Available online at www.sciencedirect.com
ScienceDirect
Remote Sensing of Environment 112 (2008) 506–521
Remote Sensing of Environment
www.elsevier.com/locate/rse

# Hierarchical image segmentation based on similarity of NDVI time series

Stefaan Lhermitte [a,*], Jan Verbesselt [a], Inge Jonckheere [a], Kris Nackaerts [b], Jan A.N. van Aardt [c], Willem W. Verstraeten [a], Pol Coppin [a]

[a] M3-BIORES, Biosystems Department, Katholieke Universiteit Leuven, Celestijnenlaan 200E, Leuven, Belgium
[b] Mapping and Geospatial Solutions, Intergraph Belgium B.V., Riverside Business Park, Internationalelaan 55, Brussel, Belgium
[c] Ecosystems, Natural Resources and the Environment, CSIR, P.O. Box 395, Pretoria, South Africa

Received 15 January 2007; received in revised form 8 May 2007; accepted 10 May 2007

**Abstract**

Although a variety of hierarchical image segmentation procedures for remote sensing imagery have been published, none of them specifically integrates remote sensing time series in spatial or hierarchical segmentation concepts. However, this integration is important for the analysis of ecosystems which are hierarchical in nature, with different ecological processes occurring at different spatial and temporal scales. Therefore, the objective of this paper is to introduce a multi-temporal hierarchical image segmentation (MTHIS) methodology to generate a hierarchical set of segments based on spatial similarity of remote sensing time series. MTHIS employs the similarity of the fast Fourier transform (FFT) components of multi-seasonal time series to group pixels with similar temporal behavior into hierarchical segments at different scales. Use of the FFT allows the distinction between noise and vegetation related signals and increases the computational efficiency. The MTHIS methodology is demonstrated on the area of South Africa in an MTHIS protocol for Normalized Difference Vegetation Index (NDVI) time series. Firstly, the FFT components that express the major spatio-temporal variation in the NDVI time series, the average and annual term, are selected and the segmentation is performed based on these components. Secondly, the results are visualized by means of a boundary stability image that confirms the accuracy of the algorithm to spatially group pixels at different scale levels. Finally, the segmentation optimum is determined based on discrepancy measures which illustrate the correspondence of the applied MTHIS output with landcover–landuse maps describing the actual vegetation. In future research, MTHIS can be used to analyze the spatial and hierarchical structure of any type of remote sensing time series and their relation to ecosystem processes.


*Keywords:* Hierarchical image segmentation; Fast Fourier transform; NDVI time series; Vegetation mapping

## 1. Introduction

The development of effective methodologies to analyze time series of satellite imagery is one of the most important issues in the understanding of temporal dynamics of vegetation cover (Bruzzone et al., 2003). The temporal component, integrated with the spectral and spatial dimensions, provides essential information on ecological systems and vegetation dynamics. However, advanced analysis methods are crucial for the proper exploration of that information; certainly with the ever increasing amount of time series data. Several methods and algorithms have already been developed based on satellite-based biophysically meaningful variables, e.g. the Normalized Difference Vegetation Index (NDVI) whose behavior follows annual cycles of vegetation growth (Running et al., 1994). These applied methods include Principal Component Analysis (PCA) (Anyamba & Eastman, 1996; Eastman & Fulk, 1993; Gurgel & Fereira, 2003), development of phenological metrics (Jönsson & Eklundh, 2004; Lee et al., 2002; Reed et al., 1994; Verbesselt et al., 2006), change detection (Coppin et al., 2004), and harmonic or Fourier analysis (Andres et al., 1994).

The fast Fourier transform (FFT) has shown to be particularly useful for NDVI time series analysis to describe and quantify fundamental temporal characteristics, since the noise-affected NDVI time series are decomposed into simpler periodic signals in the frequency domain. By performing analysis in the frequency domain, a distinction can be made between frequency

* Corresponding author. Tel.: +32 16 329750; fax: +32 16 329760.
*E-mail address:* stefaan.lhermitte@biw.kuleuven.be (S. Lhermitte).

0034-4257/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2007.05.018

terms with daily frequencies, related to atmospheric and cloud-contamination effects, and specific frequency terms related to vegetation in dynamic ecosystems (Evans & Geerken, 2006; Jakubauskas et al., 2001, 2002; Juarez & Liu, 2001; Olsson & Eklundh, 2001). Azzali and Menenti (2000) and Moody and Johnson (2001) have used the inter- and intra-seasonal periodic signals successfully in classification procedures to map vegetation–soil–climate units. These studies revealed typical temporal characteristics of vegetation complexes, but they are per-pixel approaches based on clustering procedures of temporal properties of individual pixels. Consequently, they do not take into account the spatial or hierarchical context of the data. As such, they ignore the information in the spatial domain and fail to aggregate the temporal information into hierarchical regions at different scales. These concepts are important since ecosystems are hierarchical in nature, with different ecological processes occurring at different spatial and temporal scales (Handcock & Csillag, 2004; Hay et al., 2003). For example, macro-ecological characteristics, e.g. climate, will have coarse spatial regional effects, while more localized characteristics, e.g. weather, create patterns of variability at finer spatial scales. In this context, ecological systems can be perceived as nested patch hierarchies, where patterns and dynamics at the focal scale are products of the potential behaviors of components at lower levels (smaller scales), and are bound within the environmental constraints imposed by higher levels (larger scales) (Woodcock & Harward, 1992; Wu & Loucks, 1995).

Image segmentation methods provide a valuable alternative to the conventional per-pixel classification methods, since they consider the spatial context. Segmentation methods partition a study area into adjoining clusters of pixels, called segments or regions, based on similarity or dissimilarity of their single or multiple-layer pixel values (Stuckens et al., 2000). Mathematically, most of these methods operate on the principle of minimizing the within-region variance, or other measures of internal homogeneity (Beaulieu & Goldberg, 1989). Different approaches are commonly used for this principle, ranging from threshold techniques, and boundary techniques, to region-based techniques and hybridized approaches (Fan et al., 2001). The advantages of the segmentation approach over classical per-pixel procedures are multiple. Firstly, they allow quantification of spatial heterogeneity within the data at various scale levels. Such measures can indicate spatial complexity, variability, and fragmentation, which can have a significant influence on the rate, character, and magnitude of ecosystem processes. Secondly, the delineation of homogeneous patches is possible and involves a certain spatial generalization. This reduces the effect of local spatial heterogeneity that often masks larger spatial patterns (Tilton & Lawrence, 2000). Thirdly, an explicit hierarchal structure can be implemented between segments at different spatial scales (Woodcock & Harward, 1992). The hierarchical structure provides insight into the functional ecology of ecosystems, since it presents the study area as a nested patch hierarchy. This means that the study area is divided into spatial sets corresponding to coarse regions. These coarse sets are subdivided into subsets corresponding to region subparts at smaller scales. This hierarchy can be represented by a tree where the segments at the lower level are joined to form segments at higher levels.

Although a number of hierarchical image segmentation procedures for remote sensing imagery have been published (e.g. Baatz & Schäpe, 2000; Tilton & Lawrence, 2000), none of them specifically incorporates similarity of temporal information in the algorithm. The objective of this paper is consequently the introduction of a multi-temporal hierarchical image segmentation (MTHIS) methodology that generates a hierarchical structure of segments based on spatial similarity of temporal profiles. MTHIS employs the similarity of FFT components to assess that similarity of temporal profiles. Application of the MTHIS consequently allows hierarchical clustering of image time series into spatio-temporal segments at numerous scales based on specific periodic patterns. This will provide insight in the hierarchical spatio-temporal structure of ecosystem processes, e.g. the relation of different landcover properties at various spatial scales, the relationship between climate-weather and vegetation phenological variability.

In this paper, the MTHIS methodology is applied on NDVI time series of South Africa to demonstrate the concept. Section 2 presents the study area and satellite data, while the MTHIS methodology is described based on its underlying theoretical concepts in Section 3. Since MTHIS is a general methodology that can be applied to any image time series, a specific MTHIS protocol for multi-temporal NDVI image series is introduced in Section 4. This protocol serves to select the relevant temporal characteristics that describe the majority of spatio-temporal variation in the original NDVI data (4.1), incorporates effective application (4.2) and visualization (4.3), and allows to extract the segmentation optima that relate to ecological processes occurring at different scales (4.4). Finally, the results of the MTHIS protocol are presented in Section 5 and the advantages and drawbacks of the methodology are discussed in Section 6.

## 2. Data description

### 2.1. Study area

The proposed methodology was tested on the area of South Africa, Swaziland and Lesotho, which approximately encompasses the geographic region between latitudes 21°S and 35°S and longitudes 33°W and 16°E. The elevation ranges from sea-level to more than 3300 m, while the rainfall varies from almost zero to more than 3000 mm in mountainous areas. Rainfall regimes are defined as winter rainfall in the west to strong summer rainfall regimes in the northeastern and northern parts of the study area. The vegetation in the study area is characterized by 68 different vegetation types (LR) as described by Low and Rebelo's Vegetation Map of South Africa, Lesotho and Swaziland (Low & Rebelo, 1996) and is illustrated in Fig. 1a. These broad vegetation types are principally identified by their vegetation structure, ecological processes and occurrence of important plant species. If the factor of human influence is also considered, 31 landcover–

