Remote Sensing of Environment 169 (2015) 375–389

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

A time series processing tool to extract climate-driven interannual vegetation dynamics using Ensemble Empirical Mode Decomposition (EEMD)

Pieter Hawinkel [a,b,*], Else Swinnen [a], Stef Lhermitte [b], Bruno Verbist [b], Jos Van Orshoven [b], Bart Muys [b]

[a] Flemish Institute for Technological Research (VITO), Remote Sensing Unit, Boeretang 200, B-2400 Mol, Belgium
[b] KU Leuven – University of Leuven, Department of Earth and Environmental Sciences, Celestijnenlaan 200E, B-3001 Heverlee, Belgium

ARTICLE INFO

Article history:
Received 24 March 2015
Received in revised form 4 August 2015
Accepted 22 August 2015
Available online xxxx

Keywords:
Time series analysis
Vegetation monitoring
Interannual variability
Ensemble Empirical Mode Decomposition
NDVI
Longitudinal data

ABSTRACT

Interannual changes of vegetation are crucial in understanding ecosystem dynamics under global change. However, there is no automated tool to extract these interannual changes from remote sensing time series. To fill this gap, the Ensemble Empirical Mode Decomposition (EEMD) framework was refined and implemented to decompose time series of Normalized Difference Vegetation Index (NDVI) and reconstruct their interannual components. The performance of EEMD-based interannual NDVI detection was assessed using simulated time series, and its sensitivity to model and data parameters was determined to provide a basis for remote sensing applications. The sensitivity analysis highlighted application limitations for time series with low interannual to annual amplitude ratios and high irregularity in timing of growing seasons, as these factors have the strongest effects on the overall performance. However, within these limitations, the detected interannual components correspond well to simulated input components with respect to timing of episodes and composition of time scales. The applicability on real world NDVI time series was demonstrated by mapping the coupling between precipitation variability, interannual vegetation changes, and the El Niño Southern Oscillation and Indian Ocean Dipole phenomena for ecoregions in East and Central Africa. In most areas where precipitation was found sensitive to oceanic forcing, the EEMD detected vegetation changes matched the predicted response, except in dense forest ecosystems.

1. Introduction

Interannual changes in vegetation cover and productivity are key to understanding the impacts of global climate change on ecosystem dynamics (Hilker et al., 2014; Luo et al., 2011). Such year to year variations in vegetation status critically influence ecosystem services such as terrestrial carbon sequestration (Piao et al., 2011) and regulation of the hydrological cycle (Liu et al., 2008). However, the response of vegetation to climatic anomalies is often ambiguous (Brando et al., 2010), variable across ecosystems (Holmgren, Hirota, van Nes, & Scheffer, 2013; Piao et al., 2014) and subject to critical transitions at tipping points (Hirota, Holmgren, Van Nes, & Scheffer, 2011). Therefore its quantification requires dedicated monitoring tools.

Satellite remote sensing is an excellent tool to assess land ecosystem dynamics as it provides periodic, spatially explicit and consistent measurements of biophysical processes at the Earth's surface (Baynard, 2013). It has become an established monitoring instrument in the fields of forestry (Achard et al., 2002), agriculture (Atzberger, 2013) and plant ecology (Xie, Sha, & Yu, 2008). Several vegetation indices and biophysical parameter products have been derived over the past 35 years from different satellite sensors, e.g., NOAA-AVHRR (Cracknell, 2001), SPOT-VGT (Maisongrande, Duchemin, & Dedieu, 2004), MODIS (Huete et al., 2002) and PROBA-V (Dierckx et al., 2014). A key product in these datasets is the Normalized Difference Vegetation Index (NDVI) which highlights vegetation greenness and photosynthetic activity (Tucker & Sellers, 1986). The combined data archives form a unique long term NDVI record which has been used in numerous studies dealing with the effects of climate variability on vegetation status in space and time (e.g., Fensholt et al., 2012; Guo et al., 2014).

This paper proposes and validates a framework for a tool to detect interannual NDVI fluctuations, by considering the ensemble of past efforts in remote sensing data processing and introducing promising advances from another domain of science, i.e., signal processing.

Early approaches accounted for interannual variability in time series *implicitly* without explicitly separating the interannual component. In pioneer works (Anyamba & Eastman, 1996; Nicholson & Farrar, 1994; Richard & Poccard, 1998), the forcing role of climate variability on NDVI was explored by spatial representations of the (lagged) correlation between mean monthly or annual NDVI values and corresponding aggregate rainfall values. This principle of lagged linear response was

* Corresponding author at: Flemish Institute for Technological Research (VITO), Remote Sensing Unit, Boeretang 200, B-2400 Mol, Belgium.
E-mail address: pieterhawinkel@gmail.com (P. Hawinkel).

http://dx.doi.org/10.1016/j.rse.2015.08.024
0034-4257/

further explored in a simple linear regression approach (Camberlin, Martiny, Philippon, & Richard, 2007; Zhou, Van Rompaey, & Wang, 2009; Zhou et al., 2014a), whereas other authors (De Keersmaecker et al., 2015; Myneni, Los, & Tucker, 1996; Plisnier, Serneels, & Lambin, 2000) introduced the use of standardized anomalies from the average annual season as the key to reveal interannual variability. Brown, de Beurs, and Vrieling (2010) elaborated this notion of climatic forcing at multiple interannual time scales, supported by an increasing insight in the mechanisms and the effects of regional climate oscillations (Holmgren, Scheffer, Ezcurra, Gutiérrez, & Mohren, 2001; Nicholson & Kim, 1997; Williams & Hanan, 2011).

Other tools have been developed to *explicitly* separate different time scale components in NDVI time series using techniques from the signal processing domain. Methods range from Fourier spectral analysis (Immerzeel, Quiroz, & De Jong, 2005; Lhermitte et al., 2008; Roerink, Menenti, Soepboer, & Su, 2003) and wavelet analysis (Torrence & Compo, 1998) to specific tools to detect, separate and model intra-annual variability, long term trends and abrupt changes from satellite image time series: TIMESAT (Jönsson & Eklundh, 2002), BFAST (Verbesselt, Hyndman, Zeileis, & Culvenor, 2010) and SPIRITS (Eerens et al., 2014).

Despite this wide range of tools, there is no quantitative method that focuses on interannual fluctuations, which we believe are crucial in understanding and predicting the response of ecosystems to an increasingly variable climate. For example, Fourier analysis methods generally focus on mapping the occurrence of annual unimodal or bimodal growing seasons to distinguish land use patterns (e.g., Lhermitte, Verbesselt, Verstraeten, & Coppin, 2011), while longer time scales are only represented as harmonic frequencies which are not variable over time. Wavelet analysis provides a more flexible technique to detect interannual components which are themselves variable over time, beyond the static and strictly harmonic Fourier framework. It has therefore been applied in numerous studies to identify interannual components in NDVI and climatic time series (Galford et al., 2008; Martínez & Gilabert, 2009; Quiroz, Yarlequé, Posadas, Mares, & Immerzeel, 2011). Swinnen (2008) used wavelet coherency, a form of cross-spectral analysis (Torrence & Webster, 1999) to systematically examine the climate-vegetation coupling over all occurring time scales in the time-frequency domain.

However, a generic tool to study interannual NDVI components must yield easily interpretable components in the time domain, so as to not impose constraints for further processing (e.g., co-analysis with climatic series), and to allow integrated use with other time series toolboxes. An answer to the need of breaking down an NDVI series into different time scale components while preserving the time domain flexibility of the wavelet approach is found in the Empirical Mode Decomposition (EMD) methodology, proposed and elaborated by Huang and colleagues in their key publication (Huang et al., 1998). The EMD algorithm (summarized in Section 2.2, and detailed in Appendix A) iteratively extracts the intrinsic time scales from the series, yielding a finite set of components with decreasing frequency and a residual trend component. EMD has been applied to climatic or biogeophysical time series to analyze climate variability in general (Coughlin & Tung, 2005; Molla, Ghosh, & Hirose, 2011; Pegram, Peel, & McMahon, 2008), to model slow components in climate simulations (Brisson, Demuzere, Willems, & van Lipzig, 2015), to assess climatic effects on plant phenology (Guan, 2014), to aid crop classification (Chen, Son, Chang, & Chen, 2011) and to remove the effect of platform orbital drifts in cross-sensor image time series (Pinzon, Brown, & Tucker, 2005).

EMD and its extension Ensemble Empirical Mode Decomposition (EEMD; Wu & Huang, 2009) are straightforward as a decomposition algorithm, except for the procedure to sort ensembles of components lacking explicitness, which we tackle in this paper. However, processing EEMD components to robust indicators of interannual variability has not been addressed yet and faces a number of problems. First, the capability of assigning components to their underlying physical processes (e.g., noise, annual climatology) has been demonstrated ad hoc (Coughlin & Tung, 2005), but has not been formalized to make it applicable for systematic processing of large spatiotemporal datasets. Second, distinguishing significant signal components from noise-induced components critically determines the method's performance and the conclusions that can be drawn from its output. Significance tests for (E)EMD components exist but lack validation, which is a prerequisite for a generic processing tool.

The overall aim of this work is to introduce the EEMD technique in remotely sensed image processing applications, and to propose a framework for its implementation and validation. More specifically, the objectives of this paper are:

- To define a robust procedure based on EEMD to separate interannual components in remotely sensed NDVI time series from annual and noise components;
- To assess the performance of this procedure and its sensitivity to EEMD model parameters and characteristics of the input time series;
- To demonstrate its applicability in a case study over East and Central Africa.

Based on the hypothesis that climate variability is reflected in NDVI time series as relatively weak, irregular and slow fluctuations, we composed simulated time series containing annual, interannual and noise components. In this way, we set up a validation framework in which the interannual component detected after decomposition can be compared to the original input component. After this validation, we assess the representativeness of our simulations for real world situations by examining the detected interannual NDVI response to precipitation in East and Central Africa. In the areas where precipitation is sensitive to regional climate variability, a corresponding vegetation response is expected in ecosystems with limiting precipitation.

2. Data and methods

A remotely sensed NDVI dataset spanning the 1981–2014 period and covering a wide range of climate types and ecosystems (East and Central Africa) along with series of historic climate indices were used to set up a framework to test and validate the envisaged tool for interannual component detection. We adopted the Ensemble Empirical Mode Decomposition (EEMD) to obtain a per-pixel estimate of the constituent components of NDVI series. These estimates were used to simulate annual, interannual and noise components with controlled variation of their parameters. A set of mixtures of simulated components were decomposed using EEMD, and the interannual components were reconstructed by three distinct approaches. Measures of correspondence between the input and output interannual components were calculated to evaluate the performance and sensitivity to various EEMD model parameters, data characteristics and reconstruction methods. Finally, the validated method was applied on the original NDVI dataset and historic climate indices in a regional case study.

2.1. Data

2.1.1. Remote sensing datasets

The Normalized Difference Vegetation Index (NDVI) is defined as the normalized difference in surface reflectance (*Refl*) in the near-infrared (*NIR*) and red (*R*) wavebands of a sensor and is considered to be a proxy for vegetation greenness (Tucker & Sellers, 1986).

$$\mathrm{NDVI} = (Refl_{NIR} - Refl_R)/(Refl_{NIR} + Refl_R) \quad (1)$$

Various efforts have been made to produce consistent long term NDVI records from sensor sequences, each with their particular merits and flaws (Tian et al., 2015). Within the setup of this paper, residual inconsistencies in the baseline dataset do not fundamentally affect the

data simulations or the validation results produced from them. Therefore, this section is limited to a description of the underlying datasets whereas the implications of data inconsistencies are discussed at length in the discussion section of this paper.

NASA's Long Term Data Record (LTDR) contains the reprocessed daily global AVHRR imagery from subsequent NOAA platforms (1981–1999 in version 2; extended with 2001–2013 in version 4), at 0.05° resolution. The LTDR reprocessing aims to achieve data consistency through both improved radiometric calibration (Vermote & Kaufman, 1995), corrections for orbital drift effects (Nagol, Vermote, & Prince, 2014; Vermote, Justice, & Breon, 2009) and atmospheric corrections with ancillary meteorological data (Pedelty et al., 2007).

The SPOT-VEGETATION (VGT) program (1998–2014) delivered global NDVI images of 1 km resolution as ten-daily maximum NDVI synthesis products as to minimize cloud effects present in daily recordings (Deronde et al., 2014). Residual clouded pixels typically cause sharp troughs in temporal NDVI profiles. These are detected and their values replaced by means of an NDVI interpolation algorithm (Eerens et al., 2014; Swets, Reed, Rowland, & Marko, 1999) to represent a more realistic smooth temporal NDVI profile.

