Remote Sensing of Environment 156 (2015) 182–195

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Detecting changes in vegetation trends using time series segmentation

Sadegh Jamali [a,*], Per Jönsson [b], Lars Eklundh [a], Jonas Ardö [a], Jonathan Seaquist [a]

[a] Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, SE-223 62 Lund, Sweden
[b] Division of Materials Science and Applied Mathematics, Malmö University, Sweden

ARTICLE INFO

Article history:
Received 3 December 2013
Received in revised form 8 September 2014
Accepted 8 September 2014
Available online 16 October 2014

Keywords:
Vegetation dynamics
Change detection
Satellite imagery
Trend analysis
Time series
Segmentation
DBEST
GIMMS
NDVI

ABSTRACT

Although satellite-based sensors have made vegetation data series available for several decades, the detection of vegetation trend and change is not yet straightforward. This is partly due to the scarcity of available change detection algorithms suitable for identifying and characterizing both abrupt and non-abrupt changes, without sacrificing accuracy or computational speed. We propose a user-friendly program for analysing vegetation time series, with two main application domains: generalising vegetation trends to main features, and characterizing vegetation trend changes. This program, Detecting Breakpoints and Estimating Segments in Trend (DBEST) uses a novel segmentation algorithm which simplifies the trend into linear segments using one of three user-defined parameters: a generalisation-threshold parameter δ, the m largest changes, or a threshold β for the magnitude of changes of interest for detection. 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 was tested and evaluated using simulated Normalized Difference Vegetation Index (NDVI) data at two sites, which included different types of changes. Evaluation results demonstrate that DBEST quickly and robustly detects both abrupt and non-abrupt changes, and accurately estimates change time and magnitude.

DBEST was also tested using data from the Global Inventory Modeling and Mapping Studies (GIMMS) NDVI image time series for Iraq for the period 1982–2006, and was able to detect and quantify major change over the area. This showed that DBEST is able to detect and characterize changes over large areas. We conclude that DBEST is a fast, accurate and flexible tool for trend detection, and is applicable to global change studies using time series of remotely sensed data sets.

1. Introduction

Vegetation is an important component of terrestrial ecosystems (e.g. Prentice et al., 2000; Sitch et al., 2003). It is now widely recognized that vegetation can respond to climate and disturbance in a complex, non-linear fashion (e.g. Lambin and Ehrlich, 1997). External disturbances to ecosystems, including insect outbreaks, fires, forest clear-cutting, and conversion of natural grasslands to cultivated crops, increase the complexity of vegetation variations across a wide range of spatio-temporal resolutions and scales (e.g. Higgins et al., 2002; Lambin et al., 2003). There is a corresponding need for data and tools which can properly capture and represent such complex dynamics, at a level of detail which is appropriate to the question under consideration. One flexible and simplifying idea is to allow the user to define the boundary between signal and noise. Developing analysis tools that embody this idea can help us to characterize and understand vegetation dynamics with much greater efficiency.

Vegetation dynamics are often monitored at different spatial and temporal scales through vegetation index data series, such as GIMMS (Global Inventory Modeling and Mapping Studies), MODIS (Moderate Resolution Imaging Spectroradiometer), and SPOT/Vegetation (System Pour l'Observation de la Terre) data sets. One widely used index is the Normalized Difference Vegetation Index (NDVI), which is an indicator of the vegetation's greenness (Rouse et al., 1973). Temporal trend analysis of NDVI has proved particularly useful for monitoring and characterizing the response of land cover to phenomena with a range of time scales, from abrupt natural or anthropogenic events (Verbesselt et al., 2009), to seasonal variability of plant phenology caused by changes in temperature and rainfall regimes (Heumann et al., 2007), through to gradual inter-annual climate changes (Jacquin et al., 2010). Similar analyses of NDVI trends have also been used to investigate the relationship between NDVI and Leaf Area Index (LAI), a key variable which is functionally related to plant biomass production. Studies have indicated that this relationship is not completely linear, particularly for high LAI values (e.g. Carlson and Ripley, 1997; Fan et al., 2009).

Non-linear temporal trends in NDVI can be studied by a number of techniques, including time series segmentation (Keogh et al., 2004). Series segmentation refers to approximating a time series by a set of piecewise straight-line segments, in the process eliminating noise but

* Corresponding author. Tel.: +46 46 222 1725; fax: +46 46 222 0321.
E-mail address: Sadegh.Jamali@nateko.lu.se (S. Jamali).

http://dx.doi.org/10.1016/j.rse.2014.09.010
0034-4257/

preserving the salient details of the trend. Two significant merits of segmentation are that first, it approximates complex signatures to extract their basic features (Zeileis et al., 2003) and second, it increases the efficiency of data storage and further computation (Keogh et al., 2004). The size of the output set can vary between one and the size of the original time series, and is a key parameter in determining the type of information available to the user, from most generalised/compressed information (one segment) to all possible details (all data points).

A particular type of segmentation technique, which always outputs one segment, is the ordinary least-squares (OLS) technique. It has been widely used for assessing long-duration changes in vegetation greenness using NDVI time series (Eklundh and Olsson, 2003; Fensholt et al., 2009; Fuller, 1998; Nemani et al., 2003; Peng et al., 2012). However, the implicit assumption that vegetation always changes gradually and constantly over the entire study period is a critical drawback of this method (e.g. Jamali et al., 2014). Consequently, it is impossible using OLS to identify periods in which the rate of change, or even the sign of the rate of change, varies within a given time period: for example, the existence of short-duration nonlinearities such as 'greening' or 'browning' patterns may be completely obscured because they have been averaged out altogether (de Jong et al., 2012). On the other hand, using the original time series in further analysis is not cost effective and may also, in cases when only the main features of the vegetation change are of interest, carry unused detailed information and noise.

