Analyzing Vegetation Trends with Sensor Data from Earth Observation Satellites

Jamali, Sadegh

2014

Link to publication

Citation for published version (APA):
Analyzing Vegetation Trends with Sensor Data from Earth Observation Satellites

SADEGH JAMALI
FACULTY OF SCIENCE | LUND UNIVERSITY
Analyzing Vegetation Trends with Sensor Data from Earth Observation Satellites

by
Sadegh Jamali

DOCTORAL DISSERTATION
by due permission of the Faculty of Science, Lund University, Sweden.
To be defended at Världen auditorium, Geocentrum I,
Friday, October 24, 2014, at 10.00 a.m.

Faculty opponent
Professor Ronald Eastman
Clark University, USA.
Analyzing Vegetation Trends with Sensor Data from Earth Observation Satellites

Abstract

This thesis aims to advance the analysis of nonlinear trends in time series of vegetation data from Earth observation satellite sensors. This is accomplished by developing fast, efficient methods suitable for large volumes of data. A set of methods, tools, and a framework are developed and verified using data from regions containing vegetation change hotspots.

First, a polynomial-fitting scheme is tested and applied to annual Global Inventory Modeling and Mapping Studies (GIMMS)–Normalized Difference Vegetation Index (NDVI) observations for North Africa for the period 1982–2006. Changes in annual observations are divided between linear and nonlinear (cubic, quadratic, and concealed) trend behaviors. A concealed trend is a nonlinear change which does not result in a net change in the amount of vegetation over the period.

Second, a systematic comparison between parametric and non-parametric techniques for analyzing trends in annual GIMMS-NDVI data is performed at fifteen sites (located in Africa, Spain, Italy, and Iraq) to compare how trend type and departure from normality assumptions affect each method’s accuracy in detecting long-term change.

Third, a user-friendly program (Detecting Breakpoints and Estimating Segments in Trend, DBEST) has been developed which generalizes vegetation trends to main features, and characterizes vegetation trend changes. The outputs of DBEST are the simplified trend, the change type (abrupt or non-abrupt), and estimates for the characteristics (time and magnitude) of the change. DBEST is tested and evaluated using both simulated NDVI data, and actual NDVI time series for Iraq for the period 1982-2006.

Finally, a decision-making framework is presented to help analysts perform comprehensive analyses of trend/change in time series of satellite sensor data. The framework is based on a conceptual model of the main aspects of trend analyses, including identification of the research question, the required data, the appropriate variables, and selection of efficient analysis methods. To verify the framework, it is applied to four case studies (located in Burkina Faso, Spain, Sweden, and Senegal) using Moderate-resolution Imaging Spectroradiometer (MODIS)–NDVI data for the period 2000–2013. Each of the case studies successfully achieved its research aim(s), showing that the framework can achieve the main goal of the study which is to advance the analysis of nonlinear changes in vegetation. The methods developed in this thesis can help to contribute more accurate information about vegetation dynamics to the field of land cover change research.

Key words: Change detection, Satellite sensor data, Time series analysis, Vegetation dynamics, Vegetation index

Recipient’s notes

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature  Sadegh Jamali  Date 16/09/14
A doctoral thesis at a university in Sweden is produced either as a monograph or as a collection of papers. In the latter case, the introductory part constitutes the formal thesis, which summarizes the accompanying papers already published or manuscripts at various stages (in press, submitted or in preparation).
List of Papers


IV. Jamali, S., Eklundh, L., Seaquist, J., & Ardö, J. A framework for temporal trend analysis of satellite sensor data. *(Manuscript)*

Contribution

I. SJ was responsible for study design and data analysis, and led the writing of the manuscript.

II. SJ was responsible for study design and data analysis, and led the writing of the manuscript.

III. SJ was responsible for study design and data analysis, developed the DBEST algorithm, and led the writing of the manuscript.

IV. SJ was responsible for study design and data analysis, and led the writing of the manuscript.

Paper I is reproduced with permission from Elsevier.
Abstract

This thesis aims to advance the analysis of nonlinear trends in time series of vegetation data from Earth observation satellite sensors. This is accomplished by developing fast, efficient methods suitable for large volumes of data. A set of methods, tools, and a framework are developed and verified using data from regions containing vegetation change hotspots.

First, a polynomial-fitting scheme is tested and applied to annual Global Inventory Modeling and Mapping Studies (GIMMS)–Normalized Difference Vegetation Index (NDVI) observations for North Africa for the period 1982–2006. Changes in annual observations are divided between linear and nonlinear (cubic, quadratic, and concealed) trend behaviors. A concealed trend is a nonlinear change which does not result in a net change in the amount of vegetation over the period.

Second, a systematic comparison between parametric and non-parametric techniques for analyzing trends in annual GIMMS-NDVI data is performed at fifteen sites (located in Africa, Spain, Italy, and Iraq) to compare how trend type and departure from normality assumptions affect each method’s accuracy in detecting long-term change.

Third, a user-friendly program (Detecting Breakpoints and Estimating Segments in Trend, DBEST) has been developed which generalizes vegetation trends to main features, and characterizes vegetation trend changes. The outputs of DBEST are the simplified trend, the change type (abrupt or non-abrupt), and estimates for the characteristics (time and magnitude) of the change. DBEST is tested and evaluated using both simulated NDVI data, and actual NDVI time series for Iraq for the period 1982-2006.

Finally, a decision-making framework is presented to help analysts perform comprehensive analyses of trend/change in time series of satellite sensor data. The framework is based on a conceptual model of the main aspects of trend analyses, including identification of the research question, the required data, the appropriate variables, and selection of efficient analysis methods. To verify the framework, it is applied to four case studies (located in Burkina Faso, Spain, Sweden, and Senegal) using Moderate-resolution Imaging Spectroradiometer (MODIS)–NDVI data for the period 2000–2013. Each of the case studies successfully achieved its research aim(s), showing that the framework can achieve the main goal of the study which is to advance the analysis of nonlinear changes in vegetation. The methods developed in this thesis can help to contribute more accurate information about vegetation dynamics to the field of land cover change research.
Mänskliga aktiviteter och klimatförändringar påverkar vegetationen över hela jorden, idag och i framtiden. Denna påverkan kan studeras och kvantifieras med hjälp av observationer från satelliter vilka förmedlar information om vegetationsförändringar i tid och rum. I denna doktorsavhandling har metoder för analys av vegetationsförändringar utvecklats och utvärderats. Lämpligheten hos olika typer av tidsserieanalyser och statistiska metoder har testats och utvärderats i områden med kända vegetationsförändringar. Vidare har användarvänliga gränssnitt utvecklats för att underlätta för andra användare av satellitdata att utnyttja de utvecklade metoderna. Metoderna bidrar till förenklade studier, kvantifiering och bättre förståelse av vegetationsförändringar och deras relation till miljöförändringar och mänsklig påverkan.
Contents

1. Introduction 1
   1.1. Global climate change and vegetation change 1
   1.2. Remote sensing 2
      1.2.1. Earth observation satellites 2
      1.2.2. Satellite-based vegetation sensing 3
   1.3. Vegetation time series analysis 7
      1.3.1. Analysis of linear change 10
      1.3.2. Analysis of nonlinear change 10
   1.4. Thesis objectives 11

2. Materials and methods 13
   2.1. Satellite data 13
      2.1.1. NOAA AVHRR products 13
      2.1.2. MODIS products 14
   2.2. Vegetation change hotspots 15
      2.2.1. North Africa 15
      2.2.2. Change hotspots 17
      2.2.3. Simulated data 19
   2.3. Analyzing vegetation trends 20

3. Results and discussion 21
   3.1. Paper I: Automated mapping of vegetation trends with polynomials using NDVI imagery over the Sahel 21
   3.2. Paper II: Comparing ordinary least-squares regression and a non-parametric method for estimating trends in multi-year NDVI data 22
   3.3. Paper III: Detecting changes in vegetation trends using time series segmentation 24
   3.4. Paper IV: A framework for temporal trend analysis of satellite sensor data 25
   3.5. General discussion 27
   3.6. Contribution to state-of-the-art 29

4. Conclusions 31

5. Future studies 33

Acknowledgements 34

References 35
1. Introduction

1.1. Global climate change and vegetation change

“It is certain that Global Mean Surface Temperature has increased since the late 19th century” (Hartmann et al., 2013). Over the period 1880–2012, the globally averaged combined land and ocean surface temperature data show a warming of 0.85 °C, as stated in the fifth assessment report by Intergovernmental Panel on Climate Change (IPCC) (Hartmann et al., 2013). The increasing temperature is mainly caused by increasing global atmospheric concentrations of greenhouse gases, especially carbon dioxide (CO\textsubscript{2}) which primarily results from fossil fuel burning and land use change (Solomon et al., 2007). According to long-term projections of climate change, the global mean temperature will continue to rise throughout the current century if greenhouse gas emissions continue unabated (Collins et al., 2013).