One approach to obtain a long uninterrupted baseline dataset is to merge the VGT archive (1998–2014) with the first part of the LTDR archive (1981–1999). Both image datasets were resampled to a common systematic 20 km frame and subjected to the same synthesis and smoothing procedures to yield comparable NDVI series. Finally, cross-sensor calibration reduces the inconsistencies between NDVI from the AVHRR and VGT sensors caused by differences in sensor spectral responses (Trishchenko, Cihlar, & Li, 2002), native spatial resolutions (Tarnavsky, Garrigues, & Brown, 2008), platform orbits and sensor geometries (Swinnen & Veroustraete, 2008). Empirical cross-sensor calibration was achieved following the linear VGT-to-AVHRR correction model proposed by Steven, Malthus, Baret, Xu, and Chopping (2003). We re-estimated the correction equation from corresponding pixels in images from 1999, applying geometric mean regression (Ji & Gallo, 2006).

$$NDVI_{AVHRR} = 0.799\ NDVI_{VGT} + 0.081 (RMSE = 0.069\ NDVI) \quad (2)$$

2.1.2. Climate data

Interannual climatic variability is represented in this study by two climatological indices which have been linked to interannual rainfall patterns in East and Central Africa (IPCC, 2007). The Oceanic Niño Index (ONI) (NOAA-CPC, 2014) represents sea surface temperature anomalies in the Pacific Ocean (ERSST v4; Huang et al., 2014) and is the most widely used indicator for the episodes of the El Niño Southern Oscillation (ENSO), which are known to have impacts globally. In addition, East Africa's rainfall is largely driven by the Indian Ocean dynamics, which as well display interannual episodes as summarized in the Indian Dipole Mode Index (DMI) (JAMSTEC, 2010). This is the difference in sea surface temperature anomalies between the western and eastern equatorial Indian Ocean. Monthly values for both indices were smoothed with a 12-month running window in order to obtain an estimate of the interannual climatic forcing (Fig. 1).

Gridded precipitation estimates derived from satellite observations (1983–2012) are obtained from the PERSIANN system (Sorooshian et al., 2000). The PERSIANN system merges infrared brightness images with rainfall estimates from satellite microwave data to daily 0.25° precipitation maps, with error detection and quality control using ground observations (Hsu, Gupta, Gao, & Sorooshian, 1999).

2.2. Ensemble EMD decomposition model

For a detailed description of the basic EMD model, we refer to Huang et al. (1998), whereas the algorithmic implementation and its parameters are described in Appendix A. Briefly, EMD works by connecting

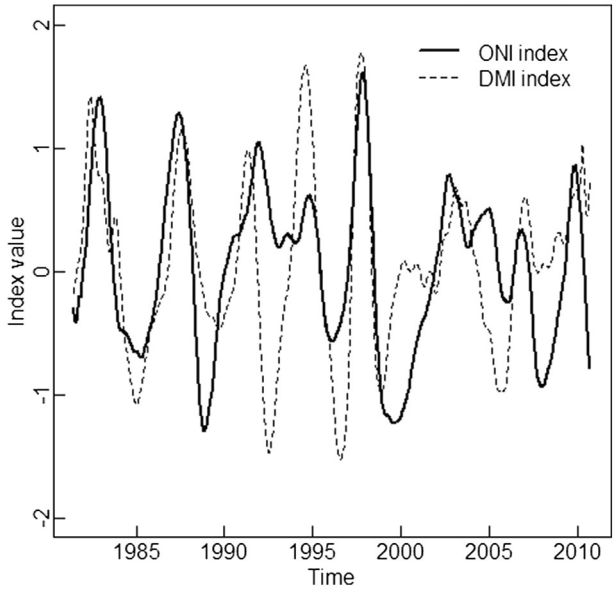

Fig. 1. The Oceanic Niño Index (ONI) and the Dipole Mode Index (DMI) are derived from sea temperature measurements and represent the positive and negative episodes of climate oscillations that affect interannual rainfall patterns over East and Central Africa (IPCC, 2007).

local maxima (minima) in a series X (t) by a cubic spline which yields upper (lower) envelopes. The time series minus the mean of the upper and lower envelopes contains the high frequency information. In a process termed *sifting*, the above steps are repeated to remove residual low frequency information until the number of extrema and zero-crossings does not change for S consecutive siftings. The outcome is a component of the original series that isolates a specific frequency of the series, termed an Intrinsic Mode Function (IMF). The residual of the data and a detected IMF is again subjected to the sifting process to iteratively yield a set of k IMFs ($IMF_{i=1\ldots k}$) with decreasing frequency, until a final residual term R (t) has less than two local extrema.

$$X(t) = \sum_i IMF_i(t) + R(t) \quad (3)$$

Although powerful, basic EMD often leads to one time scale intermittently partitioned over two or more IMFs, an issue referred to as *mode mixing* (Huang, Shen, & Long, 1999), illustrated in Fig. 2a. To cope with mode mixing, Wu and Huang (2009) proposed Ensemble EMD (EEMD) as an extension to the method. By iteratively adding a finite amount (σ_{add}) of Gaussian white noise to the series, applying EMD to the mixture, grouping corresponding IMFs into ensembles, and averaging each ensemble of IMFs over a large number N_n of noise realizations, mode mixing is largely eliminated (Fig. 2b).

Although EEMD was applied recently by various authors (e.g., Feng, Liang, Zhang, & Hou, 2012; Guan, 2014; Kuo, Wei, & Tsai, 2013; Zhou, Jiang, et al., 2014b), no formal definition of *corresponding* IMFs, required to sort them into ensembles, has been given. However, automated EEMD analysis of a large array of per-pixel NDVI time series requires a robust sorting procedure, certainly because different noise realizations can yield slightly different numbers of IMFs, and assigning each IMF to its appropriate group (or *bin*) before averaging is crucial in attaining successful mode separation. To formalize the IMF sorting step, we propose to use IMF periods p_i (i.e., the series length divided by twice the number of zero-crossings in IMF_i) as the basic criterion, where the EMD decomposition of the original series without added noise defines reference periods $p_{ref.1},\ldots,p_{ref.k}$. Since EMD is found to produce components with a general pattern of doubling (or exponentially growing) periods (Flandrin, Rilling, & Goncalves, 2004; Wu & Huang, 2004), the geometric

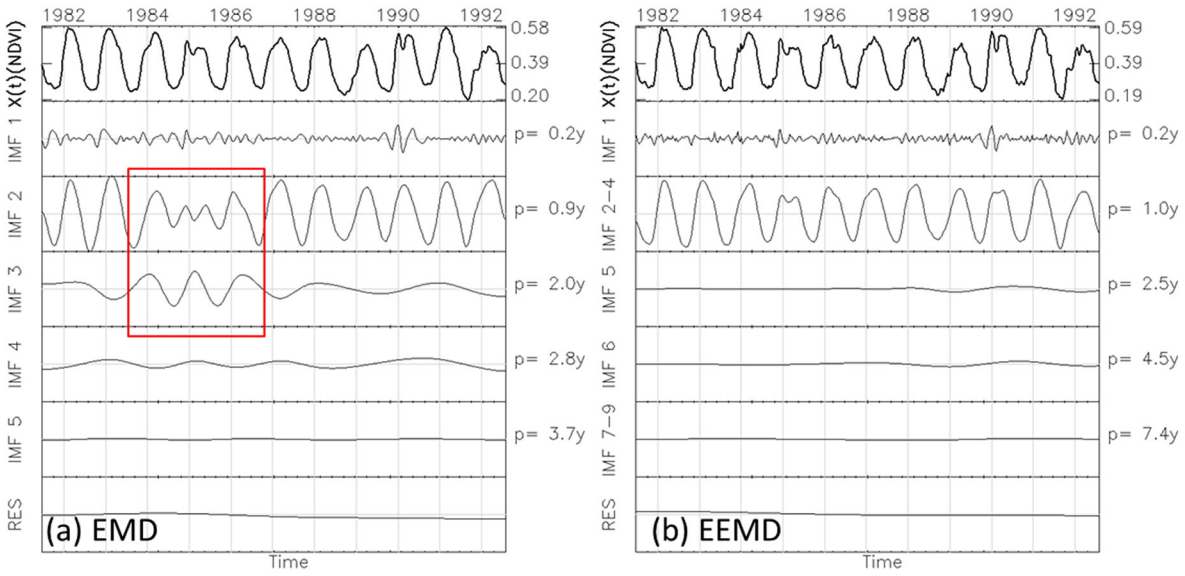

Fig. 2. A temporal NDVI profile (pixel 7 in Fig. 3, North Tanzania) and its Intrinsic Mode Functions by (a) EMD decomposition and (b) the improved EEMD decomposition. EEMD decreases 'mode mixing' whereby fragments of a single time scale occur in another component, e.g. for the annual time scale for the period 1984–1986 (box on graph (a)).

means of subsequent reference periods $p_{ref.j}$ are considered as the bin boundaries for assigning IMFs to a bin after each noise realization.

$$bin_j = \left[\sqrt{p_{ref.j-1}p_{ref.j}}, \sqrt{p_{ref.j}p_{ref.j+1}}\right] \quad (4)$$

The sorting algorithm thus evaluates the IMFs of each noise-distorted ensemble member, and assigns them to their proper predefined period bins. The IMFs in a bin are then averaged over the number of noise realizations N_n, to yield a decomposition with better mode separation. Our implementation uses parameter values for added noise amplitude [$\sigma_{add} = 0.1$] as in other EEMD applications (Kuo et al., 2013; Wu & Huang, 2009) and [$N_n = 50$] since the slight differences in outcomes for higher N_n did not offset linearly increasing processing time.

2.3. Simulated time series

Our central hypothesis is that observed NDVI time series are composed of an annual component encompassing one or two growing seasons, an interannual component including climatic influence as well as residual effects of sensor merging and orbital drift, and a high-frequent noise component reflecting residual sensor noise, radiometric and geometric errors (Lhermitte et al., 2011). Each of these components can be present in varying amplitude levels and under varying degrees of irregularity in timing. For the sensitivity analysis of the EEMD model's performance to be representative and the detection method to be generic, these forms of variability must be parameterized and the parameter ranges be estimated from a real NDVI dataset, here the intercalibrated AVHRR/VGT dataset described in Section 2.1.1.

2.3.1. Study area stratification

Fig. 3 shows the study area represented by each pixel's mean NDVI value (μ_{tot}) over the entire series, thereby revealing the major ecoregions in East and Central Africa. The series of 9 pixels with distinct NDVI levels and seasonal profiles were selected across climatic regions to optimally explore parameter ranges related to NDVI data. Fig. 4 depicts cross-calibrated temporal NDVI profiles derived from the combined AVHRR and VGT datasets for this set of pixels.

2.3.2. Estimation of variability in annual, interannual and noise components

Without any prior knowledge of relative signal amplitudes in real NDVI time series, a first estimate was made from their respective IMFs. The IMFs resulting from EEMD of each of the nine series were grouped into noise IMFs, annual IMFs and interannual IMFs, following the bin boundary logic as described in Section 2.2. Considering the occurrence of bimodal growing seasons, 0.5 year and 1 year are taken as the theoretical periods for annual cycles. As IMFs normally fall within period bins of doubling frequencies, the bin boundaries are set at $0.5/\sqrt{2}$ years and $1.0\ \sqrt{2}$ years. Shorter period IMFs are labeled as noise, whereas the sum of longer period IMFs is a first representation of the interannual component. The amplitude of each component is calculated as its standard deviation (σ). Furthermore, mean NDVI values (μ_{tot}) and the numbers of modes in the average yearly growing season were tabulated per series (Table 1). Three prototypes of NDVI seasonality were identified for further modeling: a weak bimodal season with high NDVI $^{(**H)}$ (humid conditions, e.g., East DR Congo), strongly seasonal unimodal season $^{(*)}$ (subhumid conditions, e.g., North Tanzania), and an erratic bimodal season with low NDVI $^{(**L)}$ (semi-arid conditions, e.g., East Ethiopia).