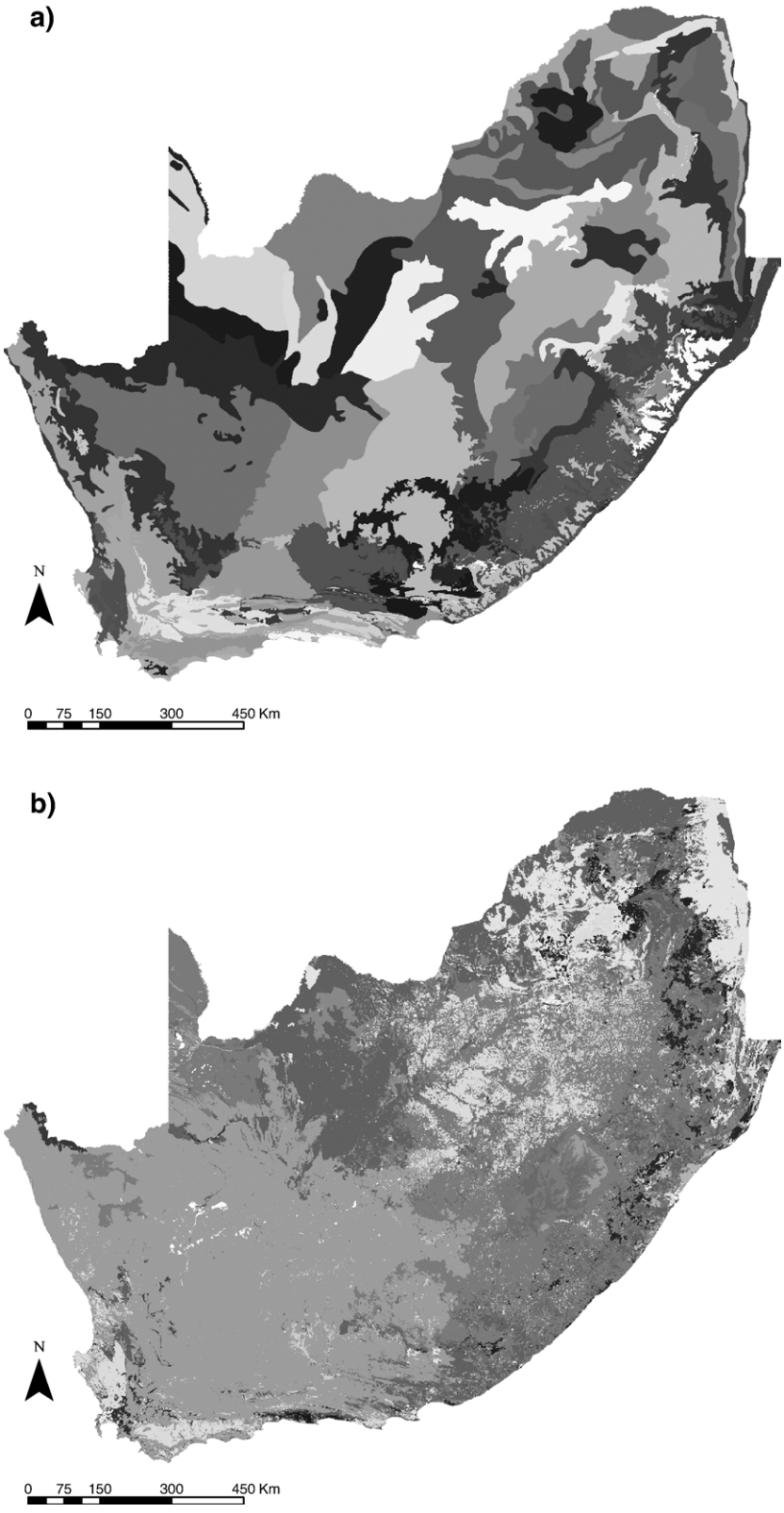


Fig. 1. a) The major vegetation types for South Africa, Lesotho and Swaziland exemplifying the spatial variation of vegetation structure, ecological processes and occurrence of important plant species. For more detailed information on the legend see Low and Rebelo (1996); b) Landcover classification for South Africa, Lesotho and Swaziland as an in indication of the spatial variation in landcover–landuse types. For more detailed information on the legend see Thompson (1999).

landuse types, ranging from natural vegetation to urban built-up land, can be distinguished as described by the National Land Cover Map of South Africa (LC) (Fig. 1b). This map was developed by the Council for Scientific and Industrial Research (CSIR) and the Agricultural Research Council (ARC) (Thompson, 1999).

### 2.2. Satellite data

Ten-daily NDVI image composites (S10) were acquired from the SPOT-VEGETATION (VGT) sensor. This low spatial resolution product (1 km) provides a very effective source for the examination of intra- and inter-annual vegetative variations, given their temporal resolution. Preprocessing of the data was performed by the Vlaamse Instelling voor Technologisch Onderzoek (VITO, Mol, Belgium) in the framework of the Global Vegetation Monitoring (GLOVEG) preprocessing chain. It consisted of the Simplified Method for Atmospheric Correction (SMAC) (Rahman & Dedieu, 1994) and compositing of daily images at ten-day intervals based on the Maximum Value Compositing (MVC) criterion (Holben, 1986). The final data set consisted of 180 ten-daily, 1 km resolution S10 composites for the period July 1998 to July 2003.

## 3. MTHIS methodology

The presented MTHIS methodology consists of a hierarchical segmentation approach to spatially cluster pixels based on their similar temporal behavior. Temporal similarity is the fundamental working principle of the MTHIS and is defined as the similarity of the FFT components. The MTHIS, consequently, contains two main phases: i) decomposition of the original image time series in FFT components, and ii) hierarchical segmentation based on the similarity of FFT components.

### 3.1. Fast Fourier transform

In the first step, the original image time series are decomposed in periodic signals using the fast Fourier transform (FFT). The FFT transforms a complex signal into a set of scaled sine and cosine waves that can be summed to reconstruct the original signal. The mixed radix FFT (Singleton, 1969) was used in this study, since it is a computationally fast variant of the discrete Fourier transform (DFT). This FFT can be used to transform any equidistant discrete time series $f(t)$ and is given by:

$$F_k = \frac{1}{N}\sum_{t=0}^{N-1} f_t e^{-2\pi i k t/N} \tag{1}$$

where $t$ is an index representing the sample number, $f_t$ is the time series value at moment $t$, $k$ is the frequency of the FFT component $F$, and $N$ is the number of samples in the time series. Eq. (1) contains a real and imaginary part, but can be decomposed into a set of cosine (real part) and sine (imaginary part) waves in rectangular notation based on Euler's equation (James, 1994):

$$F_k^c = \frac{1}{N}\sum_{t=0}^{N-1}(f_t \cos(2\pi k t/N)) \tag{2}$$

$$F_k^s = \frac{1}{N}\sum_{t=0}^{N-1}(f_t \sin(2\pi k t/N)) \tag{3}$$

where $F_k^c$ and $F_k^s$ are the cosine and sine parts, respectively. The frequency $k$ of the FFT components accordingly designates the number of cycles the sine and cosine waves complete over the time series (e.g., the fifth term completes five cycles over five years) and defines the periodicity or time between consecutive sine and cosine waves (period = $1/k$). An alternative polar notation exists wherein the time series $f(t)$ is reconstructed using only cosine waves with unique amplitude $A_k$ and phase shift $\phi_k$ (Smith, 1999):

$$f(t) = A_0 + \sum_{k=1}^{N-1} A_k \cos(2\pi k t + \phi_k) \tag{4}$$

with

$$A_k = \sqrt{F_k^{c2} + F_k^{s2}} \tag{5}$$

and

$$\phi_k = \arctan\left(\frac{F_k^c}{F_k^s}\right). \tag{6}$$

In the polar notation $A_k$ and $f_k$ jointly describe the $k$th frequency FFT component as one cosine wave in the frequency domain, whereas the sum of the cosine waves represents the original time series of each pixel. This representation as unique cosine wave allows to discriminate between processes that contribute to the original time series signal with different periodic patterns. Additionally, the influence of these processes on the signal can be quantified.

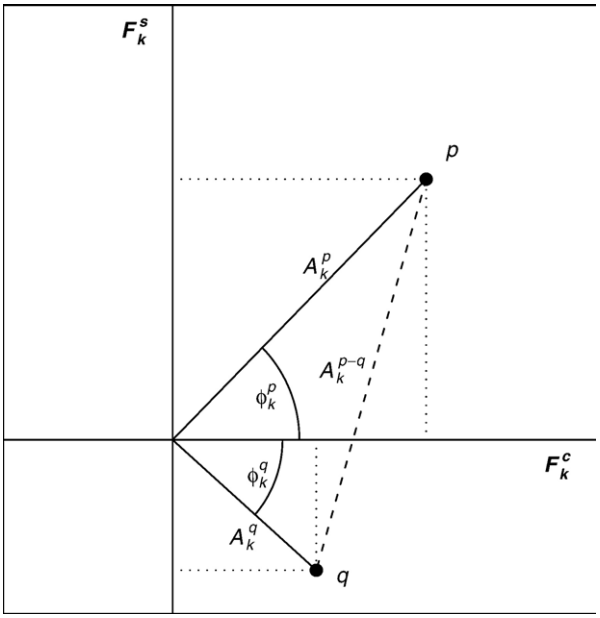


Fig. 2. Illustration of the $F_k$-distance for the $k$th frequency FFT components of two segments, represented by their mean, $p$ and $q$. $F_k^c$ and $F_k^s$ are amplitudes of the cosine and sine waves in rectangular notation, while $A_k$ and $\phi_k$ are the amplitude and phase in polar notation, respectively. $A_k^{p-q}$ is the distance between points $p$ and $q$ used in the $F_k$-distance criterion.

### 3.2. Hierarchical segmentation based on FFT component similarity

In the second step, the FFT components are imported in the hierarchical segmentation approach. The principle of MTHIS is analogous to classical hierarchical image segmentation which uses bottom-up region-merging techniques, e.g., eCognition (Baatz & Schäpe, 2000) or Recursive Hierarchical Segmentation (RHSEG) (Tilton & Lawrence, 2000). MTHIS starts with an initial partitioning of the image data into initial segments, which is an assignment of each image pixel to a separate region. Next, a segmentation run is launched in which each segment is selected once in a complete random sequence and compared with its spatially adjacent segments for similarity based on a