Climate change can significantly affect both vegetation dynamics and distribution, because climate exerts a dominant control over global vegetation activity and the distribution of species (e.g. Pearson and Dawson, 2003). Changes in vegetation dynamics and distribution (i.e. vegetation change) have a number of different causes, including: natural disturbances (e.g. fire, insect attack, and flood), climate factors (e.g. drought, variations in rainfall and temperature), extreme weather (e.g. storm, fire, and drought), human activity (e.g. deforestation, urbanization), or a combination of these (e.g. de Jong et al., 2013; Evans and Geerken, 2004; Jin and Sader, 2005; Verbesselt et al., 2010a). In many parts of the world, human activity is a growing pressure from increasing populations and their increasing demands for food, fodder, and wood for fuel and building materials (Kirby, 2013).

Terrestrial carbon uptake by vegetation is the largest global carbon flux (Beer et al., 2010), so that any change in terrestrial vegetation may also influence the carbon balance. However, it is uncertain how changes in terrestrial vegetation affect the terrestrial biosphere: it may cause the terrestrial biosphere to become a source of additional CO\textsubscript{2}, quickening the warming, or a sink for CO\textsubscript{2}, reducing the warming (Beedlow and Tingey, 2007).

The uncertainty associated with climate change processes is partly due to vegetation change. Indeed, better accounting for vegetation-climate feedbacks when analyzing change in mean climate and climate variability is an important issue for improving the representation and simulation of climate change, and for making models more realistic, accurate and spatially explicit (Wramneby, 2010). Reducing uncertainty in vegetation change knowledge, by using advanced
vegetation observation technologies and enhanced change detection methods, can improve our understanding of climate change processes and consequences, leading to better climate change predictions. One potential route to this goal is the use of time series analysis techniques applied to environmental data, and to remotely sensed data from Earth observation satellite sensors.

1.2. Remote sensing

“Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomenon under investigation” (Lillesand et al., 2008). For the remote sensing of Earth, the sensors can be handheld or mounted on towers, aircraft and satellites, and the information obtained has the potential to increase our knowledge about our planet’s land surface, oceans, atmosphere and their dynamics.

1.2.1. Earth observation satellites

Earth observation (EO) satellites\(^1\) observe the Earth from space, including the Earth’s surface, oceans, and clouds. These satellites vary according to their payload and type of orbit, and their sensors vary according to spatial resolution, spectral characteristics, and swath width. All these properties are specified during the design phase, in order to complete the satellite’s target applications. For example, in order to monitor global weather or land vegetation, the sensor may need to be in geostationary orbit, (altitude \(\sim\)36000 km) to be able to view large parts of the Earth continuously. For applications requiring high resolution imagery of smaller areas, such as land vegetation disturbance due to local events, a sensor with high spatial resolution would be combined with a satellite in Low Earth Orbit (LEO), such as a QuickBird satellite (altitude 600 km). In such an orbit it is not possible to monitor the same area continuously, because of the relative movement of the satellite with respect to the Earth.

The National Oceanic and Atmospheric Administration (NOAA)\(^2\) agency operates both geostationary and polar-orbiting satellites. One of the instruments carried on NOAA’s Polar Orbiting Environmental Satellites (POES) is the Advanced Very High Resolution Radiometer (AVHRR), which provides global collection of data

---

\(^1\) http://eospso.gsfc.nasa.gov, http://esa.int/ESA

\(^2\) http://ngdc.noaa.gov/
from the visible (VIS), near-infrared (NIR), and thermal infrared portions of the electromagnetic spectrum. The data allow global vegetation analysis.

Terra and Aqua are two further Earth observation satellites that were launched by NASA in 1999 and 2002, respectively. Terra's Sun-synchronous orbit around the Earth is timed so that it crossed the equator from north to south in the morning, while Aqua crosses the equator in the opposite direction in the afternoon. One of the instruments carried by both the Terra and Aqua satellites is the Moderate Resolution Imaging Spectroradiometer (MODIS). Terra MODIS and Aqua MODIS together view the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands ranging in wavelength from 0.4 μm to 14.4 μm and at varying spatial resolutions (250 m, 500 m, 1 km, and 5.6 km). MODIS plays an important role in meeting a wide range of scientific objectives, including the monitoring of vegetation change.

### 1.2.2. Satellite-based vegetation sensing

EO satellite sensors, irrespective of orbit, provide many types of data, including variations in the spectra of emitted radiation. To interpret radiation data gathered by satellite-mounted passive sensors, four processes must be understood: the propagation of radiation through the atmosphere; the interaction of that radiation with Earth’s surface features; the retransmission of radiation through atmosphere; and the generation of data by the sensors.

The atmosphere can strongly effect the radiation available to satellite sensors. The nature of the effects varies with the wavelength and magnitude of the radiation being sensed, the path length of atmosphere through which the radiation travels, and the atmospheric conditions. However, the atmospheric effects are caused principally by atmospheric scattering and absorption. Atmospheric scattering is the diffusion of radiation by atmospheric particles, whereas atmospheric absorption is the loss of energy to atmospheric constituents, of which carbon dioxide and particularly water vapor are the most efficient absorbers (Campbell, 1996). The acquisition of remotely sensed data is limited to ‘atmospheric windows’, which are those wavelength ranges in which the atmosphere does not block the transmission of energy. Most common remote sensors operate in one or more of the visible, infrared, or microwave portions of the electromagnetic spectrum containing atmospheric windows with high degrees of transmission (Lillesand et al., 2008). In addition to the atmosphere, several other factors create difficulties when

---

3 http://modis.gsfc.nasa.gov/
monitoring the land surface from space, including variations in sun and viewing angles, shadows, soil background (when viewing vegetation) and snow.

When electromagnetic energy hits Earth surface features, fractions of it are reflected, absorbed, and/or transmitted. The sizes of the fractions are different for different features, depending on the wavelength, and the feature’s composition and condition. The proportion of incident energy that is reflected, measured as a function of wavelength, is called spectral reflectance.

The spectral reflectance of green (healthy) vegetation differs from that of dry (non-photosynthetically active) vegetation (Fig. 1). In the visible portion of the spectrum (0.4-0.7 μm), spectral reflectance is significantly lower for green than for dry vegetation, because chlorophyll absorbs strongly in wavebands centered at about 0.45 and 0.67 μm. The reflectance of green vegetation increases dramatically at the lower limit of the NIR range (0.7-1.3 μm) whereas the reflectance of dry vegetation increases slowly.

![Typical spectral reflectance curves for green and dry vegetation in the visible (0.4-0.7 μm) and NIR (0.7-1.3 μm) portions of the electromagnetic spectrum.](image)

**Fig. 1.** Typical spectral reflectance curves for green and dry vegetation in the visible (0.4-0.7 μm) and NIR (0.7-1.3 μm) portions of the electromagnetic spectrum. Colored areas represent the red (RED) and near-infrared (NIR) wavelength bands used for AVHRR’s and MODIS’ computations of NDVI.

