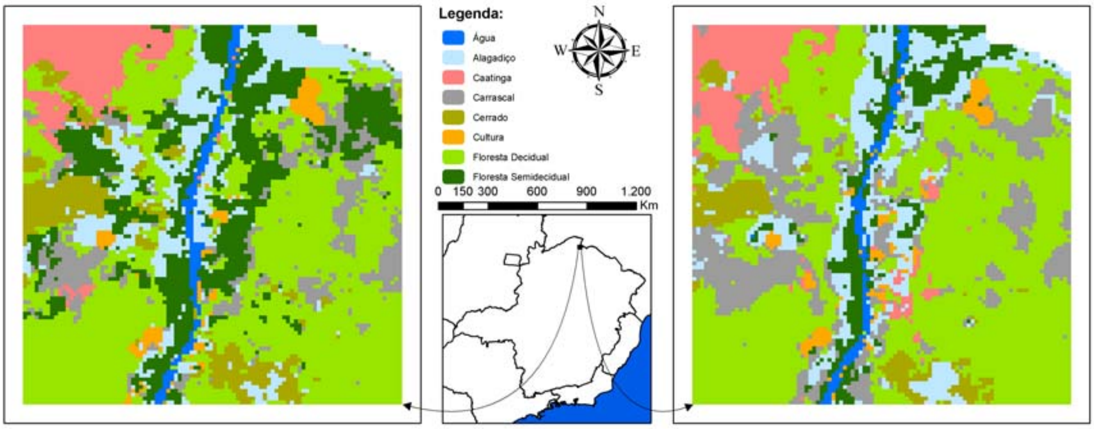

Figure 2 – Vegetation classification, wavelets filtered time series right, original time series left

4. Conclusion

The results show that forest classification can be well characterized by temporal MODIS NDVI time series. Because of the noisy original data set, filtering techniques such as FFT or Wavelet analysis can be used to remove cloud contamination and noise or shadows. These methods however can reduce or improve classification accuracy as proved in this work. For vegetation differentiation purposes, the wavelet analysis proved to improve the original data set and the FFT performed by HANTS proves the contrary, as found by Bruce et al (2006). The wavelet regression with the decomposition in one scale proved the best filtering technique with a 90% accuracy classification with the multi layer perceptron algorithm. Future research however must be done, with different wavelets, in different geographical areas and methodologies to further improve and generalize the results.

5. References

Soares, J. V.; Batista, G. T.; Shimabukuro, Y. E.; Histórico e Descrição in: Rudorff, B. F. T.; Shimabukuro, Y. E.; Ceballos, J. C. (Org.). **O sensor MODIS e suas aplicações no Brasil**. São José dos Campos: Editora Parêntese, 2007. cap. 1, p.2-22

Smith S. W. **The Scientist and Engineer's Guide to Digital Signal Processing**, copyright ©1997-1998 451-460, 1998

Kaiser, G. **A Friendly Guide to Wavelets** ISBN 0-8176-3711-7 Birkhäuser Boston, c/o Springer-Verlag New York, Inc.,175 Fith Avenue, New York, \ny 10010, USA, 1994

Carvalho, L. M. T. Mapping and monitoring forest remnants: a multi-scale analysis of spatio-temporal data. 2001. 140 p. Thesis (Doctor) - Wageningen University, Wageningen, 2001

Bloomfield, P. **Fourier Analysis of Time Series**. New York : John Wiley & Sons Inc. 257p., 1976

Bruce, L. M.; et al. Denoising and Wavelet-Based Feature Extraction of MODIS Multi-Temporal Vegetation Signatures. **GIScience & Remote Sensing**, 43. 170 – 180, 2006

Colditz, R.R. et al. Dynamics of MODIS time series for ecological applications in southern África. In Fourth International Workshop on the Analysis of Multitemporal Remote Sensing Images, Leuven, Belgium. (Multitemp 2007), 18. – 20. July, 2007

Roerink, G. J.; Menenti, M. ; Verhoef, W. 2000, Reconstructing Cloudfree NDVI composites using **Fourier analysis of time series International Journal of Remote Sensing** ISSN 0143-1161 print/ISSN 1366-5901 online © 2000 Taylor & Francis Ltd, vol. 21, no. 9, 1911–1917. , 2000

Huete, A.; et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**, Volume 83, Number 1, November. pp. 195-213(19), 2002

Jensen, J. R. Remote **Sensing of the enviroment: An Earth resource perspective**, Pretince Hall Series in geographic information science, Upper SaddleRiver, New Jersey 07458, 2000

Kanellopoulos G.G., Wilkinson F.Roli, J.Austin, **Neuro-computation in Remote Sensing Data Analysis,** 1997

Kohonen, T. The Self-Organizing Map. **Proceedings of the IEEE**, 78: 1464-80., 1990

Mather, P. M. **Computer Processing Of Remotely-Sensed Images: An Introduction**. 2. ed. Nottingham, UK: Johb Wiley, 1999, 292 p.

Moreira, M. A.; **Fundamentos de Sensoriamento Remoto e Metodologias de Aplicação**, 2ª Edição Revista e Ampliada Eidtora UFV 295p., 2003.

Morettin, P. A., Toloi, C. M., 2006 **Análise de Séries Temporais**, Segunda edição 2006 429-433 Editora Edgar Blücher LTDA. 2006 CDD-519.232

Sakamoto, T. ; Yokozawa, M. ; Toritani, H. ; Shibayama M. ; Ishitsuka N.; Ohno H. A Crop Phenology Detection Method Using Time-Series Modis Data; **Remote Sensing of Environment** 96 366 – 374, 2005

Verhoef, W. **Application of Harmonic Analysis of NDVI Time Series (HANTS). In Fourier Analysis of T emporal NDVI in the Southern African and American Continents**, edited by S. Azzali and M. Menenti, DLO Winand Staring Centre, Wageningen, TheNetherlands, Report 108, pp. 19–24, 1996

Wang Q. et al. Similarities in ground and satellite-based NDVI time series and their relationship to physiological activity of a Scots pine forest in Finland. **Remote sensing of environment**., vol. 93, nº1-2, pp. 225-237, 2004

Yu, Xinfang; et al. **Forest** Classification Based on MODIS Time Series and Vegetation Phenology. **International Geoscience And Remote Sensing Symposium**. VOL 4, pages 2369-2372, 2004.

Zhang, X.; Friedl, M.A.; Schaaf, C.B.; et al. **Monitoring Vegetation Phenology Using Modis**. **Remote Sensing of Environment** 84 471 – 475, 2003

Zhang, F. et al. An Advanced Tool for Real-Time Crop Monitoring in China. In IEEE International Geoscience and Remote Sensing Symposium. Toulouse, France, vol IV: 2242 – 2244. July 21-25, 2003.