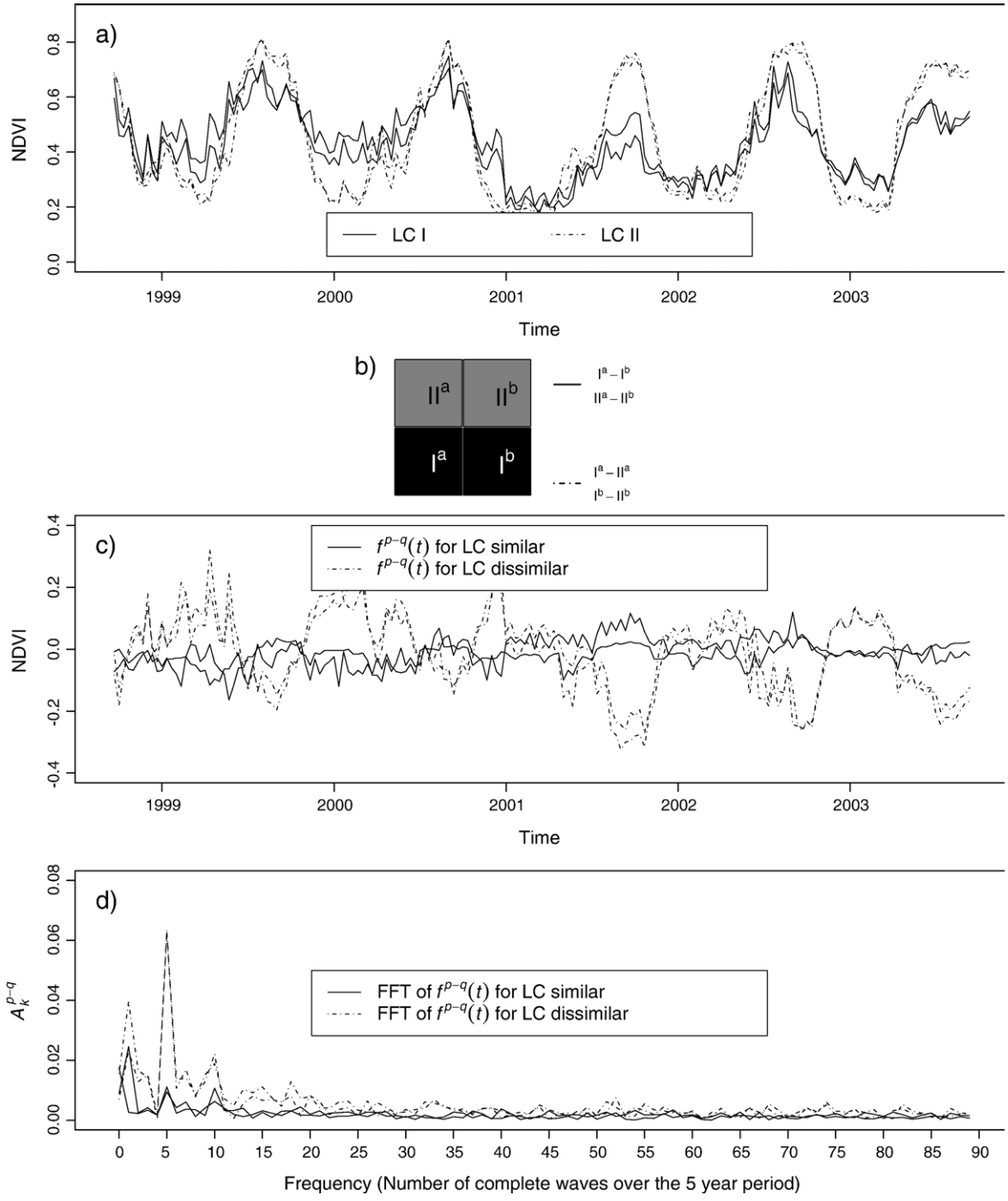


Fig. 3. Illustration of the difference vector: a) Four NDVI time series of neighboring one-pixel segments are shown. The time series with solid lines correspond to one reference unit (LC I), while the dot-dashed time series originate from a different reference unit (LC II); b) Schematic overview of relative location of one-pixel segments in (a). The line types respond to the line types of (c) and (d). Solid lines are used for the difference between pixels of similar LC, while dot-dashes are used when different LCs are compared; c) Difference vectors $f^{p-q}(t)$ for the time series of (a); d) $F_k$-distance as a function of the frequency. The $A_k^{p-q}$ is the amplitudes of the FFT components of the difference vectors of (c).

dissimilarity criterion. This dissimilarity criterion provides a measure $S$ that indicates the dissimilarity of objects. The pair of compared segments that is most similar (i.e., the minimum $S$) is subsequently merged to form a larger region if their dissimilarity $S$ remains under a user-defined similarity threshold value $T$. In this merge, the characteristics of the original segments are replaced by the mean of their original pixels. After this merge, the next segment is selected in the random sequence. If the segment has no neighbors where $S<T$, the process is halted for that segment. When each segment is handled once in the segmentation run, a new segmentation run is started for each segment. This process of segmentation runs continues until all merging possibilities end ($S>T$ for all segments). The entire process is repeatedly iterated for a range of increasing similarity threshold values $T$, until it results in a hierarchical set of spatio-temporal segments of the same image time series at different levels of detail. The bottom-up approach establishes hierarchy, since segments at coarse levels of detail are produced from simple merges of segments at finer levels of detail.

The classical hierarchical image segmentation approach was modified in this paper by introducing the $F_k$-distance criterion. This criterion employs the Euclidian distance between the FFT components of the same frequency as measure of similarity. The dashed line $A_k^{p-q}$ in Fig. 2 illustrates the $F_k$-distance for the $k$th frequency FFT component of two segments, represented by their mean, $p$ and $q$. The $F_k$-distance incorporates both parameters that represent the FFT component, $A_k$ and $\phi_k$, respectively, into one dissimilarity measure that depends on the amplitude and phase difference between the FFT components. Consequently, it describes both differences in amplitude and phase between the segments. Moreover, the $F_k$-distance differs from classical approaches that quantify the difference in amplitude or phase separately, since FFT components are not additive in polar notation. The $F_k$-distance mathematically corresponds to subtracting the NDVI time series of neighboring segments, $f^p(t)$ and $f^q(t)$, respectively, for each observation in the temporal sequence and using the amplitude of the resulting difference vector $f^{p-q}(t)$:

$$A_k^{p-q} = \sqrt{\left(F_k^c(p) - F_k^c(q)\right)^2 + \left(F_k^s(p) - F_k^s(q)\right)^2} = \sqrt{F_k^c(p-q)^2 + F_k^s(p-q)^2}. \quad (7)$$

$A_k^{p-q}$ directly includes spatial context in the MTHIS methodology as the NDVI times series of neighboring segments are subtracted before the amplitude of difference vector $f^{p-q}(t)$ is calculated. Insertion of the $F_k$-distance in the classical hierarchical image segmentation approach results in the following dissimilarity criterion $S$:

$$S = \sum_{k=0}^{N-1} w_k A_k^{p-q} \quad (8)$$

where $A_k^{p-q}$ is the $F_k$-distance between the mean FFT components of the pixels of segments $p$ and $q$ respectively, and $w_k$ is the weight of the $k$th frequency FFT component. Modification of the weights allows to enhance (high $w_k$) or diminish (low $w_k$) the influence of each component on $S$ and to accentuate specific components in the segmentation.

The $F_k$-distance criterion is illustrated in Fig. 3. Fig. 3a shows the NDVI time series of four neighboring one-pixel segments located in different landcover–landuse types (LC). The difference vectors of the time series are plotted in Fig. 3c, whereas Fig. 3d shows the amplitudes $A_k^{p-q}$ of the difference vectors. The latter represents the $F_k$-distances between the FFT components of the segments used in the dissimilarity criterion $S$. It clearly reveals three peaks for the $A_k^{p-q}$ of dissimilar LCs. The peak at the first frequency component, for example, corresponds to one amplitude cosine wave over the studied time frame (five years). It can be interpreted as a trend term that reveals the tendency differences between the compared time series. The peak at the 5th frequency component relates to a sine wave that completes five waves over the studied time frame. It reflects the annual difference of the time series, i.e., the annual cycle for each of five years and is called the first harmonic. The third peak at the 10th frequency component corresponds to the second harmonic. These harmonics originate from the FFT property that periodic signals of frequency $k$ can be decomposed in cosine waves of frequencies $k$, $2k$, $3k$, etc. However, the curves of the similar time series in Fig. 3d do not show these peaks clearly.

In other words, the difference between two similar NDVI time series is almost constant with a mean of zero. The difference of dissimilar time series (dot-dash lines) on the other hand, presents a clear annual difference, which is evident in Fig. 3c. The use of $A_k^{p-q}$ as dissimilarity criterion now allows discrimination between such similar and dissimilar time series and consequently enables the creation of a hierarchical structure of spatio-temporal segments.

## 4. MTHIS protocol

The MTHIS methodology can be applied to any image time series. Daily or monthly NDVI time series of land surfaces contain however strong systematic periodic patterns related to vegetation features and nonsystematic high frequent image noise caused by atmospheric and viewing angle effects and cloud contamination (Azzali & Menenti, 2000). These characteristics specifically can be exploited in MTHIS application to remove noise factors and enhance vegetation specific information. As a result, a close agreement can be expected between segmentation output and ecological processes related to vegetation at different scales. A specific MTHIS protocol for NDVI time series was therefore applied on the VGT NDVI time series of South Africa. The MTHIS protocol allows to select relevant FFT components that describe the majority of spatio-temporal variation in the original NDVI data, incorporates actual segmentation and visualization, and assesses the optimal MTHIS parameters for describing the ecological processes at different scales.

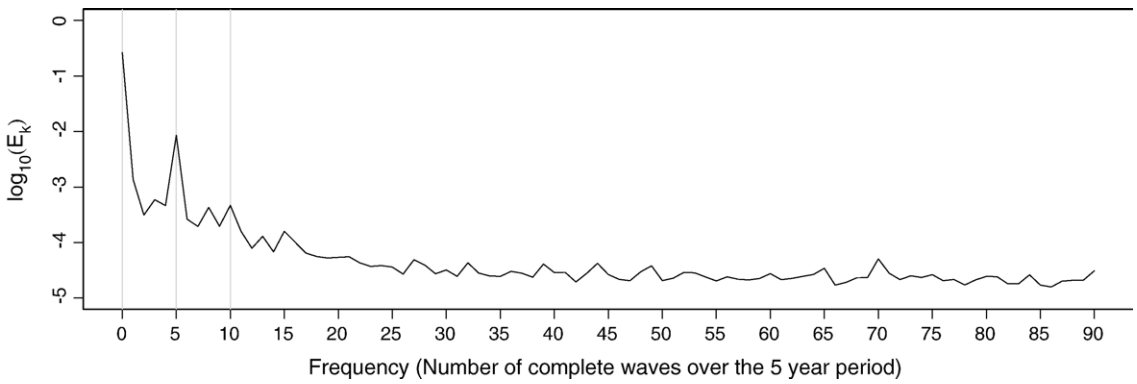


Fig. 4. Logarithmic plot of the energy density spectrum describing how the energy (or variance) of time series is distributed over the frequency FFT components. The 0th and 5th frequency term express 95% and 3% of the total energy of all terms.