Fortunately, to date at least two alternative approaches between the two mentioned extreme cases have been introduced for detecting trends and estimating changes in vegetation time series. The first approach, Landsat-based detection of Trends in Disturbance and Recovery (LandTrendR), uses arbitrary temporal segmentation to divide long-term trends into piecewise linear segments for the representation of spectral change in forested ecosystems (Cohen et al., 2010; Kennedy et al., 2010). LandTrendR focuses solely on Landsat derived data, and requires several complex control parameters to be specified. Moreover, different signal-to-noise ratios and spectral properties of different instruments may require alteration of the algorithms and parameters used in LandTrendR, and its generalisation for use with time series of other sensors in other ecosystems has not yet been implemented.

The second approach, Breaks For Additive and Seasonal Trend (BFAST), was one of the first general methods for detecting changes in time series data, focusing on the trend and seasonal components in long-term NDVI data series, at spatial scales ranging from continental to global (e.g. de Jong et al., 2013; de Jong et al., 2012; Verbesselt et al., 2010a; Verbesselt et al., 2010b). It employs a generic approach by assuming that a time series can be modelled in terms of its trend, seasonal, and remainder components, including breaks in the trend and seasonal components. In BFAST, nonlinearity in the trend component is also simplified into a number of individual trend segments, in order to identify sudden structural shifts. The simplified trend is therefore composed of segments with gradual changes, separated from each other by relatively brief, abrupt changes (de Jong et al., 2012; de Jong et al., 2013; Verbesselt et al., 2010a). Our experience using BFAST (most recent version 1.4.4) shows that although the user can control the number of breaks to be detected, they do not have full control over the level of trend generalisation (whether basic features or detailed information will be extracted), or over the change magnitude that is detected. More importantly, the discontinuities may not be instantaneous for which modelling them as one time-step shifts can be seen as a limitation.

Thus, to the best of our knowledge, no attention has yet been paid to segmenting time series of vegetation indices to generalise trends and detect both abrupt and non-abrupt changes, in a way which allows the user to control the segmentation. Ideally the user should be able to choose the generalisation scheme and extract information about both the main feature and details. The user should also be 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.

The overall aim of this paper is to propose a novel approach for segmenting and analysing trend changes in time series of vegetation indices (VI). The method, called Detecting Breakpoints and Estimating Segments in Trend (DBEST), provides both generalised trend segments and estimates of the characteristics of both abrupt events and long-term processes. DBEST has been tested and evaluated using simulated NDVI data series. We also applied DBEST to a GIMMS–NDVI dataset (1982–2006) for Iraq, to detect major changes, determine change type (abrupt or non-abrupt) and estimate the characteristics (time, magnitude) of the changes in a spatio-temporal framework. The method allows for fast, flexible and accurate estimations of trends, essential from a global change perspective.

2. Data and study area

We have developed our method using a commonly used time-series data source: the NASA GIMMS data set. Then, we applied the methods to selected study areas that highlight certain temporal features. The methods are not specific to either the data type or the sites, which merely represent commonly used data and illustrative cases.

2.1. GIMMS data

The GIMMS data set is a twice-monthly composite NDVI product with global coverage at an 8-km spatial resolution, which is available for the 25.5 year period from July 1981 to December 2006 (Tucker et al., 2004; Tucker et al., 2005). The GIMMS data were derived from the output of the AVHRR (Advanced Very High Resolution Radiometer) instrument on board NOAA (National Oceanographic and Atmospheric Administration) satellite series 7, 9, 11, 14, 16 and 17. The NDVI is computed from the red and near infrared bands of the AVHRR sensor, and is correlated with photosynthetic activity and the LAI of green vegetation. The GIMMS data have been corrected for variations arising from calibration, view geometry, volcanic aerosols, and other factors not related to actual vegetation change (Tucker et al., 2005). The GIMMS data have also been validated against the well-calibrated and atmospherically-corrected MODIS data set for the period 2000–2007 for the Earth's semi-arid regions (Fensholt et al., 2012). Despite the corrections, the GIMMS data may contain residual invalid measurements, but these are well indicated using quality flags (de Jong et al., 2012). Based on the quality flags, we only considered pixels with at least 75% valid data points (i.e., flag value zero) through the whole time span.

2.2. Study area

The case study was Iraq, selected due to the fact that large parts of the country's land surface have undergone major changes in the recent past. 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 in the north and northeast with mountains, and alluvial plain through which the Tigris and Euphrates flow from northwest to southeast. About 77.7% of Iraq's land area is not viable for agriculture. 0.3% is forest and woodlands situated along the extreme northern border with Turkey and Iran. The remaining 22% are used for a range of agricultural activities (Schnepf, 2004).

Over the last two to three decades, the landscape of Iraq has witnessed many changes (UNEP, 2001). These have been caused by frequent natural hazards, such as sand storms (Draxler et al., 2001) and floods (Hamdan et al., 2010), and human activities, such as forced migration, social conflicts, and wars (especially in south-eastern and northern

Iraq). The high probability of abrupt vegetation disturbances and recoveries made Iraq the ideal study area to test DBEST.

3. Methods

In this section, the DBEST algorithm is introduced, and then approaches used for validating and evaluating DBEST's performance are explained. The DBEST algorithm consists of two main parts: trend estimation and trend segmentation.