2.3.3. Annual component

Growing seasons generally appear as peaks in NDVI time series, with a certain degree of irregularity. A flexible parametric model for a seasonal NDVI peak was adopted from Jönsson and Eklundh (2002). Its basic approach is to fit an asymmetric Gaussian model function $f(t)$, defined by its base level b, peak position p and amplitude A, and the width $w_{l,r}$ and flatness $f_{l,r}$ of both tails (see Eqs. (5) and (6)). This approach was expanded to the case of bimodal growing seasons which are modeled as two additive peaks with different peak locations within the year. A unimodal or bimodal peak was thus fitted to each of the 31 growing seasons of the three prototype series spanning 1981 to 2012. Sample means and standard deviations (n = 31) give an account for the expected values and natural variation of the seasonal parameters throughout the study area (Table 2). These sample statistics were used to generate simulated annual components as a function of time $f(t)$ with realistic shapes (base level, amplitude, peak widths and flatness) and controlled variation in irregularity (peak positions).

$$f(t) = b + A.g(t) \quad (5)$$

$$\text{with } g(t) = \begin{cases} exp\left[-((t-p)/w_l)^{f_l}\right]; if\ t>p \\ exp\left[-((t-p)/w_r)^{f_r}\right]; if\ t<p \end{cases} \quad (6)$$

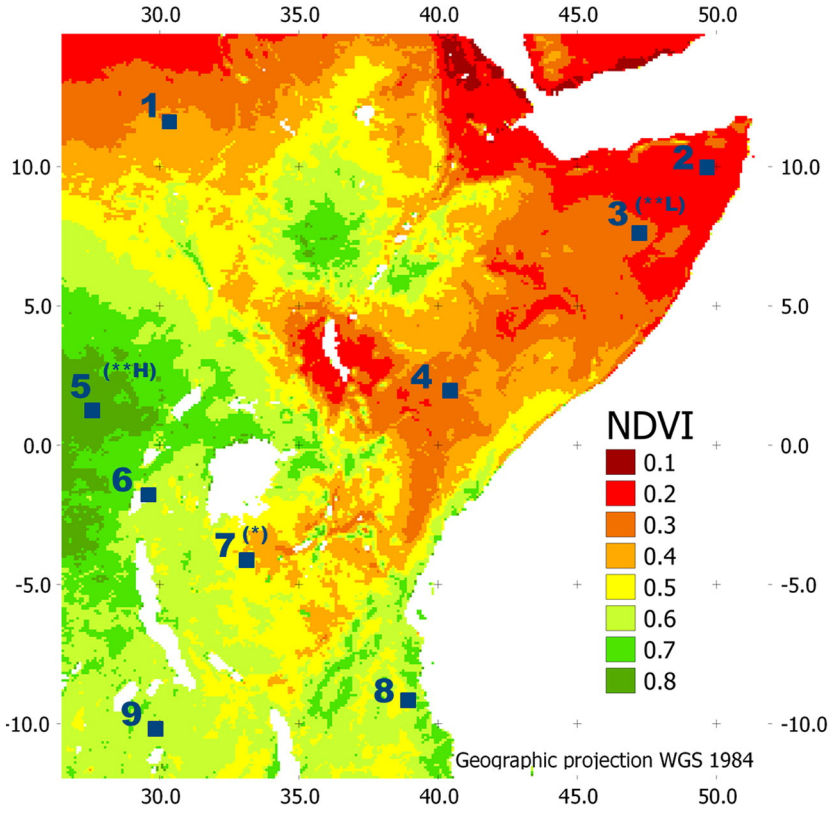

Fig. 3. Mean NDVI value over the period 1981–2012 per image pixel over East and Central Africa. The square symbols indicate the locations of the selected pixels, used to estimate parameters for NDVI time series simulation (Table 1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For each of the seasonality types, sequences of growing seasons were generated with their respective mean values of base level, amplitude and peak positions, widths and flatness. Seasonal irregularity (φ_{seas}) was introduced by adding a random variation of 0.0, 0.5, 1.0 and 1.5 times the standard deviations for the respective peak positions.

2.3.4. Interannual component

Realistic and representative a priori simulations of interannual variability are more difficult to obtain, since they are the outcome our tool aims to deliver. Therefore we combine descriptions found in literature on the impact of interannual climate variability on vegetation with

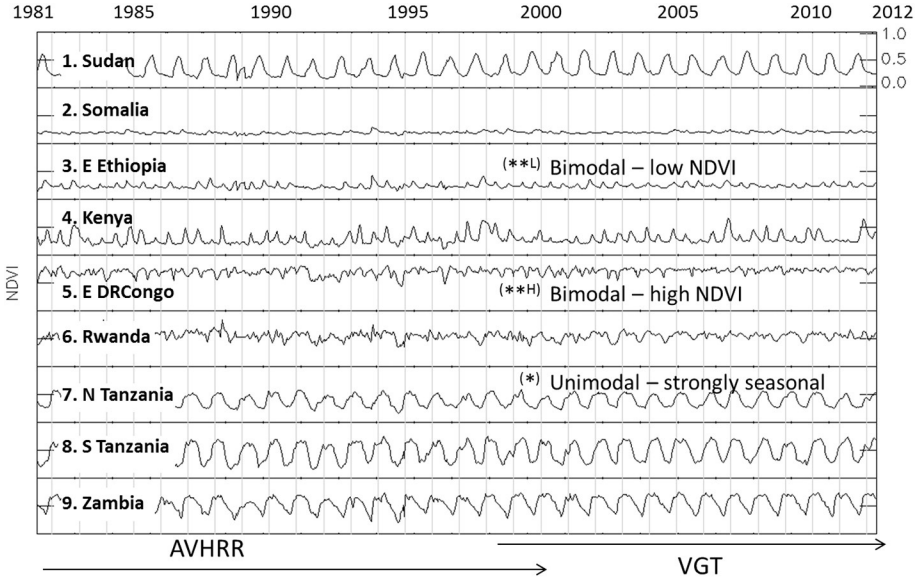

Fig. 4. Time series of the Normalized Difference Vegetation Index (NDVI) for 9 pixels in distinct climate zones in East and Central Africa (Fig. 3), derived from cross-calibrated image data from the NOAA-AVHRR (1981–1999) and SPOT-VGT (1998–2012) satellite sensors. Three prototypes of growing seasons were selected for data simulations: (*)unimodal (one rainy season) and (**)bimodal (two rainy seasons) with either overall low (L) or high (H) NDVI levels.

Table 1
Estimated parameters for the time series of 9 selected pixels: amplitudes of noise (σ_{noise}), annual (σ_{ann}) and interannual (σ_{int}) components, total amplitude (σ_{tot}) and mean value (μ_{tot}) of the series, and the number of modes in the annual growing season. Three prototype seasons for further modeling are labeled as unimodal $^{(*)}$ with strong seasonality and bimodal $^{(**)}$ with low $^{(L)}$ and high $^{(H)}$ NDVI.

Location	σ_{tot}	σ_{noise}	σ_{ann}	σ_{int}	$\sigma_{int}/\sigma_{ann}$	μ_{tot}	#modes
1. Sudan	0.14	0.02	0.12	0.03	0.27	0.35	1
2. Somalia	0.02	0.01	0.01	0.02	1.09	0.20	2
3. E Ethiopia $^{(**L)}$	0.04	0.02	0.03	0.02	0.66	0.24	2
4. Kenya	0.09	0.01	0.07	0.05	0.64	0.29	2
5. E DRCongo $^{(**H)}$	0.06	0.03	0.04	0.03	0.74	0.70	2
6. Rwanda	0.07	0.03	0.04	0.03	0.69	0.54	2
7. N Tanzania $^{(*)}$	0.10	0.01	0.10	0.02	0.19	0.39	1
8. S Tanzania	0.17	0.02	0.15	0.03	0.22	0.47	1
9. Zambia	0.12	0.03	0.11	0.03	0.23	0.53	1

our additive hypothesis and with the estimated levels of variability from the model series set of our study area. The major coupling between climatic variability and vegetation phenology is thought to act through episodes of severely increased or decreased precipitation (Holmgren et al., 2013). One well-known cause of such episodes is the El Niño Southern Oscillation (ENSO), a recurrent climate phenomenon at a 3 to 7-yearly time scale (IPCC, 2007) which is known to force vegetation growth beyond its seasonal cycle (Holmgren et al., 2013). Although the interactions are complex in nature and region dependent (Sitters, Holmgren, Stoorvogel, & López, 2012), increased precipitation episodes generally cause a rapid, but lasting response over several years (Holmgren et al., 2001).

Within the limitations of our additive model, we thus represent climate-driven interannual variability in NDVI as a weak modulation, with a steep increase and a relative slow decline. We employ again the asymmetric Gaussian simulation model to control the corresponding parameters. ENSO peak positions (t_{ONI}) and peak-to-peak amplitude variations (A_{ONI}) were derived from the historic ONI series, to reconstruct historic ENSO episodes and estimate the natural variability in duration of episodes ($\mu_{t.ONI}$; $\sigma_{t.ONI}$) for the simulation of irregular episodes with relative magnitudes ranging from 1 to 3 (R_{1-3}). Table 3 lists the resulting set of interannual models. The overall interannual amplitude levels (A_{int}) were defined as fractions of the annual amplitude ($\sigma_{int}/\sigma_{ann}$), which were in turn estimated from the interannual IMFs obtained by EEMD decomposition of the NDVI model series set. Based on the range of estimated annual/interannual ratios in Table 1, this parameter takes on values 0.1, 0.2, 0.33, 0.5 and 1.0 in the simulations.

2.3.5. Noise component

Real world NDVI noise levels (σ_{noise}) were estimated as the root mean square (*rms*) amplitudes of the sum of high frequency IMFs of each of the model set series (Table 1), as these IMFs are assumed to contain nearly all noise fragments. In order to assess the sensitivity of EEMD performance to noise in the input time series, we added random noise with *rms* amplitudes 0.00, 0.01, 0.02, 0.03 and 0.05 NDVI. The noise model present in real NDVI series is unknown a priori. As Mann and Lees (1996) argue, geophysical phenomena display a certain degree of persistence of random distortions, quantified as the noise lag-1 autocorrelation (AC_1) or *redness*. Therefore we employed both white noise ($AC_1 = 0$) and red noise ($AC_1 = \{0.25, 0.50\}$) simulations drawn from a normal distribution.

The full set of simulated series was generated by adding each combination of annual, interannual and noise component, varying over 3 seasonality types, 4 levels of seasonal irregularity (φ_{seas}), 7 interannual models, 5 interannual to annual amplitude ratios ($\sigma_{int}/\sigma_{ann}$), 3 noise models ($AC_1$) and 5 noise levels ($\sigma_{noise}$). Sensitivity of the method's performance to reconstruct interannual components was not only tested for data-related variations, but also for a method-related parameter, as described in the next paragraph.

2.4. Components post-processing

The general approach to reconstructing the interannual component from EEMD components is to add combinations of detected longer-than-annual components. Estimation of components' periods based on the number of zero-crossings was outlined above. A currently unresolved issue in the post-processing approach is the selection of *significant* interannual components. Three different approaches to selection have been evaluated through sensitivity analysis.

The need for component selection originates from the potential occurrence of low frequency components in the decomposition of random noise present in the data. As discovered by Flandrin et al. (2004) and further examined by Wu and Huang (2004), (E)EMD breaks down white noise into components of approximately doubling periods. However, the energy (squared amplitude) of a white noise (E)EMD component is shown to be related to its period, and constrained by theoretical confidence bounds (see Appendix B). Monte Carlo verification of this relationship and its confidence bounds deliver a framework to distinguish a real signal component from noise-induced components, by checking whether its energy exceeds with high confidence the expected energy of a noise component with the same period. This framework is proposed by Wu and Huang (2005) and has been adopted for climatological (E)EMD applications (e.g., Coughlin & Tung, 2005). The signal/noise distinction is however entirely dependent on the quantification of the noise level in the series. An 'a priori' test considers the entire series as potential noise, while an 'a posteriori' test can be performed by estimating the noise level from the high-frequent (E)EMD components which, under the additive hypothesis, are purely noise-induced (Wu & Huang, 2005). However, we argue that a given (E)EMD component may simultaneously contain signal information and noise information with similar periods. Therefore, exclusion as well as inclusion will introduce some error in the reconstructed interannual component. In case of sufficiently signal-to-noise ratios in the data, this balance may be in favor of retaining all detected components without significance test. We therefore consider 'no test', 'a priori' test and 'a posteriori' test in the sensitivity analysis for the reconstruction of interannual components from simulated NDVI time series.

2.5. Performance measures

Successful detection and reconstruction of interannual components from a mix of known input components, and its sensitivity to data-related and model-related parameters must be expressed in quantitative measures. We propose three different performance measures to highlight different types of correspondence between two time series (Lhermitte et al., 2011), in this case the input I_t and reconstructed output O_t interannual components.

Correlation *r* between two time series indicates whether the detected pulses are correctly localized in time and have consistently the

Table 2
Sample statistics from time series of 3 prototype growing seasons, denoted as mean (standard deviation). Five parameters represent the shape and timing of the first (I) and second (II) growing season: base level *b*, peak position *p* and amplitude *A*, and the width $w_{l,r}$ and flatness $f_{l,r}$ of the tails (as in Jönsson & Eklundh, 2002; see Eqs. (5) and (6).