### 4.1. Selection of FFT components

For time series with strong periodic patterns, such as the used NDVI time series, few principal FFT components will explain the majority of the variation in the time series. This allows a reduction of the FFT components and can be used to remove noise factors and enhance specific information. Consequently meaningful and stable characteristics of time series can be used in the MTHIS procedure. Moreover, the elimination increases the computational efficiency of the segmentation approach because fewer components have to be compared in Eq. (8). The selection and identification of these relevant spatio-temporal FFT components consisted of two steps. Firstly, the temporal informational content of the FFT components was assessed by the energy density spectrum ($E_k$) (Smith, 1999):

$$E_k = \frac{A_k^2}{2\pi}. \tag{9}$$

The energy density spectrum describes how the energy (or variance) of a time series is distributed according to frequency. It allows the distinction of relevant periodic signals in the original time series. Secondly, the spatial variability of the FFT components was assessed by means of the $F_k$-distance as a function of the segment lag distance. This measure is derived from the variogram (see Garrigues et al., 2006) and quantifies the spatial heterogeneity of the Fourier components using the amplitude of the resulting difference vector $A_k^{p-q}$. The use of $A_k^{p-q}$ allows to assess the spatial variability that is directly included in the MTHIS methodology, namely the amplitude of the difference vector, and that differs from the classical spatial heterogeneity measures due to the non-additivity of Fourier components in polar notation. Consequently, it enables the distinction of FFT components with large spatial variability for MTHIS application from components with uniform spatial distribution. For this purpose, the $F_k$-distance was calculated between all possible pairs of 500 random sample pixels in the study area and each pair was assigned a lag or distance interval class $h$:

$$\gamma_k(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} A_k^{p-q} \tag{10}$$

where $\gamma_k$ ($h$) is the mean $F_k$-distance among sample pixels separated by lag $h$, $N$ ($h$) is the number of paired pixels in lag $h$, and $A_k^{p-q}$ is the $F_k$-distance between a pair of pixels $p$ and $q$ in lag $h$. $\gamma_k$ ($h$) consequently gives an indication of how similar the FFT component is as a function of the lag distance.

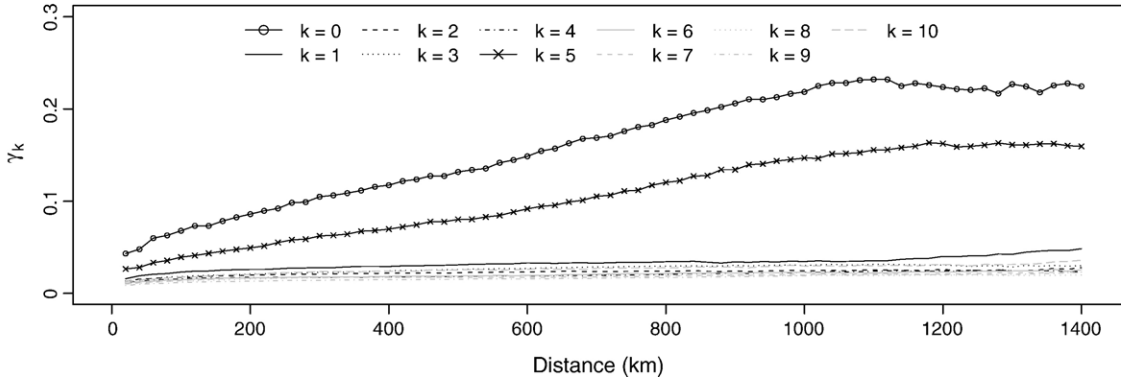


Fig. 5. The mean $F_k$-distance among sample pixels separated by lag $h$=20 km for the $k$th frequency FFT components ($k$=0, 1,..., 10).

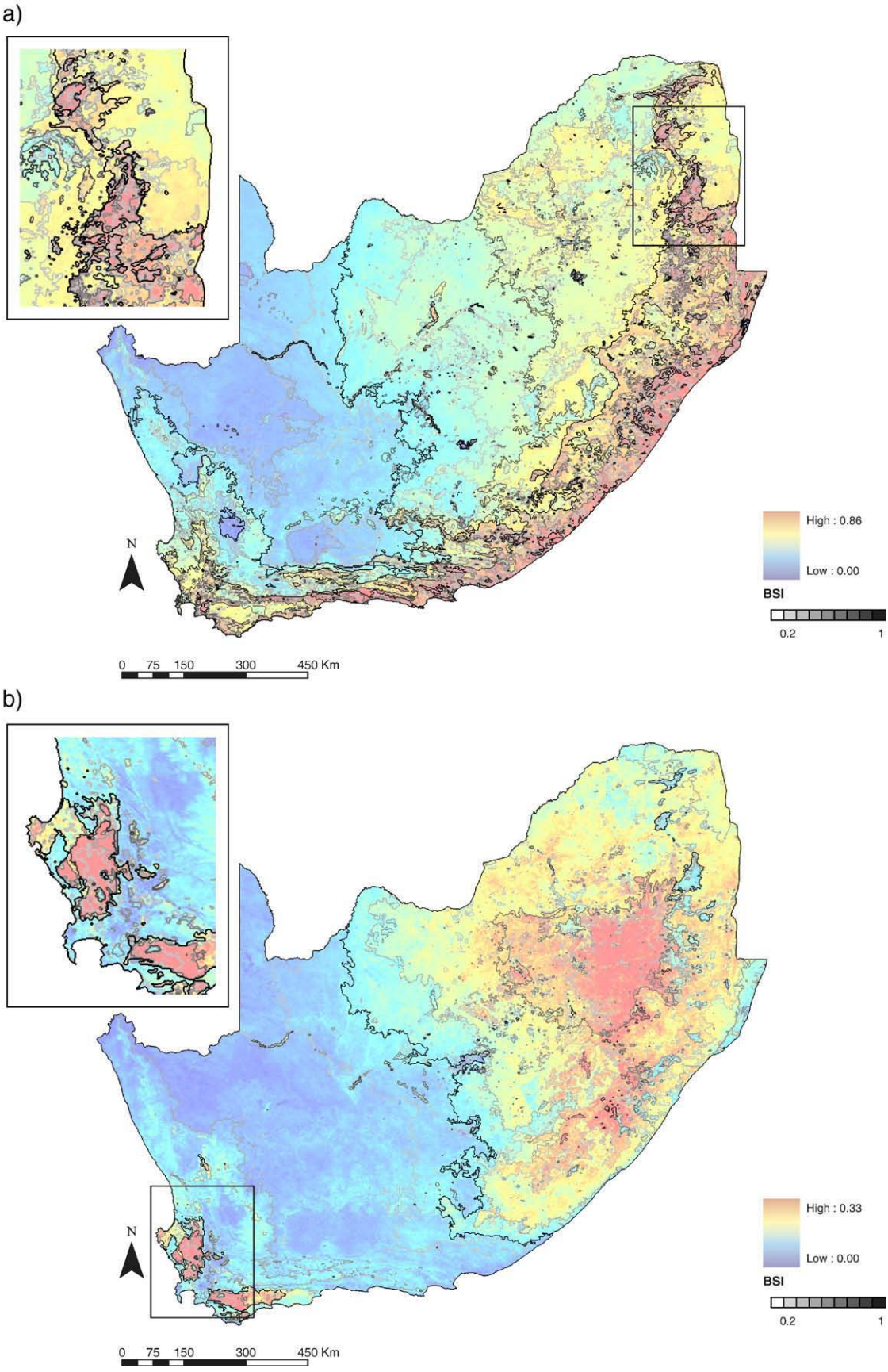


Fig. 6. a) Amplitude of the 0th frequency FFT component (average term; in color scale) and BSI overlay (in gray scale) for $w^0=1$, $w^5=0$ and $T=0.005, 0.01,\ldots, 0.14$; b) Amplitude of the 5th frequency FFT component (annual term; in color scale) and BSI overlay (in gray scale) for $w^0=0$, $w^5=1$ and $T=0.005, 0.01,\ldots, 0.14$. BSI values smaller than 0.2 are transparent. Zoom windows with a region subset are moreover provided.

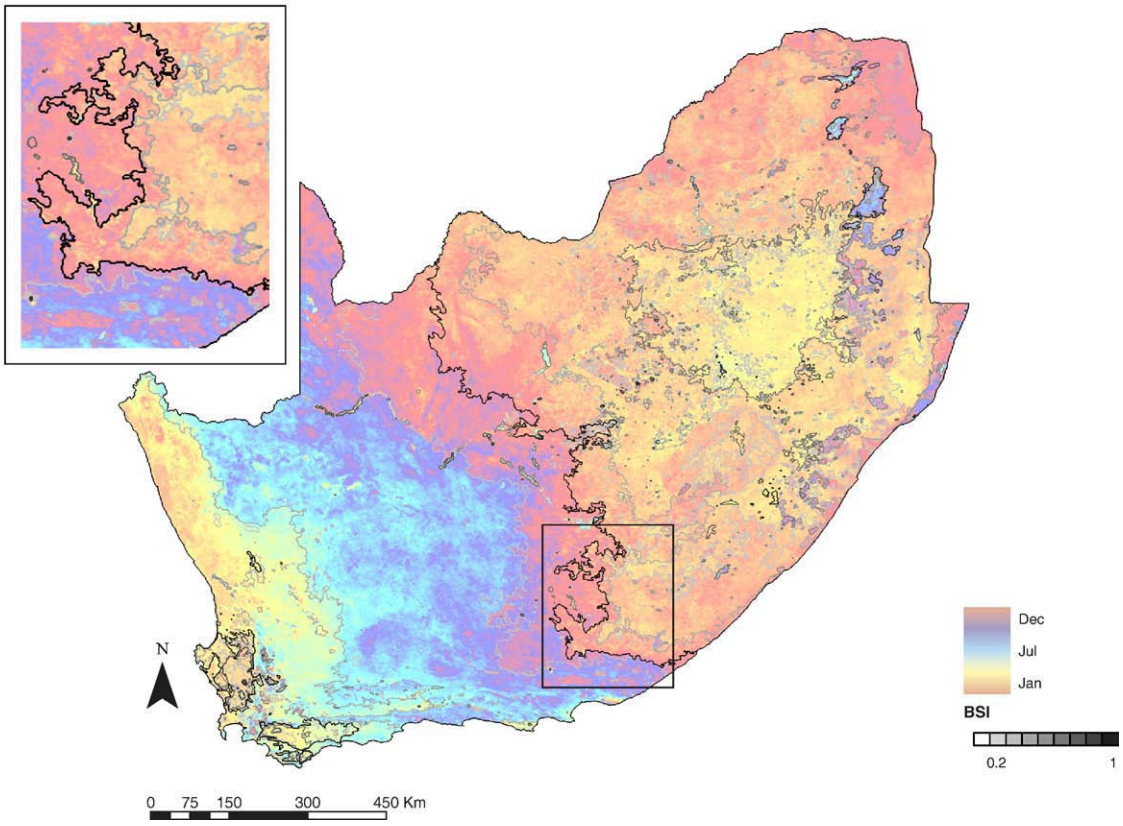


Fig. 7. Phase of the 5th frequency FFT component (annual term; in color scale) and BSI overlay (in gray scale) for $w_0=0$, $w_5=1$ and $T=0.005, 0.01,\ldots, 0.14$. BSI values smaller than 0.2 are transparent. A zoom window with a region subset is moreover provided.