3.1. The trend estimation method

For a given VI time series (including seasonal variations), the temporal trend is derived using the Seasonal-Trend decomposition procedure based on Loess (STL) (Cleveland et al., 1990), implemented in an unpublished version of the TIMESAT software package (Jönsson and Eklundh, 2004). STL is a fast filtering procedure able to deal with missing values, for decomposing a time series into trend (low frequency variation), seasonal (variation at or near the seasonal frequency), and remainder (remaining variation) components. TIMESAT is a software package for analysing time series of satellite sensor data, which makes it possible to investigate the seasonality of the data and their relationship with dynamic properties of vegetation such as phenology.

Before using STL, the existence of major discontinuities in the VI level is examined. To do this, the absolute difference in VI between each pair of consecutive data points, starting from the time series' start point, is compared with a user-defined *first level-shift-threshold* (θ_1) (Table 1). If the absolute difference is greater than the threshold value, a second criterion tests whether the change led to a considerable shift in the mean VI level which persists throughout a user-defined time period, the *duration-threshold* (ϕ). The second criterion is valid if the absolute difference in the means of the VI data, calculated over a period ϕ before and after the current data point, is greater than a user-defined *second level-shift-threshold* (θ_2).

If both tests are valid, the current data point is marked as a candidate *level-shift* point. The same procedure is repeated for every data point in the time series. The candidate points are then sorted into descending order according to the absolute value of the shift in the VI level. The first point in the sorted list is marked as the most important level-shift point. For the second and subsequent candidate points, in addition to the two mentioned criteria, a third criterion should be fulfilled in order to be marked as the next important level-shift points. The third criterion checks if the spacing between the candidate point and any previously detected level-shift point is at least the duration-threshold ϕ. Bai and Perron (2003) and Zeileis et al. (2003) also recommend a fraction of data needed between successive breaks. At any detected level-shift point, the VI time series is divided into two parts, before and after.

The STL decomposition is then performed separately for each part of the time series divided by the detected level-shift points (called hereafter *separate STL decomposition*) if one or more level-shift points have been detected: if no such points are found, it is performed once for the entire time series. The separate STL decomposition procedure often results in more precise trend and seasonal components (especially around the detected level-shift points) because the shifts in the VI level are not deformed by the smoothing methods used in the STL method (Cleveland et al., 1990).

If the given time series is already deseasonalized (i.e. we are examining the trend component) then level-shifts are similarly detected but of course no STL decomposing is needed.

3.2. The segmentation method

For the trend time series (either computed using the STL or already deseasonalized), with a number of observations N ($N > 2$), single time-step differences in the forward and backward directions are calculated at every time-point i ($2 \le i \le N-1$), as follows (Fig. 1a):

$$\Delta VI_{(i-1,\ i)} = VI_{(i)} - VI_{(i-1)} \tag{1}$$

$$\Delta VI_{(i,\ i+1)} = VI_{(i+1)} - VI_{(i)} \tag{2}$$

For each point i the *peak/valley detector* function (f) is then computed, based on the continuity of the sign of the two differences:

$$f_{(i)} = \begin{cases} 1, & \text{if } \operatorname{sign}\left(\Delta VI_{(i-1,i)}\right) = -\operatorname{sign}\left(\Delta VI_{(i,i+1)}\right) \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

As specific cases, the peak/valley detector function is set to one at the start point ($i = 1$) and zero at the end point ($i = N$) of the time series. The signum function sign (x) equals $+1$ for x greater than zero, -1 for x less than zero, and zero for x equal to zero. Clearly, the trend direction changes at time points at which the peak/valley detector function equals one (except the time series' start point). These points are called *peak* and *valley* points.

For all points of the time series, a second *turning point detector* function (g), is computed based on the peak/valley detector function and an iterative *criterion*, explained as follows. In the first iteration, for every pair of successive peak or valley points (e.g. $P_{(q)}$ and $P_{(i)}$ or $P_{(i)}$ and $P_{(z)}$ in Fig. 1a), a straight line passing through the two points is computed. Perpendicular distances (d) are computed from all the intermediate non-peak and non-valley points ($f = 0$) to the derived line. On each side of the line (above and below), the intermediate point with maximum perpendicular distance (e.g. $P_{(j)}$, $P_{(w)}$, $P_{(l)}$) is selected, and this distance is compared with a user-defined *distance-threshold* (ε), to be discussed later. If the distance is greater than the threshold (e.g. $d_j > \varepsilon$

Table 1
The user-defined thresholds used in DBEST.

Threshold	Description
First level-shift-threshold (θ_1)	The lowest absolute difference in (VI) data between the level-shift point and next data point
Duration-threshold (ϕ)	The lowest time period (time steps) within which the shift in the mean of the data level, before and after the level-shift point, persists; and, the lowest spacing (time steps) between successive level-shift points.
Second level-shift-threshold (θ_2)	The lowest absolute difference in the means of the data calculated over the period ϕ before and after the level-shift point
Distance-threshold (ε)	The lowest perpendicular distance from farthest data point to the straight line passing through every pair of successive peak and valley points. (Note: a default value for this parameter is estimated in DBEST using Eq. (7)).
Change number (m)	Number of major breakpoints included in the generalised trend (generalisation algorithm); number of greatest changes of interest for detection (change detection algorithm)
Change magnitude (β)	The largest variation within a segment after generalisation (generalisation algorithm); the lowest magnitude of changes of interest for detection (change detection algorithm)
Generalisation-threshold (δ)	The highest level of trend generalisation compared to the base level (the least simplified fit derived using all detected breakpoints in the fitting)
Statistical significance level (α)	Statistical significance level used for testing significance of detected changes

