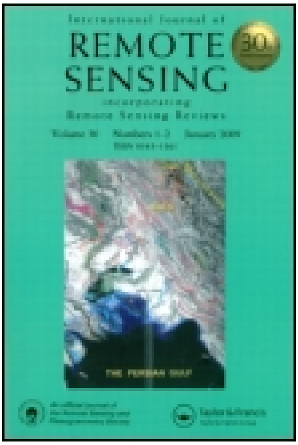

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

Reconstructing cloudfree NDVI composites using Fourier analysis of time series

G. J. Roerink , M. Menenti & W. Verhoef

Version of record first published: 25 Nov 2010.

To cite this article: G. J. Roerink , M. Menenti & W. Verhoef (2000): Reconstructing cloudfree NDVI composites using Fourier analysis of time series, International Journal of Remote Sensing, 21:9, 1911-1917

To link to this article: http://dx.doi.org/10.1080/014311600209814

INT. J. REMOTE SENSING, 2000, VOL. 21, NO. 9, 1911–1917

Reconstructing cloudfree NDVI composites using Fourier analysis of time series

G. J. ROERINK, M. MENENTI

DLO Winand Staring Centre for Integrated Land, Soil and Water Research (SC-DLO), PO Box 125, 6700 AC Wageningen, The Netherlands; e-mail: G.J.Roerink@sc.dlo.nl

and W. VERHOEF

National Aerospace Laboratory (NLR), PO Box 153, 8300 AD Emmeloord, The Netherlands

(*Received 13 January 1999; in final form 3 September 1999*)

Abstract. This letter describes the Harmonic ANalysis of Time Series (HANTS) algorithm. It performs two tasks: (i) screening and removal of cloud affected observations; and (ii) temporal interpolation of the remaining observations to reconstruct gapless images at a prescribed time. HANTS was applied to 36 AVHRR 10-days-maximum-NDVI composites covering most of Europe. The results show that cloud affected data are recognized successfully and replaced. Up to half the data points were rejected with no consequence for the successful reconstruction of seasonal NDVI profiles.

1. Introduction

For many remote sensing applications which focus on land surface properties, cloud cover has a disastrous effect on the results. A lot of time and effort has been put into the development of algorithms to detect and mask clouds (Henderson-Sellers 1982, Stowe *et al.* 1991, Simpson and Gobalt 1996). The result is always an image with data gaps.

This letter describes the implementation of an algorithm that accomplishes two tasks: (i) screening and removal of cloud contaminated observations; and (ii) temporal interpolation of the remaining observations to reconstruct gapless images at a prescribed time. The basic idea is to calculate a Fourier series to model a time series of pixelwise observations, while at the same time identifying outliers relative to the model of the time series. The algorithm eliminates such outliers and replaces them with the value given by the Fourier series.

In our previous papers (for example Menenti *et al.* (1993), Verhoef *et al.* (1996)) a fast Fourier transform (FFT) algorithm was used to perform the Fourier analysis (Fourier 1818, Main 1990). To use the FFT, observations must be relatively error (cloud) free and equidistant in time. The new 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 (Verhoef 1996).

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journals

2. HANTS

This algorithm considers only the most significant 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). 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.

The curve fitting process is controlled by five parameters, which have to be set at the beginning of each HANTS run.

(1) Number of frequencies (NOF). A curve is described by means of its average value (zero frequency) and a number of sine functions with different frequencies. By this control parameter the user defines how many frequencies are used and how large their corresponding period in time sample units is. This results in $2 \times \mathrm{NOF} - 1$ output parameters (an amplitude and phase value for each frequency), where the NOF includes the zero frequency (mean), which has no phase.
(2) Hi/Lo suppression flag (SF). This flag indicates whether high or low values (outliers) should be rejected during curve fitting.
(3) Invalid data rejection threshold (IDRT). In some cases one might know that digital numbers below or above a certain threshold should be considered invalid.
(4) Fit error tolerance (FET). During curve fitting the absolute difference in the Hi/Lo direction of the remaining (that is, not rejected) data points with respect to the current curve is determined after each iteration. Iteration stops when the difference of all remaining points becomes smaller than the FET. The FET value should not be set too low, as otherwise the fit might be based on too few points, which gives unreliable results.
(5) Degree of overdeterminedness (DOD). The number of valid observations must always be greater than or equal to the number of parameters that describe the curve ($2 \times \mathrm{NOF} - 1$). In order to get a more reliable fit the user can decide here to use more data points than the necessary minimum. The minimum number of extra data points, which have to be used in the ultimate fit, is given by the DOD value.