### 4.2. MTHIS application

After the selection and identification of relevant vegetation spatio-temporal FFT components, the actual MTHIS was performed. The iterative process was conducted in repeated runs by assigning different weights ($w_k=0, 0.1,\ldots, 1.0$) to the FFT components in the dissimilarity criterion in Eq. (8) and by varying the segmentation threshold values ($T=0.005, 0.010,\ldots, 0.14$). The FFT components that represented limited spatio-temporal variability in the original VGT time series were discarded in the whole process by assigning their $w_k=0$.

### 4.3. MTHIS visualization

The results of the MTHIS runs were visualized using a boundary stability image (BSI) (Lucieer & Stein, 2002). The BSI shows the boundaries of different scale levels in one image and, consequently, allows a visual comparison and interpretation of segmentation boundaries and the segmentation hierarchy. The BSI was established by calculating the relative presence of the boundary at different threshold levels. This was done by selecting the segment edge pixels at each threshold level $T$. At step $t$ these boundary pixels were assigned the value 1 and non-boundary pixels the value 0 and were represented on a segment-boundary image $I_t$. This resulted in a BSI, defined as:

$$\mathrm{BSI} = \frac{1}{N_t}\sum_{t=0}^{N_t} I_t. \tag{11}$$

The BSI contains values between 0 and 1, related to relative presence of the boundary at different scale levels. Segment boundaries with large BSI values are boundaries of coarse scale segments that are detected at various threshold values $T$, whereas small BSI values represent boundaries occurring only at fine segmentation scales.

### 4.4. Assessment of segmentation optimum

The MTHIS methodology generates hierarchical segments of homogeneous temporal properties at various scales. Application of MTHIS on NDVI time series however allows additional interpretation, since NDVI time series relate closely to phenological characteristics of vegetation (Justice et al., 1985; Reed et al., 1994). A close agreement can therefore be expected between MTHIS segments of NDVI time series and ecological processes that determine phenological characteristics. For example, agreement can be assumed between the MTHIS output and the reference maps that describe the vegetation characteristics at different scales

(LR and LC see Section 2.1), because the interaction of soil, physical environment, climate, and landcover–landuse defines the phenological characteristics of vegetation. The agreement will be maximal at a specific segmentation optimum where reference and segmentation coincide at the pre-defined management scale of the reference layer. The goal of this step in the MTHIS protocol is the extraction of the different segmentation optima by discrepancy measures that compare the segmented image with the reference maps. This allows the assessment of the optimal threshold and weight values and also provides quantitative indicators that estimate the agreement of the segmentation optimum with the reference maps.

#### 4.4.1. Thematic agreement

A commonly used discrepancy metric to quantify the agreement of thematic maps is Cohen's Kappa coefficient (Cohen, 1960). It requires thematic labeling of the MTHIS output, which was done by assigning the zonal majority value of the reference layer (the value that appears most often) to the MTHIS output segments. Kappa coefficients ($\hat{K}$) were calculated succeedingly to assess the disparity between MTHIS and reference segments:

$$\hat{K} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+}\cdot x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+}\cdot x_{+i})} \quad (12)$$

where $r$ is the number of labels in the reference layer, $x_{ii}$ is the number of correctly assigned pixels, $x_{i+}$ is the total number of pixels classified as $i$, $x_{+i}$ is the total number of pixels $i$ in the reference layer, and $N$ is the number of pixels in the image. $\hat{K}$ ranges from 0 to 1 indicating the percentage accuracy above chance agreement. Because of the assignment of zonal majority values of the reference layer, the Kappa coefficients overestimate the absolute thematic accuracy. Nevertheless, Kappa coefficients allow the relative comparison of geometric correspondence between segmentation output and reference layer, since high $\hat{K}$ values can be expected when the agreement between segmentation and reference is nearly perfect, whereas the assignment errors in the zonal majority procedure significantly would reduce the Kappa coefficient in cases where they show imperfect agreement.

#### 4.4.2. Boundary agreement

The Kappa coefficient, however, fails to correct for over-segmentation that results in an over-estimation of thematic correspondence. Thus, another agreement measure was applied based on the accuracy of segmentation boundaries to correct for this over-segmentation. The accuracy of the detected boundaries was formulated by a boundary accuracy measure $D(B)$ that expresses the average distance measured in pixels between a segment and reference boundary pixel (Delves et al., 1992):

$$D(B) = \frac{|B - M| + \sum_{b=1}^{B} D(b)}{B} \quad (13)$$

where $b$ is a boundary pixel in the reference map, $D(b)$ is the shortest Euclidian distance measured in pixels between $b$ and any boundary pixel in the segmented image, and $B$ and $M$ are the number of boundary pixels in the reference and segmented image, respectively. For a perfect fit, $D(B)$ equals 0, while higher values indicate higher average distances measured in pixels between segment and reference boundaries, and thus higher discrepancies. The $D(B)$ measure is corrected for over-segmented objects by penalizing segmentation outputs with high $M$ values via $|B-M|$.

## 5. Results

The results of MTHIS protocol on the study area are divided into three sections, namely i) the selection and analysis of the FFT components that describe the majority of the spatio-temporal variability, ii) the visualization of the MTHIS results using the BSI, and iii) assessment of the segmentation optima of MTHIS based on the thematic and boundary agreement in comparison with the LR and LC reference layers.

### 5.1. Selection of FFT components

The energy density spectrum in Fig. 4 shows how the variance of the original time series in the study area is distributed across the FFT components. It reflects the importance of each FFT component to describe the original time series. The average term is the term with zero frequency or the overall mean NDVI of the time series. The annual term is the term with frequency five over five years and it is related to the annual growing patterns of the time series. Both represent the largest part of the temporal variability of the original time series. Together they describe more than 98% of the variation in the logarithmic plot. Contrarily, the FFT components with frequency above ten add little information to the original VGT time series.

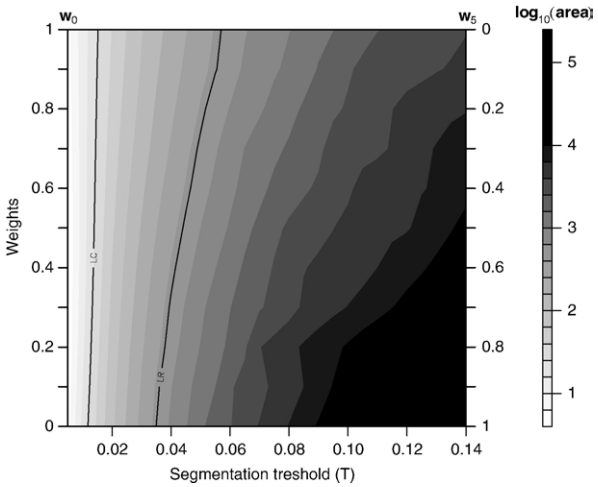


Fig. 8. Mean segment area as a function of the weights of the 0th and 5th frequency FFT components ($w_0$ and $w_5$; where $w_0+w_5=1$) and the segmentation threshold $T$. The gray scale represents the $\log_{10}$ of the mean segment area. The lines correspond to segmentation results with equal mean segment areas as LR and LC, respectively.

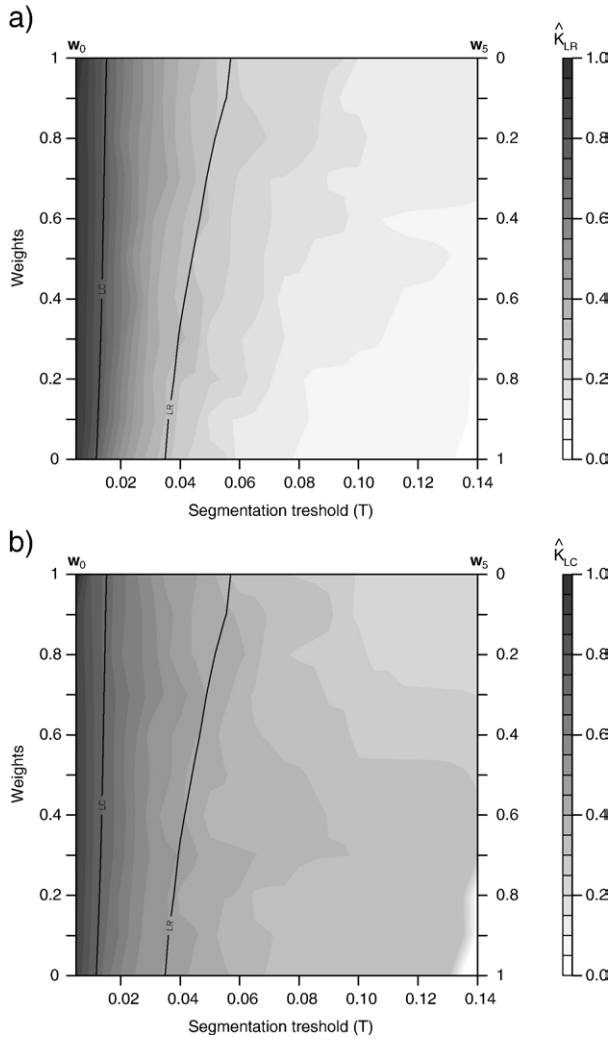


Fig. 9. Kappa coefficient as a function of the weights of the 0th and 5th frequency FFT components ($w_0$ and $w_5$; where $w_0+w_5=1$) and the segmentation threshold $T$. The gray scale represents Kappa coefficient of Eq. (12). The lines correspond to segmentation results with equal mean segment areas as LR and LC, respectively: a) $D(B)$ with LR as reference layer; b) $D(B)$ with LC as reference layer.

Fig. 5 presents the spatial variability ($\gamma_k$) of 0th to 10th frequency FFT component that describe the largest part of the temporal information in Fig. 4. It shows the spatial variability of the Fourier components by means of the mean $F_k$-distance among sample pixels separated by lag $h$ and consequently reflects the spatial variability that is directly included in the MTHIS methodology, namely the amplitude of the difference vector. It confirms the relevance of the average and annual term, as the 0th and 5th frequency FFT components express the largest part of the spatial variability for the MTHIS. These two FFT components consequently were selected for further use in MTHIS, as they clearly correspond to relevant periodic signals in the original NDVI time series. The other terms reflect a significantly smaller amount of the NDVI time series spatio-temporal variability. Comparison of Figs. 4 and 5 also reveals the distinction in spatio-temporal variability between the average and annual term. The level of spatial variability of the annual term ($\gamma_5$) is approximately 60% of the level of spatial variability of the average term ($\gamma_0$) (Fig. 5).