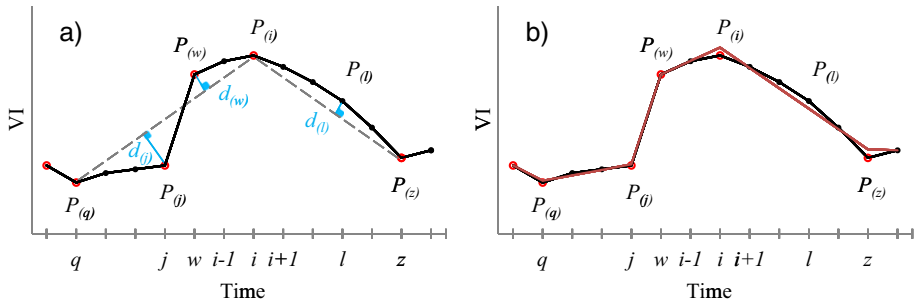

Fig. 1. A schematic representation of part of a given VI trend time series, in which $P_{(i)}$ denotes the vegetation index value at time-point *i*. (a) — Consecutive original points (solid in black), straight-lines (dashed in grey) passing through pairs of consecutive peak and valley points at which $f = 1$ (e.g. $P_{(q)}$, $P_{(i)}$ and $P_{(z)}$), perpendicular distances from other points ($f = 0$) to the straight-lines (e.g. $d_{(j)}$, $d_{(w)}$, $d_{(l)}$), and the *turning points* (rings in red). (It is assumed that $d_j > \varepsilon$, $d_w > \varepsilon$, and $d_l < \varepsilon$). (b) — the model trend (segments in red) obtained by DBEST when the change detection algorithm is used, which detects $P_{(j)}$ as the major breakpoint. The $P_{(q)}$, $P_{(j)}$, $P_{(w)}$ and $P_{(i)}$ turning points are valid (i.e. breakpoint), as confirmed by BIC. All the valid turning points plus $P_{(z)}$ (as the next turning point after $P_{(i)}$) are shared by adjoining segments.

and $d_w > \varepsilon$), the turning point detector function equals one ($g = 1$) at the selected point (i.e. $P_{(j)}$ and $P_{(w)}$), otherwise it is zero ($g = 0$) (e.g. $P_{(l)}$ with $d_l < \varepsilon$). At non-selected points, the turning point detector function equals zero as well. So the criterion is fulfilled only at the farthest points above or below the line, provided that its perpendicular distance is above the threshold value.

$$g_{(i)} = \begin{cases} 1, & \text{if } f_{(i)} = 1 \\ 1, & \text{if } f_{(i)} = 0 \text{ and the criterion is fulfilled.} \\ 0, & \text{if } f_{(i)} = 0 \text{ and the criterion is not fulfilled.} \end{cases} \quad (4)$$

As special cases, the first straight line passes through the time series' start point and the first peak or valley point; similarly the last line passes through the last peak or valley point and the time series' end point. For the second iteration, for every pair of successive points at which $g = 1$, a new straight line passing through the two points is computed, and the rest proceeds exactly as for the first iteration. The second iteration may add some new points to the set of points at which $g = 1$. This same procedure is repeated for further iterations, until no new points with $g = 1$ are detected. It is noted that if the set of points with $g = 1$ does not already contain the level-shift points, though this is unlikely, they are now added to this set ($g = 1$).

Hereafter, time points at which the turning point detector function equals one ($g = 1$) are called *turning points*: the set of turning points comprises the time series' start point, the level-shift points, the peak and valley points and the points at which the criterion is fulfilled (Fig. 1). If u is the number of turning points found for a given time series, then larger values of u indicate more peaks and valleys and/or more points fulfilling the criterion and/or more level-shift points in the trend time series. The turning points' positions (except the time series' start point) indicate times when trend changes occur. Nevertheless, all the u turning points are not necessarily associated with substantial changes. As will be discussed later, the user can select changes that have considerable magnitude, after which the validity of the turning points is tested from the fit quality standpoint.

Valid turning points are defined as turning points that significantly reduce the residual sum of squares of a least-square fit to the VI trend time series (explained in next paragraph) and do not result in over-fitting. The number of valid turning points can be determined by minimizing an information criterion. The Bayesian Information Criterion (BIC) (Schwarz, 1978) is a criterion for optimal model selection among a finite set of models. When computing the least-squares fit, adding more turning points reduces the residual sum of squares but also increases the number of model parameters, which may result in an over-fitting problem. Both BIC and the Akaike Information Criterion (AIC) (Akaike, 1974) resolve this problem by introducing a penalty term for the number of parameters in the model. We have chosen the BIC because it has been argued that the AIC usually overestimates the number of breaks, while BIC is a suitable selection in many situations (e.g. Zeileis et al., 2002; Zeileis et al., 2003). Given a finite set of estimated models, the preferred or optimal model is the one with the lowest BIC value.

Before using BIC, the *trend local change* function, designated h and defined in Eq. (5), computes the difference in VI trend between successive turning points, and assigns it to the current turning point: it is zero otherwise for non-turning points. For all points of the time series the function h is defined as follows:

$$h_{(i)} = \begin{cases} 0, & \text{if } g_{(i)} = 0 \\ VI_{(z)} - VI_{(i)}, & \text{if } g_{(i)} = 1 \end{cases} \quad (5)$$

where i and z are the indices of the current and next turning points respectively. As an exception, we choose $z = N$ for computing h for the last turning point. All u turning points are then sorted into descending order according to the magnitude of their trend local change (i.e. absolute value of h). The least-squares fit to the VI trend time series is computed for the entire time series. The fitting is started with no turning point selected from the list to be considered in the fit (i.e. linear fit), and then is repeated for cases where different numbers of the sorted turning points (from just the first, to all u of them) are included in the fit. The BIC is computed for the fit obtained for each case. The valid turning points are specified as the s ($s \le u$) points for which the corresponding fit minimizes the BIC.