---

4 Modified from Lillesand et al., (2008), page 17, Figure 1.9.
The spectral reflectance of plants in the NIR range is primarily determined by the cellular structure of plant leaves. In addition to the leaf reflectance which is influenced by leaf properties such as Leaf Area Index (LAI), Leaf Angle Distribution (LAD), and the leaf’s biochemical composition, the soil’s background signal and the canopy reflectance are important factors influencing remotely sensed vegetation spectral reflectance (Jones and Vaughan, 2010).

The spectral reflectance of plants is also influenced by temporal and spatial effects. Temporal effects are any factors that change the spectral reflectance of a specific type of plant over time, such as plant healthiness, or the amount of chlorophyll production during the growing season. Spatial effects are factors that cause differences in spectral characteristics, at a given point in time, between plants of the same type at different geographic locations (e.g. soil and climate) (Lillesand et al., 2008). This means that the spectral reflectance of plants can be used to extract information about both photosynthetic activity and the temporal and spatial variability of vegetation, in the form of vegetation indices.

A spectral vegetation index is generated by combining data from one or more spectral bands, especially the VIS and NIR wavelength bands, into a single value which provides an approximate measure of relative vegetation amount (Tucker, 1979). One widely used spectral vegetation index is the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1973).

NDVI is calculated as:

\[
NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \tag{1}
\]

where \(\rho_{NIR}\) and \(\rho_{RED}\) are the spectral reflectance values measured in the NIR and red wavelength bands, respectively. NDVI ranges from +1.0 to -1.0. Vegetated areas have positive NDVI values, whereas near 0 and negative values correspond to an absence of vegetation (Myneni et al., 1995).

NDVI as a normalized index is known to be influenced by variations in atmospheric effects, clouds, soil effects, and sensor viewing and illumination angles which make the determination of the true vegetation signal difficult (e.g. Goward and Huemmrich, 1992; Huete and Jackson, 1987; Huete et al., 2002; Kaufman and Tanre, 1992). Additional general factors which significantly affect remotely sensed information, including NDVI, include plant architecture, interaction of radiation with canopy, leaf properties, and vegetation vigor (Markon et al., 1995; Pinter Jr et al., 1985). These complicating factors are normally accounted for during pre-processing of NDVI, especially in recent versions of well-documented and quality-controlled NDVI products. In spite of the corrections, some noise still remains in the NDVI data sets, mainly arising from
effects that tend to decrease NDVI including the uncorrected effects of cloud, water, snow, and shadow, and less frequently from effects that increase NDVI, including high solar and scan angels (Pettorelli et al., 2005; Viovy et al., 1992). NDVI should therefore be used with caution, especially with respect to spatial location and vegetation density. For instance, in humid tropical regions, NDVI is highly contaminated by cloud-cover effects (Achard and Estreguil, 1995). At high latitudes (> 60°), especially during winter time, deterioration in reflectance resolution can result in false high NDVI values (Goward et al., 1991; Justice et al., 1985). Near the Equator (between 30° N and 30° S), major errors occur owning to variations in solar zenith angle (Pettorelli et al., 2005). In sparsely vegetated areas (LAI < 3), NDVI is influenced mainly by soil reflectance (Huete and Jackson, 1987), whereas in densely vegetated areas (LAI > 6), the relationship between NDVI and NIR saturates (Asrar et al., 1984; Pettorelli et al., 2005). Therefore, in sparsely vegetated areas, the soil-adjusted vegetation index (SAVI, Huete, 1988) would be a good alternative, though it requires local calibration for the effects of the soil.

The Enhanced Vegetation Index (EVI, Huete et al., 2002) is an optimized index designed to have better sensitivity in high biomass regions, and to enhance the vegetation signal by decoupling the canopy background signal and reducing atmospheric and aerosol influences. It minimizes many of the contaminations problems present in NDVI and does not become saturated as easily as NDVI, so it serves as a good alternative specifically for rainforests and other areas of the Earth with large amount of green vegetation. The physically-based plant phenology index (PPI, Jin and Eklundh, 2014) can also estimate plant canopy growth efficiently, enabling improved vegetation monitoring, particularly of evergreen needle-leaf forest phenology at high northern latitudes.

Despite all the possible perturbing factors for NDVI, it remains a useful tool with which to couple climate and vegetation at large spatial scales (Pettorelli et al., 2005), and is used in many studies in the field of land surface dynamics with good reason. It has been well documented that NDVI correlates directly with vegetation productivity (Reed et al., 1994); it is linked with the fraction of Absorbed Photosynthetically Active Radiation (fAPAR) (Asrar et al., 1984; Rautiainen et al., 2010; Sellers et al., 1992); for many ecosystems it can be used to assess the direct effect of climatic conditions on vegetation biomass and phenology (Nemani et al., 2003; Roerink et al., 2003; Wang et al., 2003a; Yu et al., 2003; Zhao and Schwartz, 2003; Zhou et al., 2003); it can also be used to assess the feedback effect of vegetation on local climate (Jingyong et al., 2003; Zhang et al., 2003). NDVI has also been used to improve the prediction and impact assessment of disturbances, such as drought (Singh et al., 2003), fire (Maselli et al., 2003), flood (Wang et al., 2003b) and frost (Tait and Zheng, 2003).

NDVI data, as a proxy for vegetation status, are mainly used in time series analysis techniques throughout this thesis. However, from the perspective of time
series analysis, time series of other relevant variables (such as LAI, fAPAR) and other vegetation indices (e.g. EVI, PPI) could also be used to gain a better understanding of the dynamics of Earth’s vegetation.

1.3. Vegetation time series analysis

A time series is a sequence of observations measured at successive time points, often at equally spaced time intervals (Madsen, 2007). Time series analysis uses a range of methods for analyzing and modeling a sequence of observations. Methods for time series analysis can extract meaningful statistics and other characteristics of the data, which can in turn be interpreted to identify the nature of the phenomenon represented by the data (Chandler and Scott, 2011). Methods for time series modeling result in models that represent the temporal behavior of the system which generated the data; these models are usually used to predict the future state of the system (Madsen, 2007). Time series analysis methods can be divided into frequency-domain methods and time-domain methods. This thesis deals with time-domain methods.

An important common first step for both time series analysis and modeling methods is system identification, which includes the identification of patterns in time series data. It is usually assumed that time series data consist of systematic patterns (identifiable components) and random noise (error). Most time series patterns can be described in terms of two basic classes of components: trend and seasonality.

A trend represents a general systematic linear or (most often) nonlinear component that changes over time and does not repeat. A trend is either monotonic (Fig. 2a, b) or non-monotonic (Fig. 2c). A monotonic trend represents gradual change over time that is either entirely non-decreasing (Fig. 2a) or non-increasing (Fig. 2b), whereas a non-monotonic trend includes both types of increasing and decreasing changes over the considered time period. A monotonic trend can be either linear or nonlinear. A nonlinear trend can consist of only gradual changes (Fig. 3a), or may also contain a single (Fig. 3b) or multiple abrupt changes.
Trends in long time series of vegetation data usually consist of gradual changes, but may also include more abrupt changes (Fig. 4a, b). Gradual changes normally reflect long-term changes in other factors such as land management, land degradation, and inter-annual climate variability such as temperature variability in boreal/arctic areas (Goetz et al., 2005) or variations in water availability in arid areas (Kawabata et al., 2001). Abrupt changes are normally caused by disturbances such as fire, flood, urbanization, insect attack, or drought (Fig. 4). They can also be produced by an ecological- or climate-tipping point, where a small perturbation triggers a large response, producing a large (and often irreversible) change such as a shift in ecosystem regime (Lenton, 2013; Scheffer et al., 2001).

Seasonality (Fig. 4c) is a cyclical component which systematically repeats at set time intervals (Hill and Lewicki, 2006). Since vegetation seasonality is mainly driven by its phenology, changes in the seasonality of vegetation can reflect changes in plant phenology drivers (e.g. temperature, rainfall, and their interactions). External disturbances to vegetation have a long-term impact on vegetation trend, but can also affect plant phenology and lead to changes in seasonality. In addition to phenology, there are other seasonality variations (e.g.
albedo changes due to snow, and moisture variations) that affect the land surface and can consequently influence plant seasonality.