The peak of the annual term, on the other hand, explains considerably less of the variance of the original time series (95% vs. 3%) in the logarithmic plot of the energy density spectrum (Fig. 4). This means that the annual term contains more spatial variation relative to the average term, since the annual term shows a variation over space that is 40% lower, whereas overall variation is considerably lower. This also indicates that higher spatial variation should be apparent, which implies more localized variability, less spatial correlation, and finer spatial patterns, resulting in more speckled images.

The color scale in Fig. 6a–b displays the amplitude of the average ($A_0$) and annual term ($A_5$), respectively. Comparison of both figures reveals the characteristics of both FFT components. The average term shows high overall variability ranging from 0 to 0.86 with coarse-smooth spatial patterns. Red colors indicate

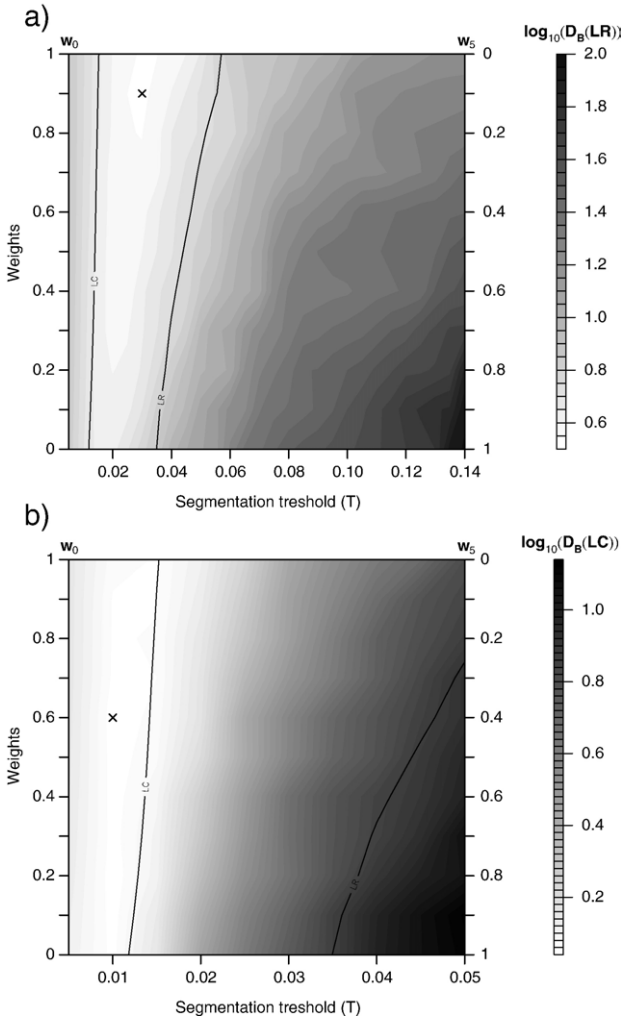


Fig. 10. Boundary accuracy measure D(B) as a function of the weights of the 0th and 5th frequency FFT components ($w_0$ and $w_5$; where $w_0+w_5=1$) and the segmentation threshold $T$. The lines correspond to segmentation results with equal mean segment areas as LR and LC, respectively: a) $D(B)$ with LR as reference layer; b) $D(B)$ with LC as reference layer and zoomed in on $T \leq 0.05$. The X represents the segmentation optimum for each reference layer.

high five year mean NDVI values related to a high vegetation cover, whereas the blue regions reflect low vegetation cover. The forested areas in the northeastern part and in the eastern coastal areas can clearly be delineated in this context, whereas the dry Karoo areas in the northwest also can be detected. The amplitude of annual term, on the other hand, presents only a third of the overall variability (from 0 to 0.33), but contains finer localized spatial patterns, resulting in a more speckled image. The red tones in Fig. 6b are related to a pronounced annual signal in the original time series, whereas the blue areas do not show this clear annual variation. This explains why agricultural and grassland regions in the southwestern and central-eastern parts contain red colors. These areas vary annually from dry soil or cured grass to high live vegetation content. Contrarily, the forested areas show little annual variation, as they represent regions with a high mean NDVI and little annual variation.

### 5.2. MTHIS visualization

Fig. 6a–b contains also the BSI output of two MTHIS runs based on only the average term ($w_0=1$ and $w_5=0$) and annual term ($w_0=0$ and $w_5=1$), respectively. The BSI boundaries are plotted in gray scale on top of the colored amplitude maps and reflect the relative presence of the boundary at different scale levels. The darker BSI values are boundaries of coarse scale segments that are detected at various threshold values $T$ and indicate the presence of the boundary at several hierarchical scale levels, while lighter values are indicators of boundaries only occurring at small threshold values; non-boundary pixels and BSI values below 0.2 are transparent. The BSI boundaries effectively indicate how the study area is partitioned at different hierarchical scale levels, since the segments with low BSI values are subsets of the coarse regions with high BSI values.

Additionally, the BSI gives an indication of the quality of the MTHIS methodology, since it allows visual comparison with the input FFT component images. This visual comparison illustrates that boundaries of clearly visible objects (e.g., forested areas in the northeast of Fig. 6a or the agricultural areas in the southwest of Fig. 6b) were identified at various scale levels, while more subtle differences were only detected at finer scales. This can be clearly seen in the zoom windows of Fig. 6a–b. Another example is the distinction between summer and winter rainfall regions that is detected at all scale levels in Fig. 7. This figure illustrates the phase instead of the amplitude of the annual term. The colors express the phase on the annual term and describe the time of occurrence of the annual peak in the original NDVI temporal profile, while the gray scale shows the BSI. The BSI illustrates how the difference between summer and winter rainfall is detected at all scale levels, while small phase differences are only recognized at finer scale levels.

A closer view of the BSI images also reveals the influence of threshold and weight values on the segment size. Small values of T lead to small segments, whereas large $T$ values result in

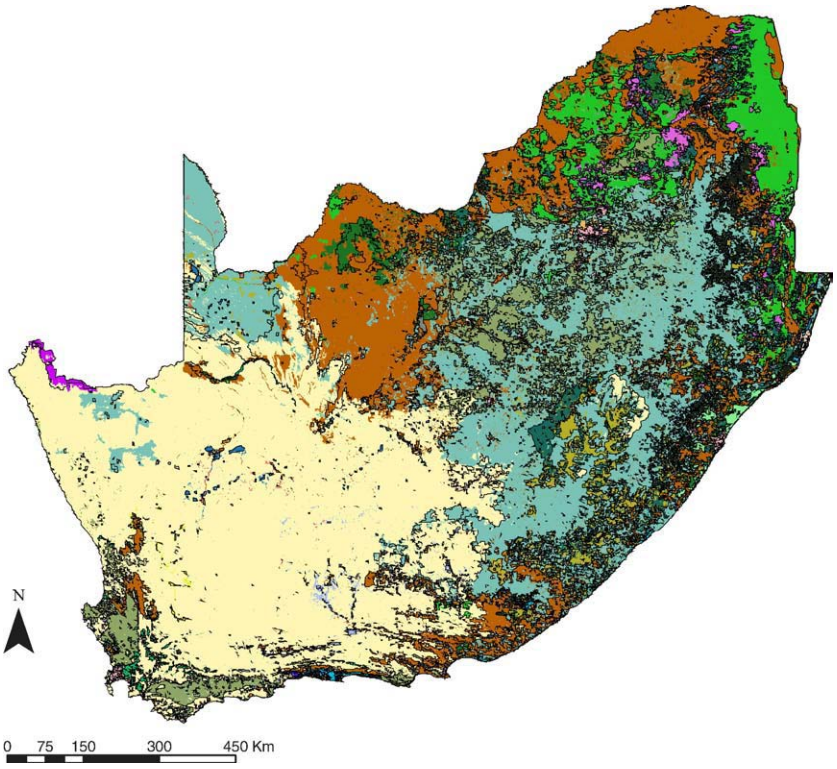


Fig. 11. MTHIS segmentation optimum in comparison with the LC map after assigning the zonal majority value of the LC map. The colors represent the original LC map, whereas the boundaries reflect the segmentation optimum after zonal majority assignment. The values for this MTHIS output are $D(B)=1.10$ and $K\hat{}=0.88$ for $w_0=0.6$, $w_5=0.4$, and $T=0.01$.

large objects. The combination of weights and spatial variability of the FFT components likewise will influence the segment size as it will determine the intra-region similarity, $S$, before reaching threshold $T$. This is confirmed in Fig. 8 where the mean segment area is plotted as a function of the weights wk and segmentation threshold $T$. It shows that the higher overall spatial variability of the average term resulted in smaller segments for the same $T$ value. A logarithmic scale is used, since an exponential area size can be expected for increasing $T$ values as each merge in the MTHIS approximately doubles the segment size.

### 5.3. Assessment of segmentation optimum

The agreement of the MTHIS output relative to the reference layers is presented in Fig. 9 based on thematic accuracy values $\hat{K}$. The Kappa coefficients decrease for increasing threshold values $T$. Very high Kappa coefficients, however, are associated with severely over-segmented results as can be seen after comparison with Fig. 8. Very low values, on the other hand, correspond to under-segmentation. A correct interpretation of $\hat{K}$ or the extraction of optimal $T$ and $w_k$ values is only reasonable when there is no over-or under-segmentation. This can be achieved when the segment size of segmentation output and reference has similar magnitudes as indicated by the lines in Fig. 9. From these lines, it appears that only the thematic labeling for LC resulted in $\hat{K}$ values above 0.7, whereas lower Kappa coefficients were obtained for the LR labeling.

Fig. 10a–b presents the boundary accuracy measure $D(B)$ from Eq. (13) that was corrected for over-segmentation. A logarithmic scale is also used here, since $D(B)$ values are also related to exponential area size. All subplots display a segmentation optimum for each combination of weights, where the correspondence between segmentation output and reference reach a maximum in comparison with the $T$ values. Analysis of the values of $D(B)$ at these optima confirms the result of Fig. 9a, showing the best correspondence between segmentation optimum and LC. The MTHIS output resembled more accurately the LC map ($D(B)=1.10$ and $\hat{K}=0.88$ for $w_0=0.6$, $w_5=0.4$, and $T=0.01$) than the LR map ($D(B)=3.47$ and $\hat{K}=0.52$ for $w_0=0.9$, $w_5=0.1$, and $T=0.03$). This also illustrates the importance of the combination of the average ($w_0=0.6$) and annual ($w_5=0.4$) temporal information for a correct landcover–landuse mapping, while the influence of the annual term is less ($w_5=0.1$) for optimal LR delineation. Fig. 11 shows the borders of the MTHIS segmentation optimum in comparison with the LC map after assigning the zonal majority value of the LC map, whereas the color scale represents the original LC map. It can be seen that the output closely resembles the original LC map, where small regions differences occur for the small regions of in the LC map due to the zonal majority procedure.