The least-squares fit to the trend time series is computed using the selected turning points, fitting either one single straight-line (when no turning point is selected) or a composite line (when one or more are selected). In computing the composite line, each selected turning point and the next turning point after it in the trend time series are treated as data points to be shared by the adjoining linear segments. This fitting method leads to a continuous, piecewise trend composed of linear segments. Piecewise linear modelling, as a special case of non-linear regression (Venables and Ripley, 2002), is often used to approximate complex phenomena and extract the basic features of the data (Verbesselt et al., 2010a; Zeileis et al., 2003).

The least-squares fitting is explained based on an example. Consider Fig. 1b with 14 data points, where $q = 2, j = 5, w = 6, i = 8$, and $z = 13$. The $P_{(q)}$, $P_{(j)}$, $P_{(w)}$, $P_{(i)}$, and $P_{(z)}$ turning points are here assumed as data points shared by adjoining segments. The composite line consists of six segments (number of the assumed turning points plus one) and is

obtained from the following regression models (6.1)–(6.14).

$$VI_{(1)} + \varepsilon_1 = b \tag{6.1}$$

$$VI_{(2)} + \varepsilon_2 = a_1 + b \tag{6.2}$$

$$VI_{(3)} + \varepsilon_3 = a_1 + a_2 + b \tag{6.3}$$

$$VI_{(4)} + \varepsilon_4 = a_1 + 2a_2 + b \tag{6.4}$$

$$VI_{(5)} + \varepsilon_5 = a_1 + 3a_2 + b \tag{6.5}$$

$$VI_{(6)} + \varepsilon_6 = a_1 + 3a_2 + a_3 + b \tag{6.6}$$

$$VI_{(7)} + \varepsilon_7 = a_1 + 3a_2 + a_3 + a_4 + b \tag{6.7}$$

$$VI_{(8)} + \varepsilon_8 = a_1 + 3a_2 + a_3 + 2a_4 + b \tag{6.8}$$

$$VI_{(9)} + \varepsilon_9 = a_1 + 3a_2 + a_3 + 2a_4 + a_5 + b \tag{6.9}$$

$$VI_{(10)} + \varepsilon_{10} = a_1 + 3a_2 + a_3 + 2a_4 + 2a_5 + b \tag{6.10}$$

$$VI_{(11)} + \varepsilon_{11} = a_1 + 3a_2 + a_3 + 2a_4 + 3a_5 + b \tag{6.11}$$

$$VI_{(12)} + \varepsilon_{12} = a_1 + 3a_2 + a_3 + 2a_4 + 4a_5 + b \tag{6.12}$$

$$VI_{(13)} + \varepsilon_{13} = a_1 + 3a_2 + a_3 + 2a_4 + 5a_5 + b \tag{6.13}$$

$$VI_{(14)} + \varepsilon_{14} = a_1 + 3a_2 + a_3 + 2a_4 + 5a_5 + a_6 + b. \tag{6.14}$$

Rewriting the models in matrix form and utilizing standard methods (Press et al., 1990) gives the values of the coefficients a_1, a_2, a_3, a_4, a_5 and a_6 (slopes of the six linear segments) and b (model value at the starting data point) and the $(1-\alpha)\times 100\%$ confidence interval for each coefficient (α is a user-defined significance level, Table 1) (Rousseeuw and Leroy, 1987). If 0 is not contained in the interval, the coefficient is significant. Note that the composite line is continuous, but it does not necessarily pass through the turning points considered in fitting.

The s valid turning points, which minimise the BIC, are called *breakpoints*. Note that the breakpoints, as a subset of the turning points, can be abrupt (i.e. level-shift) or non-abrupt.

The next step allows the user to select either all breakpoints or a subset of them for output fit. The selection depends on the reason for segmentation. If the reason is trend generalisation, there are three options for the selection. If the required trend should reflect only a specified number m ($0 \le m \le s$) of the greatest changes (first option), the first m breakpoints having the largest magnitude of the trend local change are selected (Table 1). If the trend should contain segments with variations equal to or lower than the user-defined threshold β (second option), then the k breakpoints with change magnitudes larger than β are selected. The third option is to simplify the trend up to a user-defined level (*generalisation-threshold* δ), compared to the least simplified trend fit derived using all s breakpoints as the base level. In this case, the first r ($r = \text{ceil}\,(s - \delta \times 0.01 \times s)$) breakpoints are selected.

Alternatively, if the reason for segmentation is the detection and characterization of changes (either the m greatest changes, or all changes greater than the magnitude β), then as another field application of the segmentation algorithm, all detected breakpoints (s) will be selected to be considered in the output fit (i.e. DBEST selects the best fit with the lowest BIC score). The user can choose a larger m (or lower β or lower δ) if their interest is in trend details or short local events, or choose a smaller m (or higher β or higher δ) if their interest is in the main features of the trend, or long-duration processes.

Finally, the least-squares fit to the trend time series (output fit) is computed using the selected breakpoints, fitting either a straight-line (when $m = 0$ or $k = 0$ or $r = 0$) or a composite line (when $m \ge 1$ or $k \ge 1$ or $r \ge 1$). The outputs of the trend generalisation algorithm are the generalised trend, number of segments, and fit quality measures (root-mean-square-error (RMSE) and maximum-absolute-difference (MAD)). For any detected change, the time of the corresponding breakpoint (which is called the *break date*) is considered to be the *start time*, and the time of the next turning point in the trend time series is considered to be the *end time*. The *change duration* is the time between the start and end times, and the *change value* is estimated by subtracting the fitted VI at the start time from the fitted value at the end time. The sign of the change value indicates the change direction, whether the slope is increasing or decreasing. The *change type* output denotes whether the detected change is abrupt or non-abrupt dependent on whether the corresponding breakpoint is a level-shift point or not, respectively. The *change significance* output indicates whether the detected change is significant or not dependent on whether the corresponding segment computed by the regression models (6.1)–(6.14) is statistically significant or not.