Overall, different types of changes (gradual, abrupt, seasonal) may be represented in a time series of remotely sensed vegetation index. Irrespective of the physical causes of the changes, they share underlying mathematical properties (e.g. linearity or nonlinearity), which allow them to be detected and quantified using the techniques of time series analysis.

**Fig 4.** (a) Time series of MODIS NDVI, 2000-2013, for the Spain site (Table 1). The time series was decomposed into (b) trend (c) seasonal and (d) remainder components using the DBEST program (Paper III). The abrupt, negative change in 2004 indicates the effect of drought on vegetation.
1.3.1. Analysis of linear change

Ordinary Least-Squares (OLS) regression is frequently used to analyze vegetation change, using time series of vegetation indices such as NDVI (e.g. Eklundh and Olsson, 2003; Herrmann et al., 2005; Olsson et al., 2005). Using this method, time is used as the independent variable, and the NDVI trend values, or a summary measure derived from the NDVI data (e.g. growing season sum, yearly maximum), is used as the dependent variable. To compute the strength and rate of vegetation changes, metrics are calculated including the slope and correlation coefficient, and their statistical significance.

The OLS regression is a parametric (model-based) method which assumes that observations are independent, and that errors are normally distributed with zero mean and an unknown but constant variance (e.g. Hess et al., 2001; Muhlbauer et al., 2009). Results obtained by this method might not therefore be reliable if these assumptions are not met (e.g. von Storch and Zwiers, 1999).

OLS has wide appeal because it is simple and can be easily implemented (Fuller, 1998; Peng et al., 2012). The main assumption of this method, that vegetation changes gradually and linearly (with constant rate) over a period of time, makes it specifically useful for analyzing linear changes. However, OLS is normally less suitable for analyzing nonlinear changes, such as detecting regions where the direction and rate of change varies within a given time period, or correctly identifying short-term changes as greening or browning patterns (e.g. de Jong et al., 2012).

The Mann-Kendall test (Kendall, 1938; Mann, 1945) is one of the most commonly used non-parametric methods for analyzing vegetation change using time series of vegetation indices (e.g. Alcaraz-ǦSegura et al., 2010; Neeti and Eastman, 2011; Sobrino and Julien, 2011). However, just like OLS, this method is most useful for analyzing linear changes, and also for reliable results requires that no autocorrelation should exist within the data series.

1.3.2. Analysis of nonlinear change

To date, many previous studies have used linear models for analyzing vegetation change (e.g. Eklundh and Olsson, 2003; Epstein et al., 2012; Hansen et al., 2013). However, the climate-vegetation system is highly dynamic and, as mentioned earlier, vegetation change does not always continue gradually in one directional pathway but can also occur nonlinearly (Lambin and Ehrlich, 1997; Piao et al., 2011). Causative factors for such nonlinear changes can be natural or human-induced, can arise from external single- or multi-year disturbances, and can affect across a wide range of spatio-temporal resolutions and extents. Examples include insect attacks, fires, storms, droughts, or clear cutting of forests (Lambin et al.,
The acceleration of processes like land degradation may also cause change in one direction but in a nonlinear fashion (Barnosky et al., 2012). Therefore, advanced methods for the analysis of nonlinear changes in vegetation index time series are needed. If only simple linear techniques are applied, events such as those listed which cause nonlinear changes, may be underestimated or even remain hidden.

New methods for analyzing nonlinear trends and changes in satellite time series data have recently been introduced (Kennedy et al., 2010; Verbesselt et al., 2010a). These developments add explanatory power for elaborating the interrelationships and driving forces in the climate-vegetation system. However, there are several issues associated with these methods that limit their utility, including their focus on specific types of sensor data, the requirement of several complex control parameters, or the requirement to specify which types of nonlinear change to detect.

1.4. Thesis objectives

This thesis aims at developing methods, tools, and a framework for analyzing change and trend in the time series of sensor data from Earth observation satellites, in order to quantify vegetation change more accurately, for better understanding of climate and vegetation variations and their interactions.

The main objectives are:

1. To develop a method suitable for detecting and estimating gradual, nonlinear changes in long-term trends in vegetation time series data (Paper I).
2. To compare parametric and non-parametric methods for estimating long-term trends in vegetation time series data (Paper II).
3. To develop a method suitable for analyzing nonlinear changes, both gradual and abrupt, in vegetation time series data (Paper III).
4. To develop a general framework for a comprehensive analysis of changes and trends in satellite sensor data series (Paper IV).

In order to address the specific objectives, four papers (papers I, II, III, and IV) are presented which embody the structure of the thesis (Fig. 5). Each paper targets one (or more) of the objectives.
Fig. 5. The organization of the study: the correspondence between papers in the thesis, and themes in the analysis of time series of satellite sensor data.
2. Materials and methods

2.1. Satellite data

2.1.1. NOAA AVHRR products

Papers I, II, and III use the Global Inventory Modeling and Mapping Studies (GIMMS) NDVI data set, which is derived from data obtained by the AVHRR instrument onboard NOAA satellite series 7, 9, 11, 14, 16 and 17 (Tucker et al., 2004; Tucker et al., 2005). The GIMMS-NDVI data set is a bimonthly composite NDVI product with global coverage at an 8 km spatial resolution, available from July 1981 to December 2006 (Tucker et al., 2004; Tucker et al., 2005).

NDVI is largely insensitive to variations in illumination intensity because it is a ratio of differences between two adjacent wavelength bands. However, NDVI is sensitive to some factors that may introduce significant variability in the AVHRR NDVI record. For example, NDVI is sensitive to effects that differ between the bands. Band calibrations have changed frequently between the five NOAA AVHRR instruments. NDVI can also be affected by the natural variability in atmospheric aerosols. For instance, the NDVI record was affected by the injection of large quantities of aerosols into the Earth’s stratosphere by two major volcanic eruptions, El Chichon in 1982 and Mt. Pinatubo in 1991. NDVI is also sensitive to the periodic variations in solar illumination angle and sensor view angles resulting from the NOAA orbits. However, the GIMMS-NDVI data set has been corrected for calibration, volcanic aerosols, view geometry, and other effects not related to actual vegetation change (Tucker et al., 2005).

In order to reduce inaccuracies related to frequent cloud cover presence and atmospheric contamination, compositing techniques are commonly used in NDVI data sets. Composite images are constructed at regular intervals by selecting the maximal NDVI value for each pixel during the composition period (Holben, 1986). These techniques reduce the effects of cloud cover and water vapor, which both strongly reduce NDVI values. The GIMMS data set was constructed based on two 15-day composites per month, for days 1 to 15 and for days 16 to the end of the month.

The GIMMS data set has a main characteristic that distinguishes it from other AVHRR NDVI data sets such as Pathfinder AVHRR land data set (James and Kalluri, 1994). The GIMMS data set has a satellite overpass time drift correction that largely eliminates the apparent variation of NDVI caused by changes in the
solar zenith angle due to the orbital drift of the afternoon NOAA satellites (Pinzon et al., 2005). The GIMMS-NDVI has also been validated against the well calibrated and atmospherically-corrected MODIS data set, for the period 2000–2007 for the semi-arid regions of the Earth (Fensholt et al., 2012). In another study, using Terra MODIS NDVI as a reference, Fensholt et al. (2009) conclude that the AVHRR GIMMS coarse resolution NDVI data set is well-suited for long-term vegetation studies of the African Sahel–Sudanian areas receiving less than 1000 mm rainfall annually.

Paper I used the AVHRR global land cover classification product with fourteen land cover classes at 8 km spatial resolution (Hansen et al., 1998; Hansen et al., 2000) for land cover determination purposes. This product was generated by the Department of Geography at the University of Maryland in 1998 by analyzing imagery from AVHHR satellites acquired between 1981 and 1994.