Seasonality type	b (*NDVI*)	A (NDVI)		p (years)		w_1(years)		w_r(years)		f_l		f_r	
		I	II	I	II	I	II	I	II	I	II	I	II
**L Bimodal (low NDVI)	0.17 (0.04)	0.25 (0.09)	0.30 (0.10)	0.37 (0.03)	0.97 (0.04)	0.15 (0.12)	0.14 (0.11)	0.10 (0.03)	0.07 (0.05)	1.39 (0.85)	1.86 (1.50)	1.62 (0.62)	2.19 (1.78)
**H Bimodal (high NDVI)	0.61 (0.09)	0.21 (0.07)	0.19 (0.05)	0.38 (0.09)	0.83 (0.09)	0.30 (0.17)	0.23 (0.17)	0.15 (0.09)	0.17 (0.11)	2.32 (1.64)	2.88 (2.53)	2.16 (1.32)	2.46 (1.70)
* Unimodal (strongly seasonal)	0.23 (0.03)	0.32 (0.05)	–	0.45 (0.10)	–	0.29 (0.10)	–	0.28 (0.09)	–	2.64 (1.42)	–	2.51 (1.10)	–

Table 3
Various models for interannual variability were based on the historic ONI index, defined by the timing of its episodes (t_{ONI}) and their amplitudes (A_{ONI}) with optional random fluctuations at three different levels R_{1-3}. The sample statistics of episode duration ($\mu_{t.ONI}$,$\sigma_{t.ONI}$) are used to model irregular episodes (I).

Label	Description	Parameters
H_{oni}	Historic Oceanic Niño Index	$(t_{ONI},A_{ONI}) \rightarrow \mu_{t.ONI},\sigma_{t.ONI}$
H_{cst}	Historic episodes, constant amplitude	$f(t_{ONI},A_{int})$
H_{var}	Historic episodes, historic varying amplitude	$f(t_{ONI},A_{int}{}^{*}A_{ONI})$
R_{cst}	Regular episodes, constant amplitude	$f(\mu_{t.ONI},A_{int})$
R_{rnd}	Regular episodes, random varying amplitude	$f(\mu_{t.ONI},A_{int}{}^{*}R_{1-3})$
I_{cst}	Irregular episodes, constant amplitude	$f(\mu_{t.ONI}+\sigma_{t.ONI},A_{int})$
I_{rnd}	Irregular episodes, random varying amplitude	$f(\mu_{t.ONI}+\sigma_{t.ONI},A_{int}{}^{*}R_{1-3})$

correct direction (positive or negative). It therefore indicates if statements on the occurrence in time of positive and negative climatic episodes are reliable.

$$r = \sum_t (I_t - \overline{I})(O_t - \overline{O}) \Big/ \left[\sqrt{\sum_t (I_t - \overline{I})^2} \sqrt{\sum_t (O_t - \overline{O})^2} \right] \tag{7}$$

The relative root mean squared error (*rRMSE*) between two series summarizes the *rms* deviations between input and reconstructed components, relative to the interannual input component's *rms* amplitude. It indicates if statements on the magnitude of particular episodes are reliable.

$$rRMSE = \sqrt{\sum_t (I_t - O_t)^2 \Big/ \sum_t I_t{}^2} \tag{8}$$

Finally, weighted spectral coherence (*wCOH*) between two time series measures the correspondence (from 0 to 1) of the spectral content of two time series X_t based on their Fourier spectra $F(X_t)$, weighted by the Fourier spectrum of the reference series (with F^* the complex conjugate of F and $\langle . \rangle$ expectations estimated by ensemble averaging, see Alston, Vaughan, and Uttley (2013) for a detailed description). It therefore indicates the reliability of statements on the time scales of detected oscillations.

$$COH = |\langle F^*(I_t).F(O_t)\rangle|^2 \Big/ \left(\left\langle |F(I_t)|^2 \right\rangle \left\langle |F(O_t)|^2 \right\rangle \right) \tag{9}$$

$$wCOH = \sum (\mathrm{F}(I_t) * COH) / \sum F(I_t) \tag{10}$$

The overall strategy of composing a simulated series, its decomposition and subsequent evaluation of the detected interannual component against the input component is illustrated in Fig. 5.

2.6. Sensitivity analysis

Estimating sensitivity of the above described performance measures to the parameters in the simulated series aims at identifying those factors which most affect the successful detection of interannual components. Interannual to annual amplitude ($\sigma_{int}/\sigma_{ann}$), seasonal irregularity (φ_{seas}), noise autocorrelation (AC_1), noise level (σ_{noise}), seasonality type (SEAS), interannual model (INT) and reconstruction method (METH) each partially explain the variance in performance, which is assessed through multiple linear regression. Real-valued variables ($\sigma_{int}/\sigma_{ann}$,$\varphi_{seas}$,$AC_1$,$\sigma_{noise}$) are standardized with respect to their respective sample means and standard deviations, in order to make the estimated regression coefficients comparable. Categorical variables (SEAS, INT, METH) enter the regression equation as dummy variables. The regression coefficient for each category represents the effect of the variable taking on that particular category with respect to a reference category, enabling a ranking of categories per performance

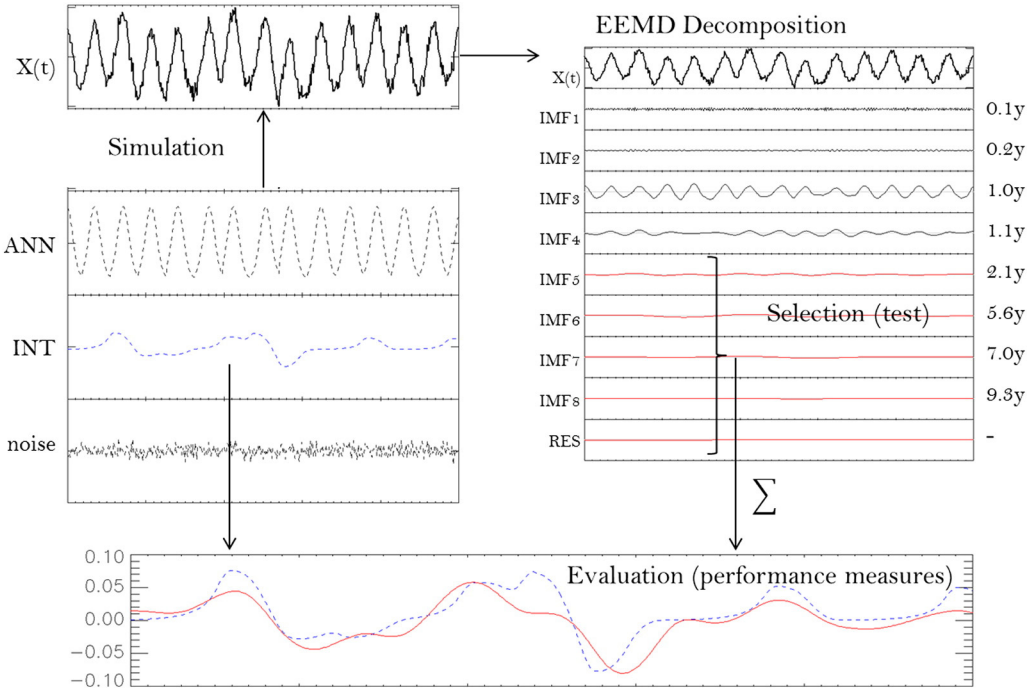

Fig. 5. Overview of the evaluation framework. Mixtures $X(t)$ of simulated annual (ANN), interannual (INT) and noise components are decomposed using EEMD. The interannual component (red solid line) is reconstructed from selected components and evaluated against the input component (blue dashed line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

measure. Within the approximation that an increase (decrease) of a real-valued variable by one standard deviation is equivalent to a categorical variable taking on a different category, all estimated coefficients are intercomparable and the factors that most affect the performance can be identified.

2.7. Case study: interannual NDVI and climate episodes

The validated performance of the EEMD method to retrieve interannual components from NDVI time series is conditioned by the degree to which the simulated series are representative for real world NDVI time series. To this end it was assessed in two steps whether the detected interannual NDVI patterns over East and Central Africa correspond to the expected long term fluctuations as predicted by historic climatic episodes.

In a first step, the zones where precipitation is most sensitive to interannual climatic and oceanic forcing are mapped by correlating the one-year accumulated Standardized Precipitation Index (SPI) with historic indices of two major climate phenomena acting in the region, i.e., ONI and DMI. The SPI index is a z-score of a precipitation

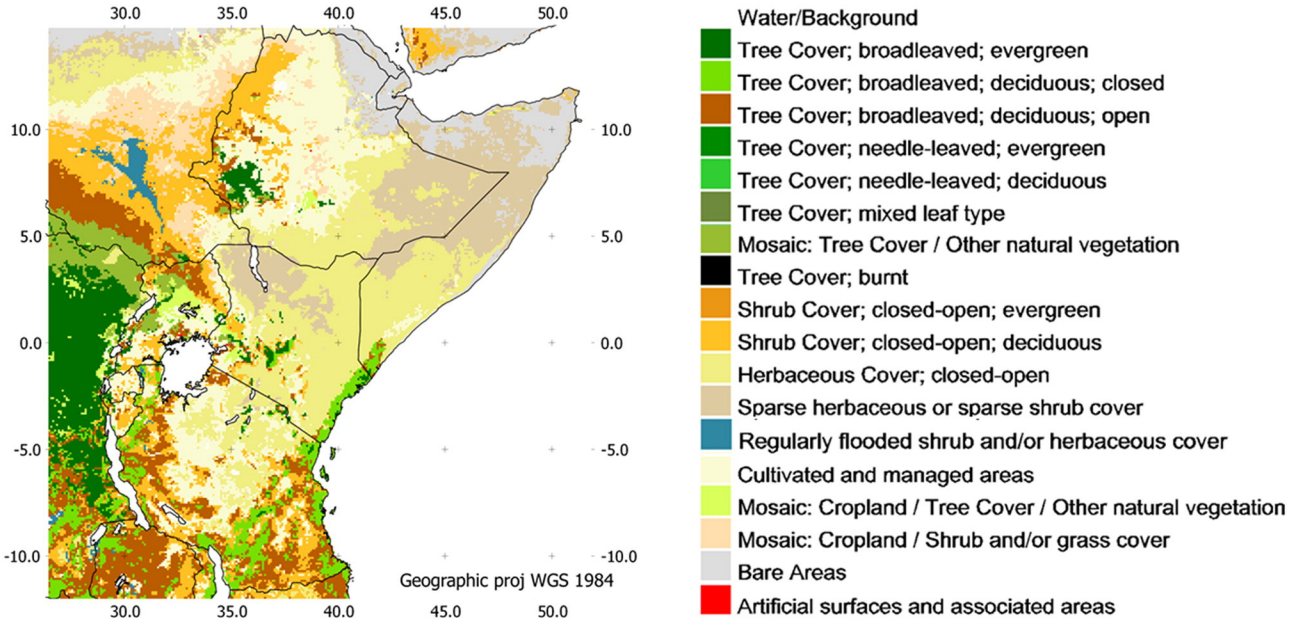

Fig. 6. The distribution of land cover types over the study area (GLC2000; Bartholomé & Belward, 2005) reflects to a large extent the different levels of mean NDVI shown in Fig. 3. Ecosystems respond differently to regional precipitation variability, as confirmed in the demonstration case study over East and Central Africa (see Discussion Section).

Table 4
Multiple linear regression of 3 performance measures against data and model parameters indicates the sensitivities of the EEMD method. Real-valued variables are scaled in order to make their coefficients comparable to the dummy coefficients of categorical variables. Regression coefficients indicate the effect of a change of one standard deviation of a real-valued variable, or the effect of a change of category with respect to a reference category for categorical variables.