The HANTS algorithm offers greater flexibility in the choice of frequencies and the length of the time series than the FFT algorithm. Also it is relatively easy to exclude certain images from the time series, since the samples are not required to be equidistant in time. The price paid is a processing time of roughly ten times that of the FFT, but this is still acceptable (approximately ten hours for a series of 36 images of 1536×2800 pixels on a work station, if two frequencies are selected). However, thanks to the removal of obvious errors, the computed amplitudes and phases at the selected frequencies are much more reliable than with the direct FFT algorithm.

3. Results

The HANTS algorithm was applied on one year of calibration AVHRR 10-days-maximum-NDVI composites, 1996 (Koslowsky 1996). The covered area is from 10°

West till 20° East and from 29° till 55° North with a resolution of 0.01° (see also figure 3). The HANTS control parameters were set as below.

NOF = 3; the selected frequencies were the zero frequency (mean NDVI) and the frequencies with time periods of 1 year and 6 months. So the output comprises five Fourier coefficients (three amplitudes and two phases).
SF = low; for NDVI images this flag was set at low, as cloudy observations lead to low NDVI values.
IDRT = +0.7; this means that NDVI values higher than 0.7 are rejected. In the input data a few high outliers existed, which were excluded in this way.
FET = 0.05 NDVI units; the absolute error in negative direction of the remaining observations should be smaller than 0.05 NDVI units with respect to the currently fitted curve.
DOD = 13; together with the minimum of five observations this means that each fitted curve is based on a minimum of 18 observations in time, which is half of all data points.

One of the disadvantages of this algorithm is that there are no objective rules to determine the HANTS control parameters. This is done on the basis of experience (for example from a preceding FFT analysis) and after running several combinations of control parameters.

Figure 1(*a*) shows the original and HANTS reconstructed NDVI time series for a pixel (2.885° East, 50.235° North) in Northern France, which is classified as non-irrigated arable farming in the CORINE land cover database (CEC 1993). Iteration stopped when 14 out of the 36 original NDVI values were rejected, that is, classified as cloud affected data points. The remaining 22 values have a maximum negative deviation from the curve of 0.05 NDVI units (= FET). Figure 1(*b*) shows the Fourier components of the three different frequencies from which the cloudfree profile is reconstructed.

Figure 2 shows another NDVI profile of a pixel in England, where the presence of cloud cover throughout the year is high. Here the HANTS iteration has stopped by reaching the DOD. The absolute error of two of the last 18 data points used for the fit is still greater than 0.05 in the negative direction, however, the resulting curve fits rather well the higher, cloudfree NDVI observations. During wintertime two positive outliers (decades 1 and 10) exist; however, HANTS cannot reject outliers in the opposite direction of the SF. To demonstrate the difference between the results of the FFT and the HANTS algorithm, the FFT reconstructed NDVI profile is shown here as well. It is generated with the same Fourier components as the HANTS profile (frequency = 0, 1 and 2). The shapes of the two profiles are similar, but it is clear that the FFT values are much too low, because this fit is based on all data points.

Figure 3 represents the HANTS Fourier components. It is a colour composite of the mean NDVI in red, the amplitude of 1 year in green and the amplitude of 6 months in blue. Within the area covered a wide variety of combinations of Fourier components can be found, due to differences in land use and climatologic conditions. Areas without vegetation cover during the year are represented by black, such as the Saharan desert and the peaks of the Alps. Irrigated areas like the Po valley and certain areas of Spain have strong 12 and 6 months' components, which results in the azure blue colour. The difference between grasslands and arable farming is remarkable; while the grass covers the land during the whole year (high mean NDVI), the arable lands are harvested once a year leaving behind bare soil (high