2.1.2. MODIS products

MODIS NDVI\(^5\) is another source of NDVI data which complements NOAA AVHRR NDVI products, and provides continuity for historical time series applications (Bédard et al., 2006). The MODIS NDVI products are computed from atmospherically corrected bi-directional surface reflectances that have been masked for water, clouds, heavy aerosols, and cloud shadows.

Paper IV used the L3 Global MODIS Terra NDVI 16-Day composites data with a 250 m spatial resolution (MOD13Q1 V005 product) for the period 2000-2013. This product provides frequent information at a similar spatial scale to the majority of changes in human-driven land cover (Townshend and Justice, 1988), giving data-quality information for each pixel. This means that each pixel contains data about aerosol quantity (low to high), cloud coverage, possible snow/ice, possible shadow, and data on whether it has been flagged as (for instance) land or ocean. This allows users to achieve reliable results by extracting and assessing the pixel quality, before they use them in their studies. The MODIS Quality layer (MOD13Q1 Pixel Reliability) data were used to select data of reliable quality.

Papers I, II, and IV used the MODIS/Terra + Aqua land cover data set (MCD12C1, V005) to determine the land cover type in the study areas. This is a yearly L3 global Climate Modeling Grid (CMG) product with a spatial resolution of 0.05° (~5600 m), which provides information on dominant land cover types and on the sub-grid frequency distribution of land cover classes (Friedl et al., 2010).

\(^5\) http://lpdaac.usgs.gov/products/modis_products_table
The CMG product is derived using the same algorithm that produced the V005 Global Land Cover Type product (MCD12Q1) with a 500 m resolution.

Paper I also used the Terra MODIS Vegetation Continuous Fields (VCF) product (MOD44B), which is a global sub-pixel-level representation of estimated surface vegetation cover. This product contains proportional estimates for woody vegetation, herbaceous vegetation, and bare ground which add up to represent 100% of the ground cover (Hansen et al., 2003). The VCF product is generated annually using monthly composites of Land Surface Reflectance data (with 250 and 500 meters resolution) from all seven bands of the MODIS sensor on board NASA's Terra satellite, and Land Surface Temperature. The Global Land Cover Facility (GLCF) edition of this product (Collection 4, Version 3) with a 0.5 km spatial resolution for the year 2001 (Hansen et al., 2006) was used.

2.2. Vegetation change hotspots

Vegetation change hotspots are defined as areas or regions with a relatively rapid or unexpected change in vegetation compared to their surroundings. North Africa was used as the study area in paper I, and a number of other regions, mostly also vegetation change hotspots, were used in papers II, III, and IV (Table 1). Paper III also used simulated NDVI data.

2.2.1. North Africa

North Africa includes the Sahelian, Sudanian, and Guinean climatic zones representing the humid tropical and dry tropical zones of sub-Saharan Africa, north of the Equator (Fig. 6). These are transition zones between the arid Sahara to the north and the humid tropics to the south. The Sahelian zone with its semi-arid tropical savanna ecosystems is marked by a steep north–south gradient in mean annual rainfall (Le Houerou, 1980).

The rainfall gradient is expressed on the ground as a continuum of change in vegetation species from the Saharan zone, with low and variable rainfall amounting to 100–600 mm per year on average and very sparse vegetation cover, to the Sudanian and Guinean zones with mean annual rainfall between 600 and 1500 mm, and with increased ground cover, taller vegetation and a greater proportion of woody species (Le Houerou, 1980; White, 1983).

Average yearly rainfall totals declined abruptly in the Sahel and Sudanian zones beginning in the 1960s. The considerable fluctuations in rainfall at inter-annual and -decadal time scales make the Sahelian region the most dramatic example of climate variability that has been directly measured (Hulme, 2001).
Table 1. Study area and study site descriptions including name, latitude and longitude (Lat, Long), dominant land cover type, and in which paper they are included. The dominant land cover type was determined using the MODIS Land Cover data with the University of Maryland (UMD) classification scheme (paper I) and International Geosphere-Biosphere Programme (IGBP) legend (MCD12C1, Collection 5, Type_1) for the years 2005 (paper IV) and 2006 (paper II).

<table>
<thead>
<tr>
<th>Paper</th>
<th>Study area/site name</th>
<th>Location (Lat, Long)</th>
<th>Land cover (MODIS IGBP/UMD)</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>North Africa (including Sahel)</td>
<td>5–25, -20–45</td>
<td>Grasslands</td>
<td>Fig. 6</td>
</tr>
<tr>
<td>II</td>
<td>Chad</td>
<td>12.900, 18.200</td>
<td>Grasslands</td>
<td>Fig. 7</td>
</tr>
<tr>
<td></td>
<td>Sierra Leone</td>
<td>8.650, -10.925</td>
<td>Woody savannas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ivory Coast</td>
<td>7.454, -6.521</td>
<td>Croplands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Burkina Faso</td>
<td>13.570, -2.520</td>
<td>Grasslands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Benin 1</td>
<td>11.151, 2.805</td>
<td>Savannas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Benin 2</td>
<td>11.296, 2.659</td>
<td>Savannas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>44.900, 9.250</td>
<td>Croplands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nigeria</td>
<td>7.223, 9.714</td>
<td>Woody savannas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>South Sudan</td>
<td>9.187, 33.350</td>
<td>Savannas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ethiopia</td>
<td>11.514, 41.714</td>
<td>Open shrublands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>37.550, -6.540</td>
<td>Woody savannas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iraq 1</td>
<td>30.850, 46.750</td>
<td>Permanent wetland</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iraq 2</td>
<td>31.240, 47.140</td>
<td>Open shrublands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iraq 3</td>
<td>30.950, 46.700</td>
<td>Sparsely vegetated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iraq 4</td>
<td>31.106, 47.159</td>
<td>Open shrublands</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Iraq</td>
<td>28–38, 38–50</td>
<td></td>
<td>Fig. 11</td>
</tr>
<tr>
<td>IV</td>
<td>Burkina Faso</td>
<td>13.820, -2.050</td>
<td>Grassland</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sweden (Sarek)</td>
<td>67.150, 17.930</td>
<td>Open shrublands</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sweden (Kindla)</td>
<td>59.750, 14.900</td>
<td>Evergreen needle-leaf forest</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>37.550, -6.300</td>
<td>Woody savanna</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senegal</td>
<td>16.050, -14.420</td>
<td>Grassland</td>
<td></td>
</tr>
</tbody>
</table>

On other hand, the climatic constraints (scarcity, variability and unpredictability of rainfall which increase from south to north) are the most important controlling factors of the Sahelian ecosystem (Herrmann et al., 2005; Nicholson, 1995). These combined with frequent drought and widespread greening trend observed in the

6 Source of data: http://lpdaac.usgs.gov
region (e.g. Eklundh and Olsson, 2003) make the Sahel a good example both of a vegetation change hotspot and of nonlinear vegetation dynamics.

Fig. 6. Land cover map for the study area used in paper I - North Africa, 2006. The MODIS land cover data set (MCD12C1, V005) was used with the University of Maryland (UMD) classification scheme. Areas with mean yearly NDVI < 0.1 were masked out.

2.2.2. Change hotspots

Paper II used GIMMS-NDVI data series for fifteen study sites located in Africa, Spain, Italy, and Iraq, covering a range of temporal vegetation trajectory patterns over the period 1982-2006 (Table 1, Fig. 7). The patterns include gradual change (Chad, Sierra Leone, Ivory Coast, Burkina Faso, Benin1, and Benin 2), no-trend (Italy, Nigeria), no net change in annual NDVI throughout the period (South Sudan, Ethiopia), and abrupt change (Spain, Iraq 1, 2, 3, 4).