	Correlation		rRMSE		wCOH	
Offset	0.68	***	1.22	***	0.41	***
Real-valued variables						
$\sigma_{int}/\sigma_{ann}$	0.21	***	−0.41	***	0.21	***
φ_{seas}	−0.05	***	0.30	***	−0.08	***
AC_1	−0.08	***	−0.03	***	−0.06	***
σ_{noise}	−0.01	***	0.04	***	−0.02	***
Categorical variables						
SEAS (reference category = 'high bimodal')						
'Low bimodal'	0.10	***	−0.42	***	0.12	***
'Unimodal'	0.05	***	−0.20	***	0.12	***
INT (reference category = Hcst)						
Rcst	0.03	***	−0.07	***	0.08	***
Rrnd	0.00		0.02		0.05	***
Honi	0.03	***	−0.03		0.02	***
Icst	−0.02	**	0.03	.	−0.01	
Hvar	−0.02	***	0.02		−0.01	
Irnd	−0.04	***	0.02	***	−0.02	**
METH (reference category = 'no test'						
'A posteriori'	−0.16	***	0.05	***	−0.09	***
'A priori'	−0.34	***	−0.00		−0.22	***

P(X > |t|) in Student t-test: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

event compared to the historic sample of events in the same period within the year, taking into account the non-normal distribution of rainfall events (Guttman, 1999). The SPI algorithm allows the definition of a time scale, i.e., the width of the past time window to consider as an accumulated precipitation event. SPI with a one-year time scale was used as a first proxy of interannual precipitation. The resulting maps reveal where the climate phenomena described by the indices have a strong coupling with the interannual precipitation patterns. This is where a strong interannual NDVI response can be expected.

In a second step, the decomposition model is applied to the NDVI time series of each pixel in the AVHRR/VGT dataset over East and Central Africa. The detected interannual components per pixel (with the choice of model parameters guided by the outcomes of the sensitivity analysis) are in turn correlated with the ONI and DMI indices. If EEMD detection on NDVI time series is effective, the footprint of ONI and DMI must be found in the interannual vegetation changes in those areas sensitive to precipitation fluctuations revealed in step one. A portion of the Global Land Cover map (GLC2000; Bartholomé & Belward, 2005) over the study area (Fig. 6) assists in the interpretation of the vegetation responses to climate variability.

3. Results

3.1. Sensitivity and performance of EEMD interannual NDVI detection using simulated time series

Following the evaluation framework (Fig. 5), the sensitivity of the EEMD method was assessed through multiple regressions of the performance measures against data and model parameters.

Table 4 shows multiple linear regression coefficients of performance measures against parameters by least-squares estimation, along with their symbolized significance in a Student-t test.

Correlation between input and reconstructed output is primarily determined by the interannual/annual amplitude ratio ($\sigma_{int}/\sigma_{ann}$) and to the choice of post-processing method, where both 'a priori' and 'a posteriori' testing tend to reduce the correlation between input and output interannual components relative to the no-IMF selection method ('no test'). To a lesser degree, more autocorrelation in the noise model tends to distort correlations, whereas the overall noise level has little impact over the range estimated from the real world dataset. While there is a limited discrepancy in input–output correlation between the different seasonal models and its degree of irregularity, the choice of interannual model does not affect correlation strongly.

The rRMSE between input and output interannual components takes on values in a range around unity. It decreases significantly only in the case of a strong interannual component, a relatively regular annual cycle, or a particular type of growing season. Noise characteristics, interannual models and selection criteria for interannual IMFs have less weight on the detection accuracy measured by rRMSE.

Weighted spectral coherence (wCOH) shows sensitivities very similar to that of correlation, favoring the inclusion of all detected interannual IMFs in the estimated interannual component. Again, the relative strength of the interannual component compared to the seasonal cycle is of more importance than their assumed underlying models or the characteristics of the noise in the series.

As a conclusion from the correlation, rRMSE and spectral coherence measures (Table 4) it can be seen that the 'no test' approach performs best to detect the input interannual components, but that the performance of this method is still strongly dependent on the interannual/annual amplitude ratio and seasonal irregularity expressed as a number of sample standard deviations in start of season. This can be seen for example in Table 5, which contains the values of the performance measures as a function of these two data-related parameters, following the 'no test' approach.

3.2. Case study over East and Central Africa

Fig. 7 displays the correlation of precipitation patterns in East and Central Africa with dominant regional climate phenomena. The ONI index, representing the El Niño Southern Oscillation (ENSO) based on Pacific sea surface temperatures, explains a part of the precipitation

Table 5
Performance measures of the 'no test' approach as a function of annual irregularity (ANN irr) and interannual/annual amplitude ratio (unimodal annual season, interannual model I_{rnd}, red noise ($AC1 = 0.25$; $\sigma = 0.03$ NDVI)). Darker shades indicate better detection performance.

		Correlation				rRMSE				Weighted spectral coherence			
ANN irr (σ)		0.0	0.5	1.0	1.5	0.0	0.5	1.0	1.5	0.0	0.5	1.0	1.5
INT/ANN ratio	0.10	0.59	0.53	0.24	0.15	0.91	1.39	3.26	4.27	0.29	0.21	0.08	0.05
	0.20	0.86	0.70	0.41	0.49	0.54	0.76	1.69	1.36	0.65	0.28	0.15	0.23
	0.33	0.90	0.81	0.75	0.53	0.45	0.61	0.82	1.10	0.71	0.58	0.59	0.25
	0.50	0.91	0.88	0.82	0.65	0.45	0.49	0.66	0.90	0.70	0.70	0.55	0.31
	1.00	0.92	0.83	0.91	0.92	0.64	0.58	0.44	0.41	0.75	0.57	0.77	0.74

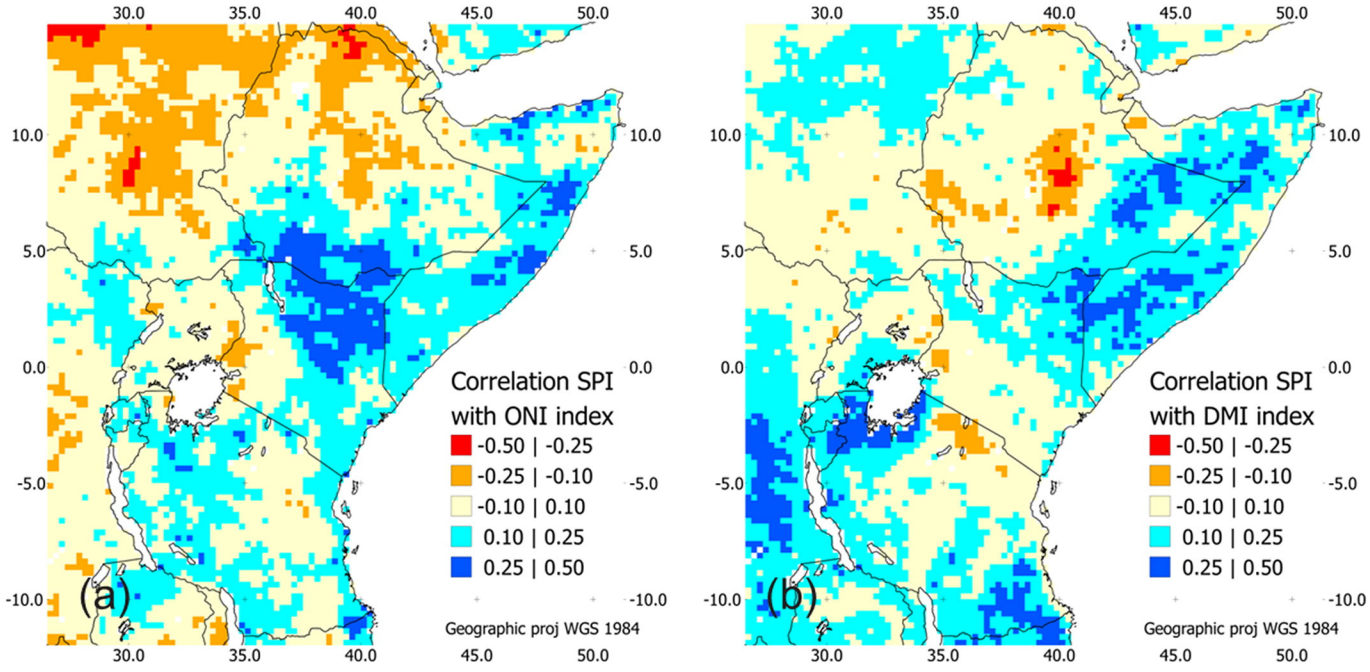

Fig. 7. Spatial representations of the correlation between one-year accumulated SPI and indices of regional climate phenomena highlight the areas where precipitation variability is sensitive to (a) the El Niño Southern Oscillation and (b) the Indian Ocean Dipole Mode. Oceanic exposure and local climate factors such as lakes partially predict sensitivity to these influences.

variability principally in the regions most exposed to oceanic influence (Fig. 7a). ENSO forcing is strongest in the Somalian, Ethiopian and Kenyan lowlands, only stretching further inland through the Turkana basin. Elsewhere its influence is limited by the Eastern and Southern rift mountains. Precipitation patterns in the Sudan lowlands towards the Sahel are rather out of phase with ENSO, pointing at the dominance of more local phenomena either not described by ENSO or negatively interfering with it (Williams & Hanan, 2011). Plisnier et al. (2000) indicated that (a) the effect of global ENSO episodes is often overruled by local climate agents such as large lakes, (b) the Indian Ocean dynamics might provide a better proxy for climate variability in certain African regions. As seen from Fig. 7b, The Indian Ocean DMI index, explains more variability of the Sahelian precipitation in the absence of direct ENSO influence. In addition, the precipitation downwind of Lake Victoria and Lake Tanganyika responds more to DMI forcing, stretching as far as the tropical rainforest in the Congo Basin (Fig. 6).

Next, it was assessed whether EEMD detected interannual NDVI is able to reveal the expected vegetation responses in the areas sensitive to climate variability. Fig. 8 shows the strength of the coupling between interannual NDVI and the ONI and DMI indices respectively. The

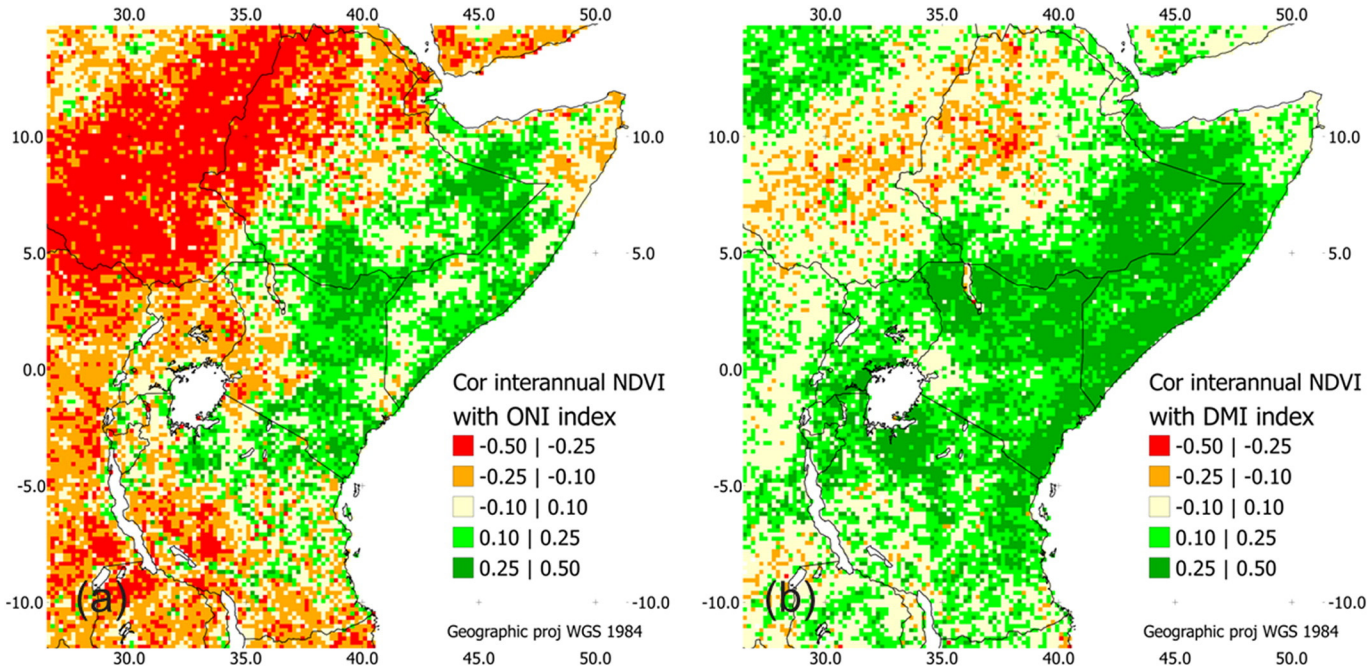

Fig. 8. Correlation maps of EEMD detected interannual NDVI with the historic ONI and DMI indices highlight the areas where vegetation changes are influenced by regional climate variability. These areas largely correspond to the areas where precipitation is forced by the same regional oceanic phenomena, demonstrating that meaningful interannual NDVI components can be extracted using EEMD decomposition.

interannual NDVI follows the ONI index in most of the lowlands identified as sensitive to ENSO (Fig. 8a), dominated by herbaceous and sparse shrub cover. However, the deciduous forest areas north of the Congo Basin do not respond to wet ENSO-driven episodes. As expected, vegetation in the continental lowlands responds negatively to lower than usual precipitation amounts during warm ENSO phases.