An overview of DBEST is shown in Fig. 2, including some practical examples. For instance, in order to generalise a trend time series at maximum level (fit one segment), set $\delta = 100\%$ in the generalisation algorithm: the rest of the analysis is performed as explained above (Fig. 2, left panel, Example 1). As a second example, to generalise the trend into main segments including only one major breakpoint, set $m = 1$ (Fig. 2, left panel, Example 2). Thirdly, to generalise into segments containing variations ≤ 0.2 in NDVI units, set $\beta = 0.2$ (Fig. 2, left panel, Example 3). Obviously, setting $\beta = 0$ or $m = s$ or $\delta = 0$ leads to the least generalised form of the trend (i.e. fit derived using all the s breakpoints). As an example for the change detection algorithm, to detect the three major changes set $m = 3$ (Fig. 2, right panel, Example 4). Here, all the s breakpoints are used in the fitting, but the outputs are characteristics of only the three changes of interest. This method, which can specify the changes of interest through either the change number or magnitude parameters, preserves the highest fitting quality as well as the actual shape of the change, because the output fit is composed of all valid segments.

As mentioned, the distance-threshold (ε) is a user-defined parameter, which can be set to a value between zero and maximum absolute difference between data values of the given time series (two for NDVI). However, values larger than the maximum influential value have no effect on the segmentation process. By the maximum influential value we mean the value that if any greater value than it is used for the distance-threshold, no turning points through fulfilling the distance criterion is detected; only peak and valley points plus time series' start point and level-shift points will be detected as turning points (minimum possible u). The lower the value of ε relative to the maximum influential value, the larger the number of turning points (u) that will be detected. Setting $\varepsilon = 0$ causes all data points (except time series's end point) to be detected as turning points (maximum possible u). A default value for ε, however, is estimated in DBEST as follows:

$$\varepsilon = C \cdot R^* \tag{7}$$

where $C = 3$ and R^* is the root mean square error for the fit obtained using the peak and valley points plus the time series' start point and the level-shift points (i.e. the distance-threshold is initially set to the maximum absolute difference value). The constant 3 in Eq. (7) is