Source of data: http://lpdaac.usgs.gov
For paper III, Iraq was selected as the study area. Iraq’s geography consists of four main zones: desert in the west and southwest, rolling upland between the upper Tigris and Euphrates rivers in north central Iraq, highlands and mountains in the north and northeast, and alluvial plain from northwest to southeast, through which the Tigris and Euphrates flow. The landscape of Iraq has been subjected to many changes over the last two to three decades (UNEP, 2001), mainly due to frequent natural hazards, including sand storms (Draxler et al., 2001) and floods (Hamdan et al., 2010), and human activities, including forced migration, social conflict, and wars (especially in northern and southeastern Iraq). Fig. 8a shows one example of a vegetation change hotspot at a sample site located in southeastern Iraq, observed using a GIMMS-NDVI data series, 1982-2006. Note how the NDVI increases irregularly in 2004 after a long period of low vegetation since 1990. Similarly, a considerable part of Iraq’s land surface has undergone major changes in the recent past. In paper III, these areas were detected, the corresponding changes were quantified; and change drivers were briefly explained.
Fig. 8. Examples of vegetation change hotspots, as indicated by NDVI. (a) GIMMS-NDVI data series, 1982-2006, at a sample site in southeast Iraq. The increase in the NDVI value in 2004 indicates the impact of action by local residents, who renewed irrigation by Tigris and Euphrates River waters after a long period of vegetation reduction starting in 1990. (b) MODIS-NDVI data series for 2000-2013 for a sample site in Spain (Table 1). The abrupt, negative change in 2004 indicates the effects of drought on vegetation.

Paper IV used MODIS NDVI data for five sites (each one with a 10×10 km land area) in different countries, including one site each in Burkina Faso, Spain, and Senegal, and two sites in Sweden (Table 1). Among these sites, the Spanish site is marked as another vegetation change hotspot due to several recent droughts, notably in 2003 (Garcia-Herrera et al., 2007; Ruiz-Sinoga and Martinez-Murillo, 2009). Fig. 8b shows MODIS NDVI time series at a sample site in the area for the period 2000-2013. As shown, NDVI decreased sharply in 2004 due to the effects of the drought on vegetation.

2.2.3. Simulated data

For testing and evaluation purposes, simulated NDVI time series including different types of change were used (paper III). NDVI time series (including seasonal cycles) and NDVI trends (without seasonal cycles) were simulated at two sites with different biomes (forest and cropland) for a period of 300 months, corresponding to the real GIMMS data, 1982-2006. The simulations contained artificial processes and events, such as sudden defoliation, drought with typical pre- and post-event life phases, and noise. The changes in NDVI, and the durations of the changes for each phase, were specified based on observations of corresponding biomes in real conditions.
To generate noise, which often originates from geometric and atmospheric errors in satellite data, a random number generator was used which had a zero-mean ($\mu=0$) normal distribution and different (at least three) values of standard deviation ($\sigma = 0.01, 0.04, 0.07$ NDVI units), corresponding to different noise levels (low, medium, strong respectively) (Lhermitte et al., 2011; Roderick et al., 1996).

2.3. Analyzing vegetation trends

Earth’s vegetation is a dynamic and often nonlinear system, so that methods which detect only linear trends may not always give useful results: in particular, some change features may remain hidden. One option for analyzing nonlinear trends to identify details of latent change features is to fit nonlinear functions. However, to the best of our knowledge, no attention has previously been given to the fitting of higher order polynomial models to long time series of remotely sensed greenness information as tools for understanding vegetation change (Paper I, Fig. 5).

In order to investigate the performance of parametric and non-parametric methods for analyzing vegetation trends, two standard methods – the OLS method and the Mann-Kendall test – were compared in terms of their performance in detecting and quantifying trends with different temporal patterns (Paper II, Fig. 5).

To further analyze temporal variations in satellite time-series data, including characterizing both linear and nonlinear changes, a time series segmentation method named DBEST (Detecting Breakpoints and Estimating Segments in Trend) was also developed (Paper III, Fig. 5).

The trend analysis studies in papers I, II, III and the methods developed within each were finally integrated with existing relevant tools in order to develop and propose a general framework underpinning the comprehensive analysis of trends in satellite sensor time series (Paper IV, Fig. 5). This final paper provides users with the most recent knowledge, theoretical guidelines and practical tools for detecting and analyzing vegetation changes using satellite imagery.
3. Results and discussion

3.1. Paper I: Automated mapping of vegetation trends with polynomials using NDVI imagery over the Sahel

This paper presented and evaluated an approach based on polynomial-fitting to account for nonlinear changes, by deriving metrics for both linear and nonlinear changes in long-term time series of remotely sensed vegetation indices. The method divides vegetation change into linear and nonlinear trend behaviors, and subdivides the nonlinear trends into cubic, quadratic, and concealed trends. Concealed trends include all cases where there exists at least a pattern of significant increase and then decrease (or decrease and then increase) in the vegetation index over the time period of the study, but without a net change in vegetation amount over the period as a whole. The method was applied to annual GIMMS-NDVI time series for the Sahel for the period 1982-2006.

The results showed that just over half of the study area (51.9%) exhibited statistically significant trends, of which 27.1% were linear and 24.8% were nonlinear (Fig. 9). The linear trends were dominated by increasing trends (22.2% of the whole), and were distributed in an east-west band across the Sahel. This is not a new finding and confirms previous studies that reported an overall increase in vegetation amount throughout much of the Sahel, as a result of a recovery from the droughts of the mid-1980s (e.g. Eklundh and Olsson, 2003; Herrmann et al., 2005; Heumann et al., 2007). Areas with nonlinear changes, especially areas with cubic and quadratic trends, were less frequent and were widely scattered throughout the study area. Nevertheless, they are credible patterns of changes. They can flag unique vegetation change hotspots that are associated with distinct vegetation types (e.g. woody shrublands) and may also be associated with direct human impact on land surface caused by, for instance, rapid urbanization or strife.
3.2. Paper II: Comparing ordinary least-squares regression and a non-parametric method for estimating trends in multi-year NDVI data

The OLS and polynomial-fitting methods are both parametric methods. However, non-parametric methods have also been widely used for trend detection. This paper systematically compares the use of parametric and non-parametric methods for analyzing long-term trends in annual GIMMS-NDVI data series, for data with a span of 25 years (1982-2006) from fifteen sites in Africa, Spain, Italy, and Iraq (Table 1). Trends were separately estimated using OLS regression (parametric) and the Mann-Kendall trend test combined with the Theil-Sen slope estimator (non-parametric) (Sen, 1968; Theil, 1950). The trend type at each site was determined using the polynomial-fitting scheme developed in paper I.

The results of this comparison show that both OLS and Mann-Kendall Theil-Sen methods performed similarly, in terms of trend slope value and its statistical significance, at sites with an overall steep trend that had gradually increased or decreased over the whole time period of the study (Fig. 10a, b, c, d, e, f). These sites had linear (subfigures a, b, c), quadratic (subfigure d), and cubic (subfigures e and f) trend types, as determined by the polynomial-fitting scheme. At sites with weak slopes of the no-trend type, both methods avoided flagging any trends (Fig. 10g, h). At sites with nonlinear changes (quadratic) but with no net change (i.e. concealed trends), both methods failed to detect any significant change (Fig. 10i, j).
Fig. 10. Annual GIMMS-NDVI time series (in blue) for the period 1982-2006, with the trend types derived by the polynomial-fitting scheme (in red) at the study sites used in paper II. The dashed lines (in black) represent non-significant linear fits in the subfigures with concealed trends.
At sites with abrupt (non-monotonic and nonlinear) changes resulting in strong departures from linearity and consequently high degree polynomials (Fig 10k, l, m, n, o), both methods failed to exhaustively flag significant trends. However, the trend slope estimated by the parametric approach reflected the impact of the abrupt change, whereas the non-parametric approach ignored this impact. All in all, this paper concludes that there is a lower risk of failing to identify the presence of trends in annual NDVI data using OLS regression, compared to the Mann-Kendall Theil-Sen method.

3.3. Paper III: Detecting changes in vegetation trends using time series segmentation

Papers I and II both focus on technical aspects of analyzing long-term trends in the time series of satellite data-derived vegetation indices. A more complete analysis of these time series also needs techniques that can detect and quantify short duration events. This paper introduces a recently developed algorithm for analyzing vegetation time series, Detecting Breakpoints and Estimating Segments in Trend (DBEST), which has two main application domains: generalizing vegetation trends to main features, and characterizing abrupt and non-abrupt changes in vegetation trends. DBEST allows the user to capture a wide range of trend features, from details to main features, by setting the number of the largest changes, or the magnitude of changes, that will be detected.