## 6. Discussion

### 6.1. MTHIS methodology

Although the MTHIS was developed for image time series, it provides a hierarchical image segmentation methodology that can be applied to any image series, e.g., image hyperspectral series. The proposed MTHIS methodology differs however in two points from the classic hierarchical segmentations. Firstly, the segmentation is not based on original data values, but on the decomposition of these values in FFT components. Secondly, the incorporation of the $F_k$-distance criterion allows effectively to cluster based on the similarity of the FFT components, since it combines both parameters that represent the FFT component, $A_k$ and $\phi_k$, respectively, into one dissimilarity measure. On the contrary, classical hierarchical segmentations such as eCognition (Baatz & Schäpe, 2000) or RHSEG (Tilton & Lawrence, 2000) consider each input parameter separately. These classic methodologies consequently do not allow to measure the similarity of the FFT components of the same frequency, since the difference between FFT components cannot be calculated by subtracting amplitude and phase separately (Smith, 1999).

The main advantage of the methodology is the use of FFT components, which enables the distinction of signals with a specific period. This is particularly amendable for detecting periodic patterns in time series of satellite data, such as daily or monthly NDVI time series of land surfaces that contain strong systematic periodic patterns related to vegetation features and nonsystematic high frequent image noise. Hence, information related to temporal vegetation characteristics can be separated from noise originating from atmospheric and viewing angle effects, cloud contamination, and other types of high frequency factors. A similar effect could be possibly achieved by applying classical hierarchical image segmentation techniques on Principal Component Analysis (PCA) components related to vegetation growth. PCA decomposition, however, does not allow the separation of different frequencies, since it is completely data dependent. It is therefore impossible to assert that a given component will reflect identical temporal properties between geographical areas, whereas FFT components always express the same specific periodicity. Additionally, this advantage of the FFT allows the accentuation of information with a specific periodicity, e.g. periods related to El Niño/Southern Oscillation (ENSO) processes like detected by Barbosa et al. (2006) and Young (2005). The periods related to these processes can be specifically selected by assigning high weights and may provide authors such as Nagai et al. (2007) a means to better assess the influence of these processes on ecosystems without noise components. Alternatively, one could argue that this selection of some periodic patterns implies an information loss and that it obscures other temporal characteristics. This represents a valid concern and, as a consequence, the segmentation results should be interpreted as segments with similar temporal properties related to the selected periodic components.

One of the limitations of the FFT, on the other hand, is the assumption that time series show a certain periodicity (i.e., that variations in the time series are repeated at an uniform time-step) and have infinite duration. Consequently, the MTHIS methodology excludes the identification of stochastic dynamics or the distinction between subsequences of time series, which can be very important for the interpretation of ecosystem processes. For example, consider a five year time series of a

grassland pixel that is burnt in the second year, but completely recovered after the third year. The MTHIS methodology will not detect this change as it assumes that variations in condition occur at the same rate in all years. The MTHIS will also not detect landcover–landuse changes between years, since it is based on an assessment of the complete time series and does not allow distinction between subsequent years. However, these changes can be derived in other strategies, e.g. by applying MTHIS to time series of individual years and considering the differences in MTHIS output between subsequent years. The development of an alternative methodology based on similarities in wavelet transforms could also serve as a solution, since the wavelet transform can be used to scale conventional Fourier components. Accordingly, it decomposes a signal in terms of both time and frequency simultaneously (Daubechies, 1990).

MTHIS application however does not only depend on the selected FFT components. The result of MTHIS will moreover be influenced by the random selection sequence of segments and by threshold values $T$ in the hierarchical segmentation process. Together they determine an arbitrary element in the segmentation sequence that will be reflected by the MTHIS output. This arbitrary element can nevertheless be minimized by correctly assigning threshold values $T$. If MTHIS is initialized with $T=x$, $2x$,..., $\infty$, where $x$ is infinitely small, the arbitrary sequence is removed, since the hierarchical stepwise segmentation algorithm of Beaulieu and Goldberg (1989) is obtained. This algorithm produces a segmentation with minimal error by allowing only the smallest merge at each segmentation run and thus guaranteeing that each segment is merged with its nearest neighbor. The major limiting factor of this algorithm is unfortunately the computing speed due to one merge per run. Therefore a balance between objectivity and speed has to be assessed by allowing multiple merges per segmentation run and a minimal arbitrariness.

$T$ values have to be selected accordingly, so they minimize the computing time and level of arbitrariness. Additional changes to the segmentation sequence of MTHIS, such as the multiple merges per pass approach as proposed by Woodcock and Harward (1992) could reduce the stochastic element even more. Future changes of the methodology should consequently focus on the implementation of these improvements.

Although the $F_k$-distance provides effectively a measure to assess the similarity between FFT components, which is impossible with the classical hierarchical segmentation methodologies, the use of the $F_k$-distance has also certain limitations. These limitations correspond to the limitations of the minimum distance to mean classifier (Lillesand & Kiefer, 2000). Both are insensitive to different degrees of variance. The use of alternative dissimilarity measures, such as the standard deviation of FFT components, causes nevertheless other difficulties, since this approach would consider both parameters of the FFT component separately. On the other hand, introduction of other similarity measures that evaluate both the variance and covariance of these parameters, such as Gaussian maximum likelihood measures, would increase computing time tremendously, since for every possible merge covariance measures need to be calculated between all pixels in neighboring segments. Given the size of the images we are dealing with in this study, this approach required far too much computing time. We therefore believe that the methodology presented here will prove useful until more sophisticated approaches for assessing the similarity between FFT components become available.

Moreover, the MTHIS methodology provides an generic methodology that can be adapted towards user requirements, for example, in remote sensing applications where the piecewise homogeneous scene model is violated. In this model it is assumed that segments comprising the landscape have both low internal variance and a common level of internal variance (Woodcock & Harward, 1992). Unfortunately, this assumption is often inadequate for images over different landcover–landuse types. This means that for an average threshold value $T$, neighboring pixels of landcover–landuse types with low internal variance will be grouped in few segments, whereas neighboring pixels of landcover–landuse types with high internal variance will not merge. Although the MTHIS methodology is partly adapted to this problem by allowing only one merge per segment per segmentation run, it is very unlikely that all the segments defined by the MTHIS output correspond to patches of the same level of the landscape hierarchy. To solve this, additional size constraints or measures of internal local variance can be incorporated by the user in the MTHIS methodology. Woodcock and Harward (1992) for example proposed the use of an additional texture channel that could also be used as additional measure in $S$.

### 6.2. MTHIS protocol

Application of the MTHIS protocol on five year VGT NDVI time series of South Africa illustrated the potential of the methodology. The average and annual terms were selected, since they represented the major spatio-temporal variability in the study area. The importance of these two terms for vegetation description corroborates the work of Azzali and Menenti (2000) and Moody and Johnson (2001). These authors discussed the average and annual term to demonstrate the usefulness of FFT components to describe vegetation phenological characteristics. They exploited the information of these terms and the sixth month FFT term to classify the vegetation of South Africa using a per-pixel iso-cluster procedure that is insensitive to nonsystematic data noise in the higher-order components.

Moreover, the selection and identification of the relevant spatio-temporal FFT components revealed the distinction in spatio-temporal variability between both the average and annual term. This distinction is critical to an understanding of vegetation dynamics, because both terms relate to different biophysical processes. The average term is mainly related to climate effects and reflects the changing amounts of vegetation cover related to rainfall per-pixel (Azzali & Menenti, 2000). The overall variability of the vegetation cover is large, ranging from almost no rainfall in the western desert, to high rainfall areas like the forest plantations in the east (Fig. 6a). Moody and Johnson (2001), on the other hand, showed that the annual term

is linked to localized effects such as structural landcover–landuse (e.g. difference between evergreen, deciduous and annual habit) with a higher relative spatial variability.

The visualization by means of the BSI allowed an interpretation of the segmentation boundaries at different scale levels. Additionally, the BSI provided an indication of the quality of the MTHIS methodology when BSI boundaries were compared with the input FFT component images. Hence, BSI boundaries are essential from a user's perspective as they allow a first analysis of the segmentation hierarchy. In this paper, the hierarchical structure of the segments was not specifically studied in detail, but it was established inherently, and can be incorporated in future research to analyze how the segments change from one scale level to another.

Comparison of the MTHIS output with the LR and LC reference maps confirmed moreover the expected agreement at a specific scale level between segmentation result and factors that describe the vegetation characteristics at this scale level. The Kappa and boundary accuracy measures also showed that this agreement was much higher for the LC map than for the LR map. This partly can be attributed to the definition and the original purpose of the reference maps. The LR map describes the biological resources from a perspective of potential natural vegetation as a functional combination of soil, physical environment, and climate, rather than actual landcover–landuse influenced by man-made transformations (Low & Rebelo, 1996). The LR maps are by definition very heterogeneous and based on vegetation structure, ecological processes and occurrence of important plant species and do not necessarily provide direct information about the dynamics of vegetation green cover and its foliar phenology. Consequently, it fails to describe the actual vegetation and is more related to potential vegetation resulting from climatological and biophysical characteristics which are often obscured by man-made alterations. The Kappa and boundary accuracy measures for the landcover–landuse map, on the other hand, confirmed the results of Moody and Johnson (2001), who showed that FFT components provide a concise and repeatable input for summarizing baseline inter-annual variability of landcover dynamics over broad regions and can be used as criteria for differentiating landcover types on the basis of temporal properties. The MTHIS methodology, however, groups the temporal properties at more scale levels, for which applicable reference maps were not readily available. It is debateable what these other scale levels represent, besides uniformity in temporal behavior, for vegetation mapping purposes. Nevertheless, certain scale levels already provide useful information, such as the difference in summer and winter rainfall regions. The description of these and all other scale levels is an important theme that should provide focus for future research on this topic. Future work should therefore concentrate on the construction of truly objective and external reference data with detailed hierarchical and spatio-temporal characteristics that effectively allow the validation of the MTHIS output at all scale levels. The use for example of artificial data sets, whose hierarchical and temporal properties can be completely controlled, could provide a great help in this context.