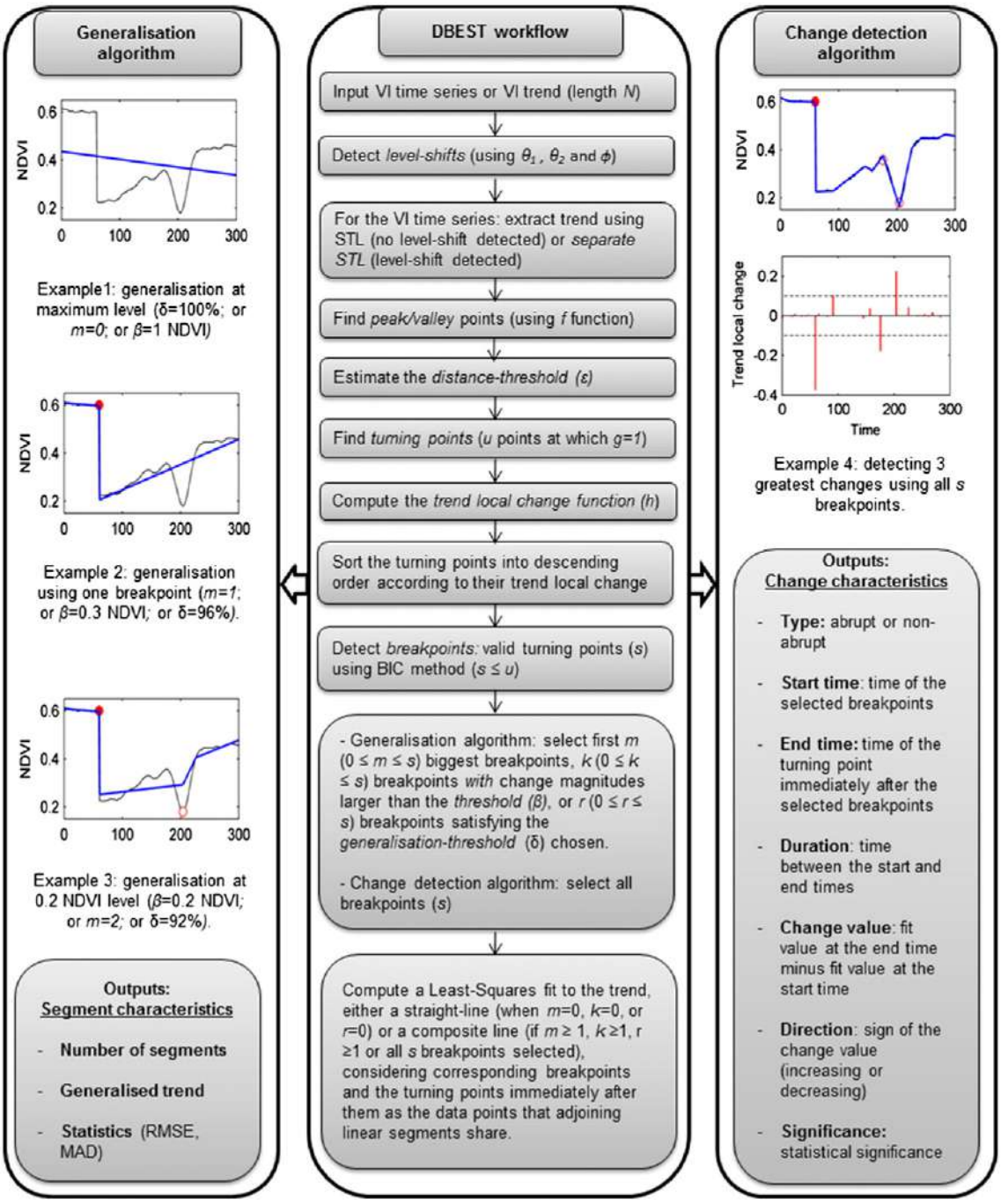

Fig. 2. Flowchart of DBEST for the two main applications of generalising trends (left panel), and detecting and characterizing changes (right panel) in VI time series.

based on empirical tests on different values for C (explained later on in Section 4.2.). However, interpretation of the estimated ε is that it is a statistical boundary ($\alpha \sim 0.01$) between outlier and normal residuals of the obtained fit. This means that data points with outlier residuals will most likely be detected as turning points through fulfilling the distance criterion, and data points with normal residuals will not be detected. Hence, it is unlikely that noise leads to false detection of turning points because they normally correspond to small residuals.

3.3. Validation

DBEST must be tested against remotely sensed satellite imagery in order to investigate its reliability. However, as Verbesselt et al. (2010a) note, the validation of change detection methods is not straightforward. This is due to the scarcity of independent reference sources for a wide range of potential changes, and also because the type and number of changes detected are never represented by field validated single-date maps (Kennedy et al., 2007). We therefore simulated NDVI data influenced by different biophysical driving processes and events (both strengthening and disturbing plant greenness, with both abrupt and gradual changes), and different types of plant response, in order to test DBEST under unambiguous and controlled conditions. In addition to using simulated data, validation based on actual satellite data remains necessary because of the challenges existent in perfectly simulating a real environment. Our satellite data were taken from Iraq, where major changes in the NDVI have been observed (Hamdan et al.,

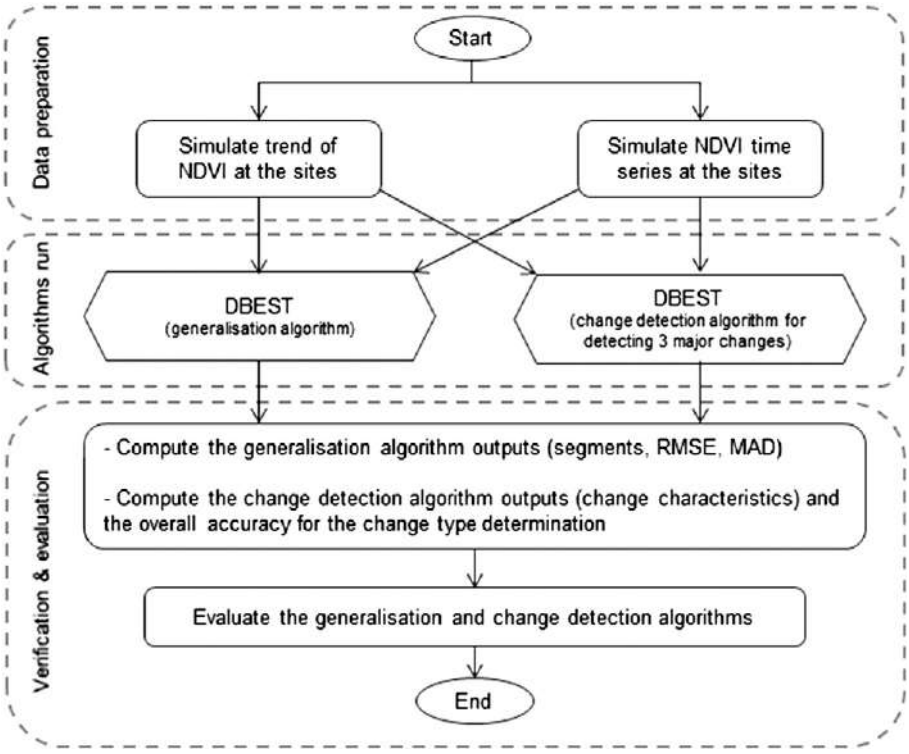

Fig. 3. Validation and evaluation workflow for DBEST. All computations were repeated for each level of noise ($\sigma = 0.01, 0.04, 0.07$). RMSE = Root Mean Squares Error, MAD = Maximum Absolute Difference.

2010; Richardson et al., 2005). Fig. 3 illustrates the overall workflow of the validation and evaluation processes.

3.3.1. Simulation

We simulated both NDVI trends (without seasonal cycles) and NDVI time series (including seasonal cycles) at two sites with different biomes (forest at site 1 and cropland at site 2) for a period of 300 months (corresponding to the real GIMMS data in Iraq) by using artificial processes or events (with typical pre- and post-event life phases) and noise (Table 2). We specified NDVI changes and their durations for each phase, based on observations of corresponding biomes in real conditions.

At site 1, we considered two typical disturbance events in a series of six phases: a stable condition followed by a sudden defoliation (causing a significant level-shift in NDVI), a gradual re-growth phase, a second defoliation, a recovery phase, and finally a return to stable plant greenness in the last phase. At site 2, one plantation and one drought, both with significant impacts on NDVI, were used in respectively second and fourth phases just after stable conditions in the first and third phases. Similar to site 1, a recovery phase and finally a return to stable plant greenness were considered in respectively fifth and sixth phases (Table 2).