DBEST was tested and evaluated using simulated NDVI data series at two sites which had undergone different types of changes. The evaluation results demonstrated DBEST’s efficiency, in that it quickly and robustly detected changes and accurately estimated the change characteristics (timing and magnitude). DBEST was also applied to a time series of GIMMS-NDVI images for Iraq for the period 1982-2006 in order to detect major changes. They were small (< 0.1 NDVI) over large areas (87% of the land area). Larger non-abrupt changes (between 0.1 and 0.3 NDVI) were detected in northern and eastern Iraq (10% of the land area). Abrupt changes (> 0.2 NDVI) were detected in relatively small areas (1%) in the southeast (Fig. 11). All the major changes detected were statistically significant, except small changes (< 0.1 NDVI) in scattered areas (2% of the land area).
3.4. Paper IV: A framework for temporal trend analysis of satellite sensor data

This paper synthesizes the Ph.D. research. It proposes a framework for the temporal trend analysis of satellite sensor data. The framework was developed from a conceptual model of the main aspects of time series trend analysis (Fig. 12). This model considers concerns raised during the major steps of trend analysis, including identifying the research problem, preparing the data, generating variables, and selecting the analysis method. Based on these considerations, the framework guides the analyst through the four steps outlined above to select efficient (families of) methods which address the research aims. The algorithms developed in paper I and paper III are part of the methods employed in the framework. Other methods, both parametric and non-parametric (as discussed in paper II), are also included in the framework to support a comprehensive analysis.

Fig. 11. (a) Major change in trend of monthly GIMMS-NDVI, 1982-2006, Iraq. (b) NDVI time series (in black), computed trend (in blue), and the detected major change (level-shift) (red dot) for one sample pixel located in an area with abrupt change in southeast Iraq, as shown in subfigure (a). Only pixels with at least 75% valid data points (i.e., quality flag value zero) throughout the whole time period were analyzed. Areas with a mean monthly NDVI < 0.1 were masked out. The changes’ characteristics were estimated by the DBEST program developed in paper III.
of trends and changes of all types, including annual, seasonal, linear, nonlinear, gradual, and abrupt (Fig. 13).

**Fig. 12.** The conceptual model of the trend analysis of satellite data series, from which the proposed framework was developed (Paper IV).

To verify the framework, it was employed for four 10 x 10 km case studies for analyzing photosynthetic activity in Burkina Faso, plant greenness in Spain, seasonal snow cover in Sweden, and biomass accumulation in Senegal. The results of the case studies demonstrated the utility of the framework for the comprehensive analysis of temporal trends in satellite sensor data.
3.5. General discussion

The polynomial-fitting scheme which was developed to detect and estimate gradual, nonlinear change in long-term trends in vegetation time series data was straightforward to implement. It was tested using annual GIMMS–NDVI observations for North Africa, including the Sahel, for the period 1982–2006. The evaluation successfully identified areas with nonlinear changes in vegetation, and quantified the changes. The method is simple to apply: it is automated to a high
degree. This is especially useful for applications with large data volumes because it does not require the careful selection of thresholds for parameterization purposes.

Despite containing trends of all types, the majority (> 75%) of the annual GIMMS-NDVI data over the large study-area (North Africa, paper I) meet the basic assumption of normality, as tested by the Jarque–Bera (Jarque and Bera, 1987), Student's t and Bartlett tests. This confirms the validity of the trend results obtained. On the other hand, paper II showed that for trend detection in time series of NDVI data, the Mann–Kendall test combined with the Theil–Sen slope estimator dose not generally outperform OLS. These results reconfirm the suitability of the polynomial-fitting method, a development of OLS, for the analysis of nonlinear change.

Prior to this study (papers I, II), to the best of the author’s knowledge, no attention had been given to comparing parametric and non-parametric methods for detecting and estimating gradual, nonlinear changes in long-term trends in satellite-derived vegetation data series. However, awareness of the importance of identifying shorter-term fluctuations and sudden events has recently increased. For example, LandTrendr is a recently developed software package that uses arbitrary temporal segmentation to divide long-term trends into piecewise linear segments, in order to represent spectral change in forested ecosystems (Cohen et al., 2010; Kennedy et al., 2010). BFAST (Breaks For Additive Seasonal and Trend) is another tool for monitoring change features embedded in long time series of NDVI data (Verbesselt et al., 2010a). In BFAST, nonlinearity in the trend component is simplified into a number of individual trend segments in order to identify sudden structural shifts in a NDVI time series.

Despite some merits, both these methods have weaknesses. LandTrendr has been specifically developed to use Landsat data for monitoring forest disturbance and its generalization to use time series of other sensors in other ecosystems has not yet been implemented. BFAST does not require assumptions regarding land cover types or reference periods (e.g. de Jong et al., 2012; Verbesselt et al., 2010a; Verbesselt et al., 2010b) but it does assume that discontinuities can always be treated as instantaneous changes, to be modelled as single time-step shifts. This assumption can be seen as a weakness of BFAST, which limits its ability to approximate nonlinear trends in real-world data, where changes of different types and durations may exist.

To overcome these weaknesses, the DBEST method was developed (paper III). In contrast to LandTrendr, DBEST can be applied to the time series of different remotely sensed vegetation indices for different ecosystems. The user of DBEST is able to detect and characterize changes of interest, including realistic abrupt and non-abrupt changes of different durations and in any order within a particular sequence of occurrences. These developments allow the user to choose the
generalization scheme and extract information about both the main features and details (i.e. long-term processes and short duration events) at different spatial and temporal scales.

The framework for temporal trend analysis of satellite sensor data (paper IV) has been shown to be a useful development: it accounts for the different elements of trend analysis in order to address a wide range of research problems, as demonstrated by the four case studies. The framework provides a systematic approach to the trend analysis of satellite sensor data, and provides guidelines and practical tools to assist with the major steps of an analysis. This leads to more accurate trend/change information, because using the framework may reduce the uncertainty associated with a trend analysis process.

Despite any merit which may be possessed by the tools and framework developed in this thesis, the results obtained by them must be interpreted with caution. This is because there is always a possibility that some of the trends are due to factors other than actual changes in vegetation. For example, Kaspersen et al. (2011) caution against performing trend analysis on NDVI for the northern Sahel due to the increased influence on the signal of soil color variations, sensor effects and dust. As another example, there is concern about deleterious effects on the resultant NDVI caused by the fact that the GIMMS-NDVI data set comprises data from six NOAA satellites, combined with orbital drift artifacts. Moreover, Forkel et al. (2013) corroborate that the estimation of trends from NDVI time series may differ substantially depending on the source satellite dataset, the corresponding spatio-temporal resolution, and the applied statistical method. As another example, work by Wessels et al. (2012) on the limits of VI trends to quantify physical processes demonstrates that trend results should be carefully interpreted.

Continuous efforts to enhance trend detection techniques and improving their efficiency are needed. High spatial resolution images have become easily available in the past decade and will hopefully remain so in future, such as for the forthcoming Sentinel-2 satellite. This has motivated the remote sensing community to focus increasingly on the time-series analysis of these data. Therefore, supplementing the current techniques with additional tools for purposes such as gap filling, handling irregular data series, and processing very large volume data is valuable for improving change detection performance, especially at fine spatial resolutions.

3.6. Contribution to state-of-the-art

On many occasions, the OLS method is applied with little consideration as to its suitability for monitoring vegetation over large areas, where there may exist areas with different change types. This can lead to poor metrics for the detected changes,
especially in the latter areas. Paper I considers these issues and presents an entirely automated method, as a development of OLS, for detecting change metrics in a long time series of satellite-derived vegetation index data that would not be identified by OLS method. The problem and proposed solution is not unique to vegetation trend analysis, but can be used for the trend analysis of general time series data (e.g. temperature or precipitation). This paper also highlights the importance of areas with vegetation change hotspots – areas to which no attention has yet been given – by investigating associations the hotspots tend to have with land cover type and land cover change estimates. In addition, the paper pays attention to areas that have undergone no long-term net change, but show significant fluctuations over the time period. Overall, this paper contributes precise and more accurate information to the emerging field of Land Change Science.