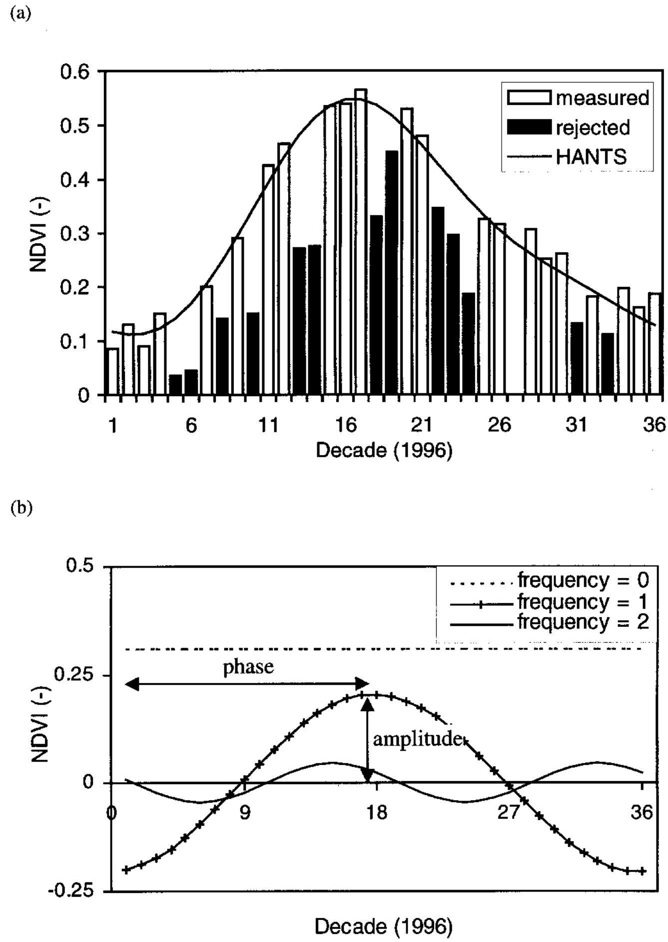

Figure 1. (*a*) Original cloud affected and reconstructed cloudfree NDVI time series of a location (arable farming; 2.885° West, 50.235° North) in Northern France and (*b*) the different Fourier components from which the cloudfree profile is reconstructed. The arrows represent those values for the annual component (frequency = 1).

annual amplitude). The Netherlands show this difference very clearly between the green polder (arable lands) and the reddish Friesland (grasslands).

In figures 1 and 2 the temporal NDVI behaviour under cloudfree conditions was reconstructed for one pixel. Using this technique for the whole area of figure 3, it becomes possible to reconstruct cloudfree NDVI images for any moment in the temporal sequence. Figure 4 illustrates this cloud removal procedure. Figure 4(*a*) is an image of the original NDVI 10-days-composite of decade 33, 1996 with clouds all over Europe. By using the three Fourier components, the cloudfree NDVI image is reconstructed in figure 4(*b*). All clouds disappeared and the natural land cover patterns become clear, while the original cloudfree areas in figure 4(*a*) maintain the same NDVI values within 0.05 NDVI units. To prove that the white areas (NDVI values between 0.01 and 0.05) in Europe in figure 4(*a*) are indeed clouds and not, for example, snow figure 4(*c*) is included. The brightness temperature map derived from AVHRR channel 4 shows that the white areas in figure 4(*a*) have temperatures

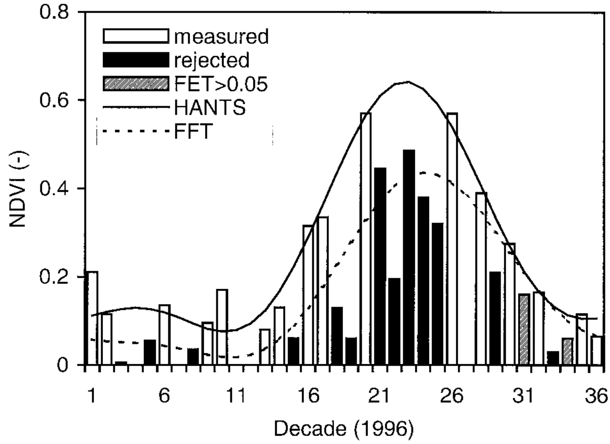

Figure 2. Measured NDVI time series and the Fourier reconstructed cloudfree time series of a pixel in England (1.845° West, 53.485° North).