Correlation between interannual NDVI and DMI emphasize these findings. Although the Indian Ocean effects act differently throughout the study area, the detected responses of vegetation to DMI episodes matches the areas identified as sensitive to DMI well in most areas. The ocean–exposed lowlands under low vegetation cover display greenness anomalies in response to increased precipitation during positive DMI episodes. The same holds for other DMI-sensitive areas in the Eastern Sahel and around Lake Victoria. Remarkably, the coupling of vegetation to the DMI episodes cannot be explained everywhere by its effect through precipitation, for example in vast areas of South Kenya and North Tanzania. Again, the forest ecosystems of the Congo Basin are neutral towards precipitation variability driven by Indian Ocean phenomena. A similar observation is made for the partly forested areas in South Tanzania: despite strongly DMI-forced precipitation, the NDVI response is rather ambiguous. Finally, the out of phase character of precipitation in the Ethiopian highlands with respect to DMI is also detected in the interannual NDVI.

The lower response of tropical forest ecosystems reflects the lower importance of water as a limiting factor, as well as the saturation effect of NDVI for high leaf area canopies (Sellers, 1985). In addition, forest-climate interactions (Makarieva & Gorshkov, 2007) and moisture recycling dynamics (van der Ent, Savenije, Schaefli, & Steele-Dunne, 2010) are known to cause particular rainfall dynamics above forests, partly independent from the dominant oceanic moisture supply.

4. Discussion

Ensemble Empirical Mode Decomposition (EEMD) proves to be a suitable algorithm for the automated decomposition of large sets of NDVI time series into their characteristic time scales. The original EEMD decomposition technique (Huang et al., 1998; Wu & Huang, 2009), complemented by the IMF sorting procedure proposed in §2.2 and clear choices of parameters for *sifting* and *splining* (Appendix A), is able to deliver the interannual components of any given time series.

With respect to reconstructing the interannual component from detected components, we reject the existing significance test to distinguish signal from noise-induced components for interannual component detection, since neither the 'a priori' or 'a posteriori' version of the test yield better results than the 'no test' approach. Instead we weighted the capabilities of these significance tests to identify noise-induced low frequency components against the risk that these components may in fact contain signal as well as noise fragments (hereafter called *mixed* components). Although neither approach ensures a precise reconstruction of interannual components in terms of rRMSE, the linear and spectral properties of such components were best recovered by retaining all interannual IMFs ('no test'). This implies that mixed components, if occurring in the decompositions of our simulated time series, contain on average more signal than noise fragments. Attempts to filter noise-affected components by subjecting individual IMFs to significance tests decreased performance measured by correlation and spectral coherence.

However, two nuances must be added to this finding. First, it may hold true for the tested ranges of $\sigma_{int}/\sigma_{ann}$ and absolute σ_{noise}, which were in turn based on estimated ranges from a real NDVI dataset. Outside these ranges noise may start to dominate low frequency EEMD components, prompting a significance test to retain the true signal components. This may be commonly the case for other biogeophysical or climatological datasets and applications, so our conclusion should be limited to the application on integrated time series of remotely sensed vegetation indices and its typical signal to noise ratios. Second, it can be argued that the purpose of IMF significance testing in a climatological context, as in the work of Coughlin and Tung (2005), is to isolate *pure* time scales with a high degree of probability to originate from signal rather than noise. This justifies strict filtering of potentially noise-induced IMFs. Conversely, when using multiple IMFs to reconstruct interannual components with an unknown combination of time scales, retaining both *pure* and *mixed* IMFs may be favorable as a strategy.

The values of the performance measures for given (inter)annual models and noise characteristics under the 'no test' approach (Table 5) give an account of the major sources of variability in performance, i.e., the relative strength of the interannual component compared to the dominant annual cycle ($\sigma_{int}/\sigma_{ann}$), and the irregularity in timing of the annual season (φ_{seas}). It is worth noting that both factors do not act independently. The distorting effect of irregular growing seasons is of less impact when the superimposed interannual modulation is sufficiently high in amplitude, for example in arid ecoregions with a strong oceanic influence (Fig. 6). A practical implication is that both factors must preferably be estimated upon detection. Low $\sigma_{int}/\sigma_{ann}$ ratios, corresponding to weak interannual changes compared to the NDVI range of the growing season, may hamper the detection and subsequent processing of a pixel's interannual component, unless the annual cycle displays a sufficiently regular timing. If threshold values for both parameters are considered, pixels can be masked out from further interpretation. Whereas the choice of thresholds is arbitrary and dependent on the further steps in a particular application, the results presented in Table 5 offer some guidelines. For example, a threshold of $[\sigma_{int}/\sigma_{ann} > 0.33]$ will generally ensure a good reproduction of the interannual component measured by correlation and spectral coherence. As seen from Table 1, this applies for all ecoregions with a dry or wet bimodal growing season. For unimodal (semi-arid to subhumid) ecoregions, the interannual NDVI components are buried under a more dominant seasonal component. However, if the estimated variability (measured as standard deviation) of the growing season peak position does not exceed half of the sample standard deviation used in our simulated prototype (0.10 years, see Table 2), then the performance level is largely maintained. We therefore recommend that automated use of EEMD to reveal interannual components from large gridded time series data (biogeophysical, climatological) be preceded by a calibration step, whereby preliminary estimates of the series' composition and characteristics are collected in order to assess the proper limits of detection.

A major implication for the further use of detected interannual components of NDVI (or other remotely sensed vegetation indices) in a causal analysis with gridded climate datasets, is the fact that detection performs relatively well when measured by correlation or spectral coherence, but proves imprecise with respect to its rRMSE deviation from the true component. This suggests that positive and negative episodes in interannual vegetation change can be localized, and that the time scales present in the series are represented well. Conversely, the high rRMSE values (~0.40 under optimal conditions) indicate that an EEMD detected interannual component of a time series does not necessarily represent the actual interannual NDVI change at each discrete point in time. Neither can the absolute amplitudes of individual positive or negative episodes be interpreted as such. The power of the decomposition tool lies in the capability to filter the desired time scales from a mixed signal, and to produce a global representation of the occurrence and relative intensities of positive and negative episodes. Further development of a technique to identify and isolate a climate-forced component in the interannual NDVI component must take this knowledge into account. It must not rely entirely on instantaneous responses in the time domain, but rather consider sufficiently large time windows as well as the spectral domain for bivariate analysis with climate series.

A first effort to shift the EEMD analysis beyond the simulated setting towards real world NDVI series demonstrated that spatially coherent

patterns of vegetation dynamics can be revealed and explained (partially) by episodes of regional climate oscillations. Gridded precipitation datasets - remotely sensed as well as modeled using observational data – are expected to explain the detected interannual vegetation dynamics at a more local scale. Eventually, explained interannual variability can be separated from non-explained patterns, for further interpretative analysis over particular study areas.

The case study on real NDVI and climate datasets over East and Central Africa demonstrates that EEMD detection is able to produce meaningful indicators of interannual NDVI changes, as their spatial distribution corresponds largely to what can be expected from empirical climate data. As for any processing tool, the quality and the real world value of the output are inevitably limited by the quality of the input data. In the case of the EEMD decomposition of NDVI time series, inconsistencies in long term NDVI archives are the principal source of spurious patterns in the detected interannual NDVI, which may jeopardize conclusions in thematic studies like the case study in this paper. The most likely causes of inconsistency can be categorized as within-sensor inconsistencies (e.g., orbital drift or imperfect corrections for it) or between-sensor inconsistencies (e.g., historic processing differences, imperfect intercalibration of spectral response, or differences in sun-target-sensor geometry). Other sources of error with a less systematic occurrence, such as sensor noise and undetected clouds, will have less effect on the detected outcome, as the EEMD technique efficiently isolates high frequent components. With this knowledge, the quantitative validation of the detection tool was based on simulated data rather than on any existing real world dataset containing residual inconsistencies. The LTDR/VGT dataset contains at least two potential sources of inconsistency. The first one is the empirical linear sensor intercalibration between AVHRR and VGT as described in Section 2.1.1 yields slightly biased residuals for NDVI values above 0.65. In practice, this may cause an upward jump in NDVI at the 1999 breakpoint for pixels in densely forested areas. Moreover, 1999 may not be representative for the full NDVI record in case of residual between-sensor inconsistency in the LTDR version 2 data, which may cause an artifact interfering with the former effect. The second source of inconsistency stems from between-sensor inconsistency of the VGT-1 and VGT-2 sensors related to spectral response discrepancies. This effect has been identified to cause an upward break in 2003 throughout all NDVI strata (Tian et al., 2015). For future applications, alternative datasets can be considered. The LTDR version 4 provides improvement upon the earlier version with respect to geolocation as well as an extension in time with successor NOAA platforms (NASA, 2014). The updated 15-day composite NDVI product of the Global Inventory Modelling and Mapping Studies (GIMMS3g) results from efforts to recalibrate historic NOAA-AVHRR data by correcting for effects introducing spurious NDVI trends such as orbital drift and volcanic eruptions (Fensholt & Proud, 2012; Pinzon et al., 2005). GIMMS3g is thus likely to display improved stability over time, although potential artifacts have been identified over various NDVI strata (Tian et al., 2015). Finally, the MODIS instruments have provided NDVI records from 1999 onwards, considered an improvement upon the AVHRR products due to their narrow band definitions and wide dynamic range (Huete et al., 2002). Although generally deemed the alternative dataset with highest consistency, the MODIS NDVI series was found to contain the effects of sensor degradation (Wang et al., 2012).

Regardless the eventual choice of dataset, the EEMD decomposition and reconstruction approach in its current form is deemed operational to conduct thematic follow-up research. It is complementary to existing time series processing tools in the sense that 'interannual NDVI' representing slow year-to-year evolution of vegetation greenness is a new key variable in remote sensing-driven environmental research. Its user community may invoke extraction tools according to their specific needs: BFAST (Verbesselt et al., 2010) for the detection of stepwise trends and breakpoints relating to abrupt and gradual vegetation disturbances, such as wildfires and recovery; TIMESAT (Jönsson & Eklundh, 2002) for a detailed parameterization and quantification of parameters related to the growing season; SPIRITS (Eerens et al., 2014) for near-real time agricultural monitoring through processing of images and the extraction of vegetation indices related to crop status. EEMD is proposed to complement this list of tools as one to extract interannual vegetation changes in response to climatic oscillations and other sources of periodic disturbance episodes.

Despite this applicability, some points for methodological improvement can be defined. A review of existing methods to estimate a priori the nature and the level of noise from an NDVI time series may contribute to the calibration phase necessary before EEMD application on a new dataset. If noise can be modeled and estimated more accurately, the performance of the 'a posteriori' significance test for IMFs may surpass that of the current 'no test' approach. Further refinement of the EEMD application for automated analysis of gridded time series data can consist of evaluating the effects of series length and temporal resolution, which were held constant at 1114 ten-day time steps in this analysis.

5. Conclusions

Available archives of more than thirty years of remotely sensed optical imagery provide large spatiotemporal datasets, showing changes in condition of terrestrial vegetation as per-pixel NDVI trajectories. Beside seasonal variation, weaker but considerable interannual fluctuations exist in the NDVI series, and their identification is of high interest for research aiming at quantifying the long-term effects of climate variability on ecosystems. Thereto, signal processing techniques can be employed to detect interannual NDVI components within quantified limits of uncertainty.

We found that the Ensemble Empirical Mode Decomposition (EEMD) technique, extended with a sorting procedure for IMF components, is suitable to split a given NDVI time series into its constituent components with distinct time scales. The longer-than-annual EEMD components can be combined to reproduce the overall interannual oscillation present within the series. Implementation for automated processing of large NDVI datasets requires some complementary algorithmic formalization, which we proposed in this paper.

In order to gain insights in the performance of this method to accurately represent the true interannual component, and its sensitivities to data and model parameters, we tested its performance in a controlled setting with simulated NDVI time series. With respect to performance, we concluded that EEMD-detected interannual components generally represent well the occurrence of positive and negative episodes with proper time scales, whereas the local amplitudes must not be interpreted for individual episodes. This finding provides useful insights for subsequent bivariate analysis of interannual NDVI with precipitation series. With respect to sensitivities, the strength of the interannual component relative to the seasonal cycle, and modality and timing of the growing seasons, proved to be the principal factors affecting the detection performance. In addition, they interact and should therefore be assessed upon detection so that time series with low confidence in detection outcome can be excluded from the analysis.