Although both the OLS (parametric) and Mann-Kendall (non-parametric) approaches are very common and have been widely used, paper II is the first study to compare their performances when analyzing trends in the time series of vegetation indices. Comparisons made from different perspectives (e.g. slope steepness, statistical significance, method assumption validity, trend linearity, abrupt or gradual change type) can be of specific interest for those working with trend and change detection approaches, and shed light on the way for future research. In a broader usage, this paper raises analysts’ awareness about the strengths and weaknesses of each method when applied to NDVI time series, especially when applied to large areas that experience multiple types of change due to the range of local climates or human made events.

The main goal of paper III was to fill an important need when detecting trends in satellite sensor time series, by developing and implementing a new change detection algorithm (DBEST) which can identify and characterize both abrupt and non-abrupt changes, in a highly configurable fashion. This paper demonstrated the success of DBEST as a fast, accurate and flexible tool for both detecting abrupt and non-abrupt changes and estimating their timing and magnitude. This advances the ability to detect and map both short duration events and long-term processes on the Earth’s surface, at different spatial and temporal scales.

Paper IV helps users to structure their knowledge with regard to the trend analysis of satellite sensor data series. The paper raises awareness among analysts about the important aspects of change detection, and presents a decision-making system that guides them through a series of steps to select appropriate processing methods/tools. It also constitutes a general structure to which new methods can be added. The framework enables more accurate and efficient analysis of vegetation dynamics as observed from space. Improved analysis can help better explain and understand land ecosystem dynamics, bio- and geophysical processes and their responses to climate variation and natural and anthropogenic activities. Better analysis therefore serves as an initial step to better forecasts through better understanding of climate-vegetation feedback mechanisms.
4. Conclusions

The main conclusion of this thesis is that developing improved time series methods for identification and analysis of vegetation change hotspots, where vegetation responds to climate and disturbance in a complex and nonlinear fashion, provides more accurate information. This can lead to a better understanding of vegetation dynamics, especially in sensitive regions of the Earth.

The main conclusions for each of the listed objectives and their corresponding papers are as follows:

1. The polynomial-fitting scheme, an entirely automated method, successfully derives nonlinear change metrics from long time series of vegetation index data from satellite sensors. The method, a development of OLS, is used where the aim is to detect gradual changes, related for example to inter-annual climatic variability, land degradation or land management changes. Assessment of the method showed that it generates credible patterns of vegetation change that would not be detected by OLS. Although the nonlinear trends were infrequent and scattered throughout the Sahel they can flag unique vegetation change hotspots that are associated with distinct vegetation types (e.g. woody shrublands) and may also be associated with direct human impact on the land surface caused by, for example, rapid urbanization or strife. (Paper I)

2. OLS regression (parametric) and Mann-Kendal trend test combined with Theil-Sen slope estimator (non-parametric) methods performed similarly for identifying and estimating trends for gradually increasing or decreasing changes with steep slopes. For abrupt (non-monotonic and nonlinear) changes, both methods failed to exhaustively flag significant trends. However, OLS captured the impact of the abrupt change on the trend slope whereas the Mann-Kendall Theil-Sen approach did not. Both methods failed to detect concealed trends, which highlights the importance of the polynomial-fitting method for detecting and classifying nonlinear changes. (Paper II)

3. The validation and evaluation results demonstrated that DBEST detects trend changes, determines their type (abrupt or non-abrupt), and estimates their timing, magnitude, number, and direction. DBEST's was shown to be efficient for providing accurate, robust and cost effective results. Vegetation change hotspots in the form of abrupt changes (> 0.2 NDVI) leading to significant shifts in the level of NDVI were detected in relatively small areas (~1% of the land area) in southeast Iraq. This was
consistent with the likely impact on vegetation of the uncontrolled replenishment of the Tigris and Euphrates River waters, which was implemented by local residents. (Paper III)

4. The framework, developed for temporal trend analysis of satellite sensor data, carefully considers the concerns involved in the major steps of trend/change detection process, including defining the research problem, selecting appropriate satellite sensor data and the appropriate variable from that data, and selecting suitable analysis methods. The application of the framework to the four case studies demonstrated its utility, both in supporting users with guidelines and in providing them with the most suitable (family of) methods for each specific research aim. (Paper IV)
5. Future studies

Possible future studies include:

- integrating the developed trend analysis methods into existing software for fast processing of remotely sensed time series data (TIMESAT, Jönsson and Eklundh, 2004),

- applying the new methods to time series of vegetation index data across the globe to identify vegetation change hotspots in time and space,

- elucidating the main drivers of change in regions with known vegetation changes (e.g. arctic and boreal areas), and

- providing insights about vegetation change to agencies dealing with drought early warning and natural resource planning in sensitive regions of the Earth.
Acknowledgements

I would like to thank Jonathan Seaquist, my main supervisor, and Lars Eklundh, Jonas Ardö, and Per Jönsson, my supporting supervisors, for their excellent scientific guidance, encouragement, and great support throughout the time of this thesis.

A special thanks to Ford Cropley, for proofreading the final version of thesis overview and the last two manuscripts.

Many thanks to Petter Pilesjö and Ulrik Mårtensson, who gave me the opportunity to come to Sweden, and for their support and help, especially when I was new to Lund.

I acknowledge all colleagues, PhD candidates, personnel, and administration staff at the Department of Physical Geography and Ecosystem Science and GIS Centre, for creating a friendly atmosphere and helping whenever needed.

I am very grateful to my dear parents; to Soraya, my beloved wife; to Soraya’s parents; and Sarina and Sevina, my lovely daughters.

This thesis was funded by the European Union funding programme Erasmus Mundus “External Cooperation Window” (EMECW lot8), and also by Lund University.
References


Zhenlin Yang, 2012: Small-scale climate variability and its ecosystem impacts in the sub-Arctic.

Ara Toomanian, 2012: Methods to improve and evaluate spatial data infrastructures.

Michal Heliasz, 2012: Spatial and temporal dynamics of subarctic birch forest carbon exchange.


Julia Bosiö, 2013: A green future with thawing permafrost mires? : a study of climate-vegetation interactions in European subarctic peatlands. (Lic.)

Anders Ahlström, 2013: Terrestrial ecosystem interactions with global climate and socio-economics.

Kerstin Baumanns, 2013: Drivers of global land use change: are increasing demands for food and bioenergy offset by technological change and yield increase? (Lic.)

Yengoh Genesis Tambang, 2013: Explaining agricultural yield gaps in Cameroon.


Wårlind, David, 2013: The role of carbon-nitrogen interactions for terrestrial ecosystem dynamics under global change: a modelling perspective.

Sundqvist, Elin, 2014: Methane exchange in a boreal forest: the role of soils, vegetation and forest management.

Falk, Julie Maria, 2014: Plant-soil-herbivore interactions in a high Arctic wetland: feedbacks to the carbon cycle.

Hedefalk, Finn, 2014: Life histories across space and time: methods for including geographic factors on the micro-level in longitudinal demographic research. (Lic.)

The wealth of remotely sensed vegetation data made available by satellites in recent decades has allowed researchers to look for vegetation trends and changes over short and long time series, and at spatial scales ranging from local to global. The majority of studies to date have used linear analysis to detect the grossest features of change: that is, whether vegetation has on average increased or decreased, and the rate of that change. However it is now widely recognized that vegetation often responds to climate and disturbance in a complex, nonlinear fashion.

This thesis aims to advance the analysis of nonlinear trends in time series of vegetation data from Earth observation satellite sensors. This is accomplished by developing fast, efficient methods suitable for large volumes of data. A set of methods, tools, and a framework are developed and verified using data from regions containing vegetation change hotspots.