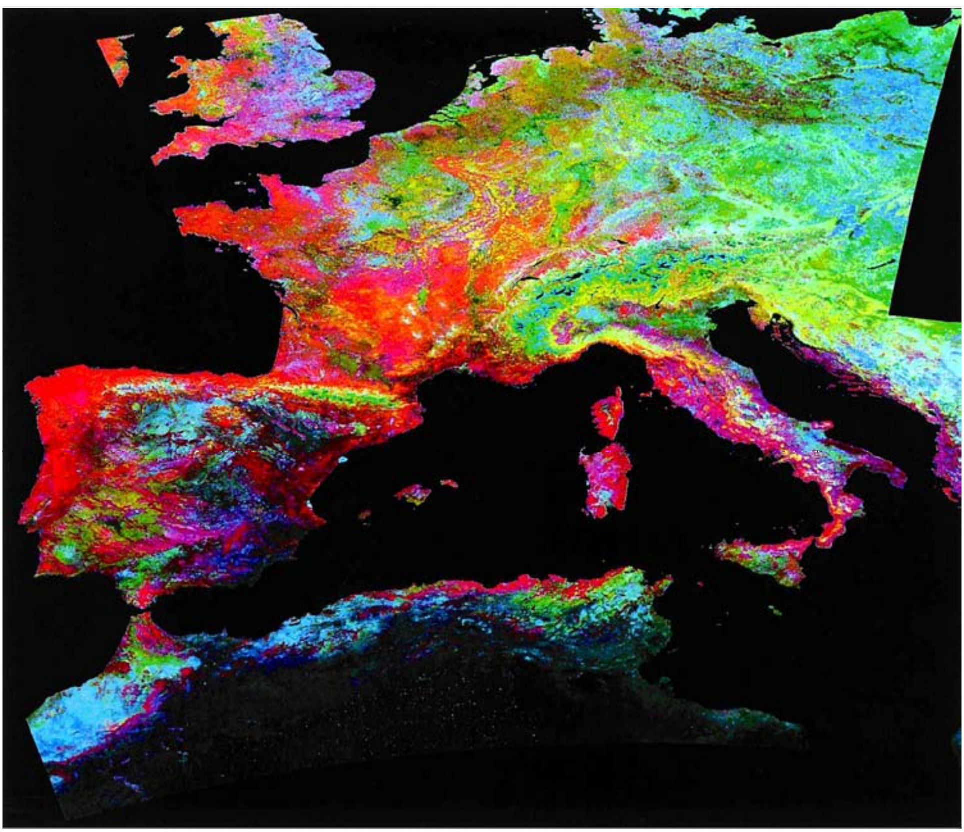

Figure 3. Map of HANTS Fourier components of NDVI of 1996: red, mean NDVI; green, amplitude 1 year; blue, amplitude 6 months.