The effectiveness of the EEMD decomposition method to detect interannual vegetation was demonstrated on real NDVI and climate data over East and Central Africa. In most areas where precipitation variability is sensitive to regional climatic and oceanic forcing, the method reveals interannual vegetation changes that show the expected responses, except in dense forest ecosystems. Finally, some methodological points of action were defined for further development of this tool to extract environmentally relevant information from growing remote sensing data sources.

Acknowledgments

This work was supported by the Flemish Institute for Technological Research (VITO) in Mol, Belgium through a PhD grant (1410215) to

Pieter Hawinkel, and by the Belgian Science Policy Office (Belspo) through the CVB contract with VITO (CB/67/08). Stef Lhermitte was supported as postdoctoral researcher by Fonds Wetenschappelijk Onderzoek-Vlaanderen (FWO). Bruno Verbist was supported by the VLIR-UOS/ARES-CCD Acropolis platform KLIMOS on sustainable development funded by the Belgian Development Cooperation (DGD). We gratefully thank the anonymous reviewers for their thorough reading and constructive comments which helped to improve this article.

Appendix A. The Empirical Mode Decomposition algorithm (Huang et al., 1998)

A series X(t) is decomposed into a finite number of Intrinsic Mode Functions (IMFs) and a residual term R(t) following the additive model (Eq. (3)).

The locations and values of all local minima and maxima in a series X(t) are detected using an extrema-finding function. Endpoints of the series are not considered as extrema unless they lie outside the range defined by the nearest minimum and maximum. For flat sections, the (left) middle location is returned.

The lower (upper) envelope of X(t) is obtained by connecting the local minima (maxima) of X(t) using a cubic spline. Cubic splines are uniquely defined when values for the second derivatives at the endpoints are provided. We follow the two-step approach of Pegram et al. (2008), here described for the left end of the lower (upper) envelope. The first original minimum (maximum) is mirrored around the left endpoint, and assigned the same second derivative. A cubic spline is calculated through the minima (maxima), ignoring the left endpoint. If the left endpoint is found to lie outside the calculated envelopes, it is added to the extrema set and the spline is recalculated. This is to avoid splines to vary freely at the ends, which would cause end effects propagating through the following iterative sifting procedure.

The mean of lower and upper envelopes M(t) contains the low-frequent information of X(t). The first ProtoMode Function (PMF_1) of X(t) is obtained by subtracting M(t) from X(t). This step is termed 'sifting' and PMF_1 is enriched in high frequency information. Sifting PMF_1 yields PMF_2 and the procedure is repeated k times until the number of extrema and the number of zero-crossings have not changed for S consecutive siftings. PMF_k is taken as the first Intrinsic Mode Function (IMF_1) of X(t). We follow the guidelines by Huang et al. (2003) and use [S = 5] as a stopping criterion for the sifting process.

The difference of X(t) and IMF_1 is the (preliminary) residual. It is ingested in the sifting procedure to yield IMF_2, which will contain the high frequency information after subtraction of IMF_1. More IMFs are obtained iteratively with increasing frequencies, until the (final) residual R(t) is monotonic or has one extremum. It can be regarded as the residual trend in the series.

The computing cost of the EMD algorithm increases quadratically with the length N_t of the time series. Its most costly step is the cubic spline interpolation, which involves manipulations on $N_t x N_t$ matrices. For $N_t = 1114$, the spline interpolation accounts for nearly 80% of the total computing cost of 1163 msec. In our IDL implementation, the computing time t_{comp} (in msec) for one series of length N_t is on average given by

$$t_{comp} = \left(0.000858 * {N_t}^2 + 0.282 * N_t + 23.2\right). \quad \text{(A1)}$$

For the EEMD extension, there is an additional linear increase with the number of noise-added iterations (set to 50 in this study). The vast number of pixels in a spatio-temporal dataset as the image sets over East and Central Africa (>14000 land pixels at 20 km resolution) further pushes the total processing time into the order of magnitude of weeks.

Appendix B. Significance test for IMFs based on period–energy relationships of white noise

The theoretical and derived empirical relationships to distinguish signal from noise, are based on two key characteristics of an IMF of a series with length N_t: its energy density

$$E = \frac{1}{N_t} \sum \mathrm{IMF}^2 \quad \text{(B1)}$$

and its theoretical mean period T, estimated as T by counting the number of extrema divided by twice the series length. Wu and Huang derived that the energy of a population of white noise series normalized to unit energy density has a mean Ẽ, related to its mean period T:

$$\ln(E) = -\ln(T). \quad \text{(B2)}$$

Moreover, the probability distribution of the energy density of a white noise IMF can be derived from the finding that $N_t{}^* E$ follows a χ^2-distribution with $N_t{}^* E$ degrees of freedom:

$$p(N_t * E) = (N_t * E)^{N_t * E/2 - 1} \exp\left(-\frac{N_t * E}{2}\right) \Big/ \left(2^{N_t * E/2} \Gamma(N_t * E/2)\right). \quad \text{(B3)}$$

Through log-transformation of the energy density

$$y = \ln(E) \quad \text{(B4)}$$

and substitution of Eqs. (B2) and (B4) into the probability distribution of Eq. (B3), the confidence levels for the log-energy density of an IMF originating from white noise can be estimated from its mean period $\hat{T}$.

$$p(y) = (N_t * \exp(y))^{N_t/(2\hat{T})} \exp(-N_t * \exp(y)/2) / \left(2^{N_t/(2\hat{T})} \Gamma\left(N_t/\left(2\hat{T}\right)\right)\right) \quad \text{(B5)}$$

The bias in the estimation of $\hat{T}$ affects Eq. (B2) but preserves its linear relationship. Monte Carlo verification by a large number of random white noise simulations yielded the empirical relationship for series length [N_t = 1114]:

$$\ln(E) = -1.08 \ln(T) + 0.12. \quad \text{(B6)}$$

The energy distributions and confidence levels were recalculated accordingly, and verified with the simulation result. The (shifted) theoretical 99% confidence limit proved to bound 95% of the white noise simulations. Similar outcomes with slightly different coefficients were found for red noise simulations with [AC1 = {0.25, 0.50}].

As white noise is assumed to contain no information, real signal components will display energies higher than can be expected from white noise with a chosen level of confidence. Under the 'a priori' approach, a given time series consisting of signal and noise is normalized to unit energy density and treated as being potentially pure noise. If however the noise level in the series can be estimated, the normalization step can be performed more accurately 'a posteriori' and more signal components can be distinguished. Both approaches were evaluated in this paper.

Appendix C. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2015.08.024.

Appendix C. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:http://dx.doi.org/10.1016/j.rse.2015.08.024.

These data include Google maps of the most important areas described in this article.

References

Achard, F., Eva, H.D., Stibig, H. -j., Mayaux, P., Gallego, J., Richards, T., & Malingreau, J. -p. (2002). Determination of deforestation rates of the world's humid tropical forests. *Science*, *297*, 999–1002.

Alston, W.N., Vaughan, S., & Uttley, P. (2013). The flux-dependent X-ray time lags in NGC 4051. *Monthly Notices of the Royal Astronomical Society*, *435*, 1511–1519.

Anyamba, A., & Eastman, J.R. (1996). Interannual variability of NDVI over Africa and its relation to El Nino Southern Oscillation. *International Journal of Remote Sensing*, *17*, 2533–2548.

Atzberger, C. (2013). Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs. *Remote Sensing*, *5*, 949–981.

Bartholomé, E., & Belward, A.S. (2005). GLC2000: A new approach to global land cover mapping from earth observation data. *International Journal of Remote Sensing*, *26*, 1959–1977.

Baynard, C.W. (2013). Remote sensing applications: Beyond land-use and land-cover change. *Advances in Remote Sensing*, *2*, 228–241.

Brando, P.M., Goetz, S.J., Baccini, A., Nepstad, D.C., Beck, P.S.A., & Christman, M.C. (2010). Seasonal and interannual variability of climate and vegetation indices across the Amazon. *Proceedings of the National Academy of Sciences*, *107*, 14685–14690.

Brisson, E., Demuzere, M., Willems, P., & van Lipzig, N.M. (2015). Assessment of natural climate variability using a weather generator. *Climate Dynamics*, *44*, 495–508.

Brown, M.E., de Beurs, K., & Vrieling, A. (2010). The response of African land surface phenology to large scale climate oscillations. *Remote Sensing of Environment*, *114*, 2286–2296.

Camberlin, P., Martiny, N., Philippon, N., & Richard, Y. (2007). Determinants of the interannual relationships between remote sensed photosynthetic activity and rainfall in tropical Africa. *Remote Sensing of Environment*, *106*, 199–216.

Chen, C. -F., Son, N. -T., Chang, L. -Y., & Chen, C. -C. (2011). Monitoring of soil moisture variability in relation to rice cropping systems in the Vietnamese Mekong Delta using MODIS data. *Applied Geography*, *31*, 463–475.

Coughlin, K., & Tung, K.K. (2005). Empirical mode decomposition of climate variability in the atmosphere. In N. Huang, & S. Shen (Eds.), *Hilbert-Huang Transform and Its Applications. World Scientific.* (pp. 149–166).

Cracknell, A.P. (2001). The exciting and totally unanticipated success of the AVHRR in applications for which it was never intended. *Advances in Space Research*, *28*, 233–240.

De Keersmaecker, W., Lhermitte, S., Tits, L., Honnay, O., Somers, B., & Coppin, P. (2015). A model quantifying global vegetation resistance and resilience to short-term climate anomalies and their relationship with vegetation cover. *Global Ecology and Biogeography*, *24*, 539–548.

Deronde, B., Debruyn, W., Gontier, E., Goor, E., Jacobs, T., Verbeiren, S., & Vereecken, J. (2014). 15 years of processing and dissemination of SPOT-VEGETATION products. *International Journal of Remote Sensing*, *35*, 2402–2420.

Dierckx, W., Sterckx, S., Benhadj, I., Livens, S., Duhoux, G., Van Achteren, T., ... Saint, G. (2014). PROBA-V mission for global vegetation monitoring: standard products and image quality. *International Journal of Remote Sensing*, *35*, 2589–2614.

Eerens, H., Haesen, D., Rembold, F., Urbano, F., Tote, C., & Bydekerke, L. (2014). Image time series processing for agriculture monitoring. *Environmental Modelling & Software*, *53*, 154–162.

van der Ent, R.J., Savenije, H.H.G., Schaefli, B., & Steele-Dunne, S.C. (2010). Origin and fate of atmospheric moisture over continents. *Water Resources Research*, *46* (n/a-n/a).

Feng, Z., Liang, M., Zhang, Y., & Hou, S. (2012). Fault diagnosis for wind turbine planetary gearboxes via demodulation analysis based on ensemble empirical mode decomposition and energy separation. *Renewable Energy*, *47*, 112–126.

Fensholt, R., & Proud, S.R. (2012). Evaluation of earth observation based global long term vegetation trends — Comparing GIMMS and MODIS global NDVI time series. *Remote Sensing of Environment*, *119*, 131–147.

Fensholt, R., Langanke, T., Rasmussen, K., Reenberg, A., Prince, S.D., Tucker, C., ... Wessels, K. (2012). Greenness in semi-arid areas across the globe 1981–2007 — An earth observing satellite based analysis of trends and drivers. *Remote Sensing of Environment*, *121*, 144–158.

Flandrin, P., Rilling, G., & Goncalves, P. (2004). Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters*, *11*, 112–114.

Galford, G.L., Mustard, J.F., Melillo, J., Gendrin, A., Cerri, C.C., & Cerri, C.E.P. (2008). Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment*, *112*, 576–587.

Guan, B.T. (2014). Ensemble empirical mode decomposition for analyzing phenological responses to warming. *Agricultural and Forest Meteorology*, *194*, 1–7.

Guo, L., Wu, S., Zhao, D., Yin, Y., Leng, G., & Zhang, Q. (2014). NDVI-based vegetation change in Inner Mongolia from 1982 to 2006 and its relationship to climate at the biome scale. *Advances in Meteorology*, *2014*, 12.

Guttman, N.B. (1999). Accepting the standardized precipitation index: A calculation algorithm. *Journal of the American Water Resources Association*, *35*, 311–322.

Hilker, T., Lyapustin, A.I., Tucker, C.J., Hall, F.G., Myneni, R.B., Wang, Y.J., ... Sellers, P.J. (2014). Vegetation dynamics and rainfall sensitivity of the Amazon. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 16041–16046.