## 7. Conclusion

A MTHIS methodology was proposed to integrate remote sensing time series in a hierarchical image segmentation approach. MTHIS clusters adjoining pixels with similar temporal properties into hierarchical segments at various scales. Therefore, the similarity of temporal behavior was defined as similarity of FFT components and an $F_k$-distance criterion was introduced that employs the Euclidian distance between FFT components of the same frequency as similarity measure. This choice was based on the duality between the frequency domain of the FFT components and time domain of the NDVI time series. They contain exactly the same information, but in a different form. The use of FFT components and $F_k$-distance, however, allowed the elimination of components that represented little more than noise contained in the original time series, resulting in an increased computational efficiency of the segmentation methodology.

Application of the methodology in a specific MTHIS protocol for VGT NDVI time series demonstrated the concept of MTHIS. The actual MTHIS was performed on the average and annual FFT term, since these components contained the majority of the spatio-temporal variability in the NDVI time series. The selection of these components was assessed by means of the energy density spectrum and $F_k$-distance as a function of lag distance. The results of the MTHIS implementation were visualized by means of BSI overlays. These overlays provided an indication of the quality of the MTHIS. Additionally, they confirmed the usefulness of MTHIS to partition the study area at different hierarchical scale levels.

Finally, the correspondence between MTHIS results and reference layers of vegetation characteristics at different scales was assessed to determine the specific segmentation optimum where both coincide at the pre-defined management scale of the reference layer. The comparison of these optima revealed a close relationship between segmentation output and landcover–landuse reference map. This relationship was less clear for the vegetation type reference maps due to the reference map purpose and definition. On the other hand, the accuracy measures for the landcover–landuse map strengthened the findings of earlier studies with FFT components that also successfully mapped landcover types based on these components. The MTHIS methodology, however, adds much more information as its provides results at several scale levels. The description and evaluation of these other scale levels based on various data with a wide range of reference layers is however crucial for future research.

## Acknowledgements

This work was performed in the framework of a research project on satellite remote sensing of terrestrial ecosystem dynamics, funded by the Belgian Science Policy Office (GLOVEG-VG/00/01). I. Jonckheere is 2006 Laureate of the Belgian Stichting Roeping, and is currently a Postdoctoral fellow of the FWO Flanders. The SPOT VGT S10 data set was generated by the Vlaamse Instelling voor Technologisch Onderzoek (VITO), whereas the National Land Cover Map of South Africa was supplied by the Agricultural Research Council (ARC). The authors

would like to thank T. Harding and D. Rizopoulos for feedback on processing and statistical questions. We are indebted to the editor and referees for their detailed reviews that led to an improved version of the manuscript.

## References


Andres, L., Salas, W. A., & Skole, D. (1994). Fourier analysis of multitemporal AVHRR data applied to a landcover classification. *International Journal of Remote Sensing*, *15*(5), 1115−1121.

Anyamba, A., & Eastman, J. R. (1996). Interannual variability of NDVI over Africa and its relations to El Niño/Southern Oscillation. *International Journal of Remote Sensing*, *17*(13), 2533−2548.

Azzali, S., & Menenti, M. (2000). Mapping vegetation–soil–climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. *International Journal of Remote Sensing*, *21*(5), 973−996.

Baatz, M., & Schäpe, A. (2000). Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. In J. Strobl, & T. Blaschke (Eds.), *Angewandte geographische informationsverarbeitung*, *vol. xii*. (pp. 12−23)Heidelberg: Wichmann-Verlag.

Barbosa, H., Huete, A., & Baethgen, W. (2006). A 20-year study of NDVI variability over the Northeast Region of Brazil. *Journal of Arid Environments*, *67*, 288−307.

Beaulieu, J. M., & Goldberg, M. (1989). Hierarchy in picture segmentation: A stepwise optimization approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, 150−163.

Bruzzone, L., Smits, P. C., & Tilton, J. C. (2003). Foreword special issue an analysis of multitemporal remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(10), 2419−2420.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37−46.

Coppin, P., Jonckheere, I., Lambin, E., Nackaerts, K., & Muys, B. (2004). Digital change detection methods in ecosystem monitoring: A review. *International Journal of Remote Sensing*, *25*, 1565−1596.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, *36*(5), 961−1005.

Delves, L. M., Wilkinson, R., Oliver, C. J., & White, R. G. (1992). Comparing the performance of SAR image segmentation algorithms. *International Journal of Remote Sensing*, *13*(11), 2121−2149.

Eastman, J. R., & Fulk, M. (1993). Long sequence time series evaluation using standardized principal components. *Photogrammetric Engineering and Remote Sensing*, *59*(8), 1307−1312.

Evans, J. P., & Geerken, R. (2006). Classifying rangeland vegetation type and coverage using a Fourier component based similarity measure. *Remote Sensing of Environment*, *105*, 1−8.

Fan, J., Yau, D. K. Y., Elmagarmid, A. K., & Aref, W. G. (2001). Automatic image segmentation by integrating color-edge extraction and seeded region growing. *IEEE Transactions on Image Processing*, *10*, 1454−1466.

Garrigues, S., Allard, D., Baret, F., & Weiss, M. (2006). Quantifying spatial heterogeneity at the landscape scale using variogram models. *Remote Sensing of Environment*, *103*, 81−96.

Gurgel, H. C., & Fereira, N. J. (2003). Annual and interannual variability of NDVI in Brazil and its connections with climate. *International Journal of Remote Sensing*, *24*(18), 3595−3609.

Handcock, R. N., & Csillag, F. (2004). Spatio-temporal analysis using a multiscale hierarchical ecoregionalization. *Photogrammetric Engineering and Remote Sensing*, *70*, 101−110.

Hay, G. J., Blaschke, T., Marceau, D. J., & Bouchard, A. (2003). A comparison of three image-objects for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*, 327−345.

Holben, B. (1986). Characterization of maximum value composites from temporal AVHRR data. *International Journal of Remote Sensing*, *7*, 1417−1434.

Jakubauskas, M. E., Legates, D. R., & Kastens, J. H. (2001). Harmonic analysis of time series AVHRR NDVI data. *Photogrammetric Engineering and Remote Sensing*, *67*(4), 461−470.

Jakubauskas, M. E., Legates, D. R., & Kastens, J. H. (2002). Crop identification using harmonic analysis of time series AVHRR NDVI data. *Computers and Electronics in Agriculture*, *37*, 127−139.

James, G. (1994). *Advanced Modern Engineering Mathematics* (pp. 288−291). Wokingham, England: Addison-Wesley.

Jönsson, P., & Eklundh, L. (2004). TIMESAT— A program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, *30*, 833−845.

Juarez, R. I. N., & Liu, W. T. (2001). FFT analysis on NDVI annual cycle and climatic regionality in Northeast Brazil. *International Journal of Climatology*, *21*, 1803−1820.

Justice, C. O., Townshend, J. R. G., Holben, B. N., & Tucker, C. J. (1985). Analysis of the phenology of global vegetation using meteorological satellite data. *International Journal of Remote Sensing*, *6*(8), 1271−1318.

Lee, R., Yu, F., Price, K. P., Ellis, J., & Shi, P. (2002). Evaluating vegetation phenological patterns in Inner-Mongolia using NDVI time-series analysis. *International Journal of Remote Sensing*, *23*(12), 2505−2512.

Lillesand, T. M., & Kiefer, R. W. (2000). *Remote sensing and image interpretation* (pp. 538−539). New-York: John Wiley & Sons.

Low, A. B., & Rebelo, A. G. (1996). Vegetation of South Africa, Lesotho, and Swaziland. Pretoria: Dept. Environmental Affairs and Tourism.

Lucieer, A., & Stein, A. (2002). Existential uncertainty of spatial objects segmented from satellite sensor imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(11), 2518−2521.

Moody, A., & Johnson, D. M. (2001). Land-surface phenologies from AVHRR using the discrete Fourier transform. *Remote Sensing of Environment*, *75*, 305−323.

Nagai, S., Ichii, K., & Morimoto, H. (2007). Interannual variations in vegetation activities and climate variability caused by ENSO in tropical rainforests. *International Journal of Remote Sensing*, *28*(6), 1285−1297.

Olsson, L., & Eklundh, L. (2001). Fourier series for analysis of temporal sequences of satellite sensor imagery. *International Journal of Remote Sensing*, *15*(18), 3735−3741.

Rahman, H., & Dedieu, G. (1994). SMAC: A simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing*, *15*, 123−143.

Reed, B. C., Brown, J. F., Vanderzee, D., Loveland, T. S., Merchant, J. W., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, *5*, 703−714.

Running, S. W., Loveland, T. R., & Pierce, L. L. (1994). A vegetation classification logic based on remote sensing for use in global biogeochemical models. *Ambio*, *23*, 77−81.

Singleton, R. C. (1969). Algorithm for computing the mixed radix fast Fourier transform. *I.E.E.E. Transactions on Acoustics, Speech, and Signal Processing*, *17*, 93−103.

Smith, S. W. (1999). *The scientist and engineer's guide to digital signal processing* (pp. 161−165). San Diego, CA: California Technical Publishing.

Stuckens, J., Coppin, P. R., & Bauer, M. E. (2000). Intergrating contextual information with per-pixel classifications for improved land cover classifications. *Remote Sensing of Environment*, *71*, 282−296.

Thompson, M. W. (1999). South African national landcover database project, data users manual: Final report (phases 1, 2, and 3). *Client report ENV/P/C 98136* Pretoria: CSIR.

Tilton, J. C., & Lawrence, W. T. (2000). Interactive analysis of hierarchical image segmentation. *Proceedings of the IGARSS*, 733−735.

Verbesselt, J., Jönsson, P., Lhermitte, S., van Aardt, J., & Coppin, P. (2006). Evaluating indices derived from satellite and climate data as fire risk indicators in savanna ecosystems. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(6), 1622−1632.

Woodcock, C., & Harward, V. J. (1992). Nested-hierarchical scene models and image segmentation. *International Journal of Remote Sensing*, *13*, 3167−3187.

Wu, J., & Loucks, O. L. (1995). From balance-of-nature to hierarchical patch dynamics: A paradigm shift in ecology. *Quarterly Review of Biology*, *70*, 439−466.

Young, S. S., & H.R. (2005). Changing patterns of global-scale vegetation photosynthesis, 1982–1999. *International Journal of Remote Sensing*, *20* (20), 4537−4563.