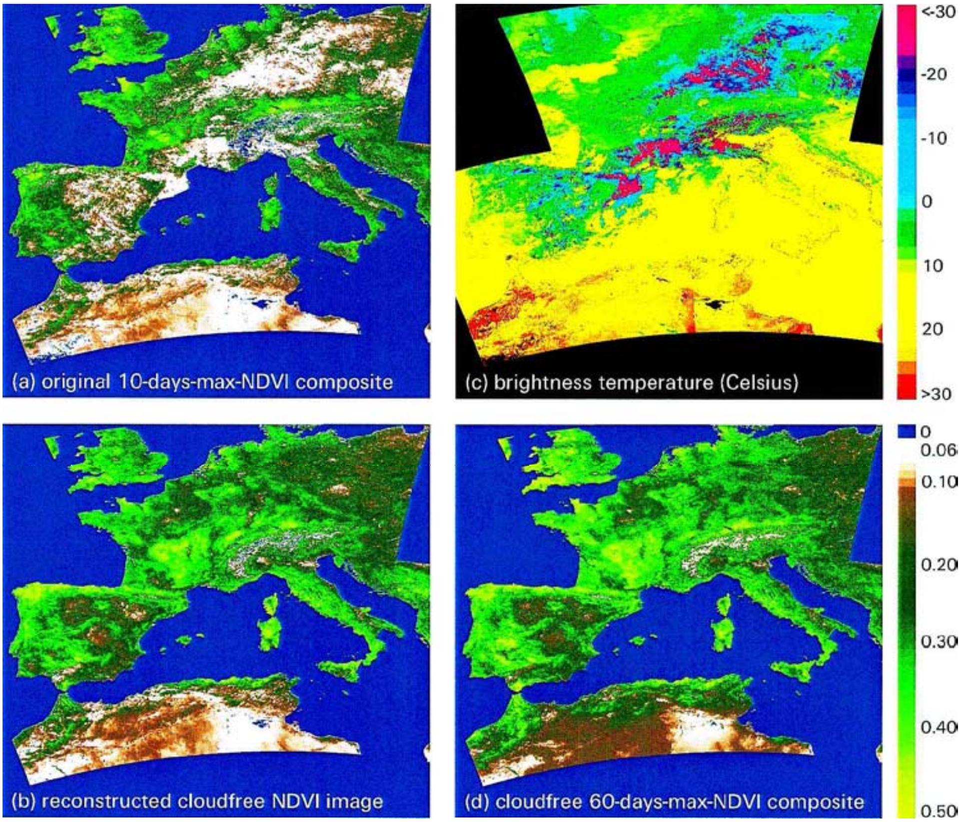

Figure 4. Maps of decade 33 (9–18 November), 1996 of (*a*) the original NDVI composite, (*b*) the reconstructed cloudfree NDVI, (*c*) the brightness temperature in AVHRR channel 4 and (*d*) the maximum NDVI composite of a 60 days period.

of − 20°C or less, while the rest of Europe has temperatures greater than 0°C. These low temperatures can be explained only by the presence of clouds. In addition figure 4(*d*) is shown here to prove that the reconstructed land cover patterns in figure 4(*b*) reflect the reality. Figure 4(*d*) is a maximum NDVI composite of the decades 33 and 34 of the years 1995, 1996 and 1997 and can be considered cloudfree. It is clear that the land cover patterns in both figures correspond very well; areas with (questionable) reconstructed low NDVI values, like Eastern Europe, the area south of Paris and the Umbro valley in Spain, are indeed areas with low vegetation cover. The NDVI values in figure 4(*d*) are in general slightly higher than in figure 4(*a*) and (*b*) due to the fact that the maximum NDVI value of a period of 60 days is most of the time higher than the maximum NDVI value of a 10 day period, which is part of the considered 60 days.

4. Conclusions

The new HANTS algorithm has proven to be a powerful tool not only to classify and climate cloud affected data in time series of NDVI data, but also to reproduce cloudfree images during any moment in that time series.

A secondary advantage is the reduction of the amount of data. In this case the major characteristics of the temporal NDVI behaviour could be described by only three Fourier components (three amplitudes and two phases), which is a reduction of 5/36.

HANTS has to be steered by five control parameters. The problem is that there are no objective rules to determine the magnitude of these control parameters. However, based on experience and after running several combinations, the control parameters can be set at the appropriate magnitudes.

Acknowledgements

The authors would like to thank the Free University of Berlin, Berlin (Germany) for the NOAA/AVHRR 10-days-maximum-NDVI composites.

References

CEC, 1993, CORINE Land Cover technical guide. ISBN 92-826-2578-8, European Union, Directorate General Environment Nuclear Saftey and Civil Protection, Luxembourg.

Fourier, J., 1818, Note relative aux vibrations des surfaces élastiques et au mouvement des ondes, *Bulletin des Sciences par Société Philomatique*, **1**, 126–136.

Henderson-Sellers, A., 1982, Defogging cloud determination algorithms. *Nature*, **298**, 419–420.

Koslowsky, D., 1996, Mehrjaehrige validierte und homogenisierte Reihen des Reflexionsgrades und des Vegetationsindexes von Landoberflaechen aus Taeglichen AVHRR-Daten hoher Aufloesung. Meteorologische Abhandlungen, Neue Folge, Serie A Monographien, Band 9/Heft 1, Free University of Berlin, Berlin, Germany.

Main, I. G., 1990, *Vibrations and Waves in Physics* (Cambridge: Cambridge University Press).

Menenti, M., Azzali, S., Verhoef, W., and van Swol, R., 1993, Mapping agro-ecological zones and time lag in vegetation growth by means of Fourier analysis of time series of NDVI images. *Advances in Space Research*, **5**, 233–237.

Simpson, J. J., and Gobat, J. I., 1996, Improved cloud detection for daytime AVHRR scenes over land. *Remote Sensing of Environment*, **55**, 21–49.

Stowe, L. L., McClain, E. P., Carey, R., Pellegrino, P., Gutman, G. G., Davis, P., Long, C., and Hart, S., 1991, Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data. *Advances in Space Research*, **3**, 51–54.

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

Verhoef, W., Menenti, M., and Azzali, S., 1996, A colour composite of NOAA-AVHRR-NDVI based on time series (1981–1992). *International Journal of Remote Sensing*, **17**, 231–235.