Hirota, M., Holmgren, M., Van Nes, E.H., & Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. *Science*, *334*, 232–235.

Holmgren, M., Hirota, M., van Nes, E.H., & Scheffer, M. (2013). Effects of interannual climate variability on tropical tree cover. *Nature Climate Change*, *3*, 755–758.

Holmgren, M., Scheffer, M., Ezcurra, E., Gutiérrez, J., & Mohren, G.F.J. (2001). El Niño effects on the dynamics of terrestrial ecosystems. *Trends in Ecology & Evolution*, *16*, 89–94.

Hsu, K. -l., Gupta, H.V., Gao, X., & Sorooshian, S. (1999). Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation. *Water Resources Research*, *35*, 1605–1618.

Huang, B., Banzon, V.F., Freeman, E., Lawrimore, J., Liu, W., Peterson, T.C., ... Zhang, H. -M. (2014). Extended reconstructed sea surface temperature version 4 (ERSST.v4). part I: Upgrades and intercomparisons. *Journal of Climate*, *28*, 911–930.

Huang, N.E., Shen, Z., & Long, S.R. (1999). A new view on nonlinear water waves. *Annual Review of Fluid Mechanics*, *31*, 417–457.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., ... Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London*, *454*, 903–995.

Huang, N.E., Wu, M.C., Long, S.R., Shen, S.S., Wendong, Q., Gloersen, P., & Fan, K.L. (2003). A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proceedings of the Royal Society of London*, *459*, 2317–2345.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., & Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213.

Immerzeel, W.W., Quiroz, R.A., & De Jong, S.M. (2005). Understanding precipitation patterns and land use interaction in Tibet using harmonic analysis of SPOT VGT-S10 NDVI time series. *International Journal of Remote Sensing*, *26*, 2281–2296.

IPCC (2007). Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, & H.L. Miller (Eds.), Intergovernmental Panel on Climate Change (p. 996).

JAMSTEC, & Indian Ocean Dipole Mode Index (DMI) (2010). *Japan Agency for Marine-Earth Science and Technology.* URL: http://www.jamstec.go.jp/frcgc/research/d1/iod/DATA/dmi_HadISST.txt, last accessed 15 December 2014.

Ji, L., & Gallo, K. (2006). An agreement coefficient for image comparison. *Photogrammetric Engineering & Remote Sensing*, *72*, 823–833.

Jönsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, *40*, 1824–1832.

Kuo, C. -Y., Wei, S. -K., & Tsai, P. -W. (2013). Ensemble emprirical mode decomposition with supervised cluster analysis. *Advances in Adaptive Data Analysis*, *5*.

Lhermitte, S., Verbesselt, J., Jonckheere, I., Nackaerts, K., van Aardt, J.A.N., Verstraeten, W.W., & Coppin, P. (2008). Hierarchical image segmentation based on similarity of NDVI time series. *Remote Sensing of Environment*, *112*, 506–521.

Lhermitte, S., Verbesselt, J., Verstraeten, W.W., & Coppin, P. (2011). A comparison of time series similarity measures for classification and change detection of ecosystem dynamics. *Remote Sensing of Environment*, *115*, 3129–3152.

Liu, M., Tian, H., Chen, G., Ren, W., Zhang, C., & Liu, J. (2008). Effects of land-use and land-cover change on evapotranspiration and water yield in China during 1900–2001. *JAWRA Journal of the American Water Resources Association*, *44*, 1193–1207.

Luo, Y., Melillo, J., Niu, S., Beier, C., Clark, J.S., Classen, A.T., ... Torn, M.S. (2011). Coordinated approaches to quantify long-term ecosystem dynamics in response to global change. *Global Change Biology*, *17*, 843–854.

Maisongrande, P., Duchemin, B., & Dedieu, G. (2004). VEGETATION/SPOT: An operational mission for the earth monitoring; presentation of new standard products. *International Journal of Remote Sensing*, *25*, 9–14.

Makarieva, A.M., & Gorshkov, V.G. (2007). Biotic pump of atmospheric moisture as driver of the hydrological cycle on land. *Hydrology and Earth System Sciences*, *11*, 1013–1033.

Mann, M.E., & Lees, J.M. (1996). Robust estimation of background noise and signal detection in climatic time series. *Climatic Change*, *33*, 409–445.

Martínez, B., & Gilabert, M.A. (2009). Vegetation dynamics from NDVI time series analysis using the wavelet transform. *Remote Sensing of Environment*, *113*, 1823–1842.

Molla, M.K.I., Ghosh, P.R., & Hirose, K. (2011). Bivariate EMD-based data adaptive approach to the analysis of climate variability. *Discrete Dynamics in Nature and Society*, *2011*, 21.

Myneni, R.B., Los, S.O., & Tucker, C.J. (1996). Satellite-based identification of linked vegetation index and sea surface temperature anomaly areas from 1982 to 1990 for Africa, Australia and South America. *Geophysical Research Letters*, *23*, 729–732.

Nagol, J.R., Vermote, E.F., & Prince, S.D. (2014). Quantification of impact of orbital drift on inter-annual trends in AVHRR NDVI data. *Remote Sensing*, *6*, 6680–6687.

NASA (2014). *Land long term data record.* URL http://ltdr.nascom.nasa.gov/cgi-bin/ltdr/ltdrPage.cgi?fileName=LTDR_update, last accessed 24 June 2015.

Nicholson, S.E., & Farrar, T.J. (1994). The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. I. NDVI response to rainfall. *Remote Sensing of Environment*, *50*, 107–120.

Nicholson, S.E., & Kim, J. (1997). The relationship of the El Niño-southern oscillation to African rainfall. *International Journal of Climatology*, *17*, 117–135.

NOAA-CPC (2014). Historic El Niño/La Niña episodes (1950-present). *Climate Prediction Center, National Oceanic and Atmospheric Administration* URL http://www.cpc.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml, last accessed 24 June.

Pedelty, J., Devadiga, S., Masuoka, E., Brown, M., Pinzon, J., Tucker, C., ... Pinheiro, A. (2007). Generating a long-term land data record from the AVHRR and MODIS instruments. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 23–28 July 2007* (pp. 1021–1025). Barcelona, Spain: IEEE.

Pegram, G.G.S., Peel, M.C., & McMahon, T.A. (2008). Empirical mode decomposition using rational splines: An application to rainfall time series. *Proceedings of the Royal Society A*, *464*, 1483–1501.

Piao, S., Nan, H., Huntingford, C., Ciais, P., Friedlingstein, P., Sitch, S., ... Chen, A. (2014). Evidence for a weakening relationship between interannual temperature variability and northern vegetation activity. *Nature Communications*, *5*.

Piao, S.L., Wang, X.H., Ciais, P., Zhu, B., Wang, T., & Liu, J. (2011). Changes in satellite-derived vegetation growth trend in temperate and boreal Eurasia from 1982 to 2006. *Global Change Biology*, *17*, 3228–3239.

Pinzon, J.E., Brown, M.E., & Tucker, C.J. (2005). EMD correction of orbital drift artifacts in satellite data stream. In N.E. Huang, & S.S.P. Shen (Eds.), *Hilbert-Huang transform and its applications* (pp. 167–186). World Scientific Publishing.

Plisnier, P.D., Serneels, S., & Lambin, E.F. (2000). Impact of ENSO on East African ecosystems: a multivariate analysis based on climate and remote sensing data. *Global Ecology and Biogeography*, *9*, 481–497.

Quiroz, R., Yarlequé, C., Posadas, A., Mares, V., & Immerzeel, W.W. (2011). Improving daily rainfall estimation from NDVI using a wavelet transform. *Environmental Modelling & Software*, *26*, 201–209.

Richard, Y., & Poccard, I. (1998). A statistical study of NDVI sensitivity to seasonal and interannual rainfall variations in Southern Africa. *International Journal of Remote Sensing*, *19*, 2907–2920.

Roerink, G.J., Menenti, M., Soepboer, W., & Su, Z. (2003). Assessment of climate impact on vegetation dynamics by using remote sensing. *Physics and Chemistry of the Earth*, *28*, 103–109.

Sellers, P.J. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, *6*, 1335–1372.

Sitters, J., Holmgren, M., Stoorvogel, J.J., & López, B.C. (2012). Rainfall-tuned management facilitates dry forest recovery. *Restoration Ecology*, *20*, 33–42.

Sorooshian, S., Hsu, K. -L., Gao, X., Gupta, H.V., Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, *81*, 2035–2046.

Steven, M.D., Malthus, T.J., Baret, F., Xu, H., & Chopping, M.J. (2003). Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment*, *88*, 412–422.

Swets, D.L., Reed, B.C., Rowland, J.D., & Marko, S.E. (1999). A weighted least-squares approach to temporal NDVI smoothing. *ASPRS Annual Conference, 17–21 May 1999* (pp. 526–536). Oregon, Portland: American Society for Photogrammetry and Remote Sensing.

Swinnen, E. (2008). Vegetation dynamics in Southern Africa from NOAA-AVHRR and SPOT-VGT time series. *Faculté des sciences. École doctorale geosciences* (pp. 208). Louvain-la-Neuve: Université catholique de Louvain.

Swinnen, E., & Veroustraete, F. (2008). Extending the SPOT-VEGETATION NDVI time series (1998–2006) back in time with NOAA-AVHRR data (1985–1998) for Southern Africa. *IEEE Transactions on Geoscience and Remote Sensing*, *46*, 558–572.

Tarnavsky, E., Garrigues, S., & Brown, M.E. (2008). Multiscale geostatistical analysis of AVHRR, SPOT-VGT, and MODIS global NDVI products. *Remote Sensing of Environment*, *112*, 535–549.

Tian, F., Fensholt, R., Verbesselt, J., Grogan, K., Horion, S., & Wang, Y. (2015). Evaluating temporal consistency of long-term global NDVI datasets for trend analysis. *Remote Sensing of Environment*, *163*, 326–340.

Torrence, C., & Compo, G.P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, *79*, 61–78.

Torrence, C., & Webster, P.J. (1999). Interdecadal changes in the ENSO–Monsoon system. *Journal of Climate*, *12*, 2679–2690.

Trishchenko, A.P., Cihlar, J., & Li, Z. (2002). Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. *Remote Sensing of Environment*, *81*, 1–18.

Tucker, C.J., & Sellers, P.J. (1986). Satellite remote sensing of primary production. *International Journal of Remote Sensing*, *7*, 1395–1416.

Verbesselt, J., Hyndman, R., Zeileis, A., & Culvenor, D. (2010). Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. *Remote Sensing of Environment*, *114*, 2970–2980.

Vermote, E., & Kaufman, Y.J. (1995). Absolute calibration of AVHRR visible and near-infrared channels using ocean and cloud views. *International Journal of Remote Sensing*, *16*, 2317–2340.

Vermote, E., Justice, C.O., & Breon, F.M. (2009). Towards a generalized approach for correction of the BRDF effect in MODIS directional reflectances. *IEEE Transactions on Geoscience and Remote Sensing*, *47*, 898–908.

Wang, D., Morton, D., Masek, J., Wu, A., Nagol, J., Xiong, X., ... Wolfe, R. (2012). Impact of sensor degradation on the MODIS NDVI time series. *Remote Sensing of Environment*, *119*, 55–61.

Williams, C.A., & Hanan, N.P. (2011). ENSO and IOD teleconnections for African ecosystems: Evidence of destructive interference between climate oscillations. *Biogeosciences*, *8*, 27–40.

Wu, Z., & Huang, N.E. (2004). A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society of London*, *460*, 1597–1611.

Wu, Z., & Huang, N.E. (2005). Statistical significance test of intrinsic mode functions. In N.E. Huang, & S.S.P. Shen (Eds.), *Hilbert-Huang transform and its applications* (pp. 107–127). World Scientific.

Wu, Z., & Huang, N.E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, *1*, 1–41.

Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: A review. *Journal of Plant Ecology*, *1*, 9–23.

Zhou, Q., Jiang, H., Wang, J., & Zhou, J. (2014b). A hybrid model for PM2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Science of the Total Environment*, *496*, 264–274.

Zhou, L., Tian, Y., Myneni, R.B., Ciais, P., Saatchi, S., Liu, Y.Y., ... Hwang, T. (2014a). Widespread decline of Congo rainforest greenness in the past decade. *Nature*, *509*, 86–90.

Zhou, H., Van Rompaey, A., & Wang, J. a. (2009). Detecting the impact of the "grain for green" program on the mean annual vegetation cover in the Shaanxi province, China using SPOT-VGT NDVI data. *Land Use Policy*, *26*, 954–960.