DBEST works for both types of trend (non-seasonal) and VI time series (including seasonal cycles). We therefore simulated both types of data series. For each site, we simulated NDVI trend data by computing a straight-line segment for each phase using an appropriate change and duration to calculate the segment's slope and intercept (Table 2). The initial NDVI levels, as starting points in the first phases, were assumed to be 0.6 (site 1) and 0.15 (site 2). For generating noise – which often originates from geometric and atmospheric errors in satellite data – we used a random number generator with a zero-mean ($\mu = 0$) normal distribution and different values (at least three) for 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). The generated noise for each level was then separately summed with the computed trend at each site. Finally, the derived trends including high frequency variations were smoothed using a locally-weighted regression (loess) smoother (Cleveland and Devlin, 1988) after which likely level-shifts were detected. As recommended for the STL model (Cleveland et al., 1990), we used a first degree polynomial model for the loess smoothing (with a smoothing span of 23 divided by the total number of data points within each part of the time series). For evaluation purposes, the NDVI trends were simulated with 500 iterations for each level of noise at each site.

To generate VI time series with seasonal trends and different types of noise, we simulated seasonal periodic cycles of NDVI at each site, and added them to the earlier simulated noisy trends (before applying the loess smoothing). Finally, we used the separate STL procedure ($\theta_1 = 0.1, \theta_2 = 0.2$ NDVI units and $\phi = 24$ monthly time steps) to detect likely level-shifts and extract trend components. In order to simulate the

Table 2
Magnitude and duration (year) of changes used for simulating trends of monthly NDVI at site 1 (forest) and site 2 (cropland).

	1st phase	2nd phase	3rd phase	4th phase	5th phase	6th phase
Site 1	Stability	Sudden defoliation	Gradual regrowth	Defoliation	Recovery	Stability
Change/duration	0/5	−0.4/(1/12)	0.15/(9 + 11/12)	−0.2/2	0.3/2	0/6
Site 2	Stability	Planting	Stability	Drought	Recovery	Stability
Change/duration	0/5	0.3/1	0/13	−0.2/2	0.15/2	0/2

seasonal component for each biome, we used a symmetric Gaussian function with amplitudes of maximum 0.2 (site 1) and 0.25 (site 2) for healthy seasons. The maximum values were derived by using the STL function for some real forest and cropland GIMMS–NDVI time series from the study area, extracting their seasonal components, and approximating the amplitudes. The Gaussian type function has performed well when used to calculate seasonality parameters by fitting the function to NDVI time series (Jönsson and Eklundh, 2002). The amplitude values used for simulating the disturbed seasons and seasons in the first phase at site 2 were assumed to be lower than the corresponding values for healthy seasons.

3.3.2. Spatial application of the major change characterization

We generated monthly NDVI image series for a period of 25 years (1982–2006) by converting the original twice-monthly GIMMS–NDVI data ($25 \times 12 \times 2 = 600$ images) to monthly data (300 images) using the maximum-value composite (MVC) method (Holben, 1986), which is similar to the method used in GIMMS. Although the compositing method can reduce cloud and other noise effects, the conversion was solely made because of the easier interpretation of monthly over twice-monthly results, and is not a prerequisite for DBEST. The temporal trend of the monthly NDVI data series was derived using the separate STL decomposition procedure ($\theta_1 = 0.1$, $\theta_2 = 0.2$ NDVI units and $\phi = 24$ monthly time steps). We then applied DBEST's change detection algorithm to the trends of the monthly NDVI data series in Iraq in order to detect, estimate, and map major change characteristics (change type, timing and magnitude) during the period 1982–2006.

3.4. Evaluation

To evaluate DBEST's performance under different generalisation schemes, we assigned two values (zero and one) to the number of breakpoints, and two values (0.1 NDVI and 0) to the trend local change threshold (β), in order to create four different schemes (run settings) respectively called: *no-breakpoint*, *major breakpoint*, *0.1 NDVI threshold*, and *all-inclusive breakpoints*. As the fifth scheme, we generalised the trends using different values (from 0 to 100%) for the generalisation-threshold (δ). We then applied DBEST's generalisation algorithm to the simulated trends under these schemes (Fig. 3). For the evaluation, we quantified the accuracy of the data fit of each scheme by computing the average RMSE and the MAD between the simulated trend values and the fitted values (on the generalised trend) obtained by DBEST. To do this, the RMSE and MAD values obtained for each individual trend (simulated) were averaged over the 500 iterations of the simulation. In addition, the number of segments in the generalised trend, averaged over the 500 iterations, was computed for each scheme.

Furthermore, we assessed DBEST's performance for detecting and characterizing changes through quantifying the accuracy of the break date, change magnitude estimates and output fits. For this exercise, errors were defined as differences between the estimated (DBEST) and true values (Table 2). First we computed the error of the time and magnitude of the three largest changes detected by DBEST for each individual simulated trend, and calculated the RMSE for the estimates over the 500 iterations. These RMSEs were finally averaged over the three changes. To quantify the data fit accuracy, the average RMSE and MAD for the output fits were computed, in the same way as in the previous paragraph. We also assessed DBEST's performance for determining change type (abrupt or non-abrupt) using the overall accuracy descriptive statistic (e.g. Congalton, 1991). To compute this statistic, the types of the three changes detected by DBEST for each individual trend were compared with the types of the simulated changes. If the change type for all changes was determined correctly we assigned 100% to the statistic for the individual trend, otherwise lower dependent on the number of correct change types. The statistics became 0% if no change type was determined correctly. We then averaged the statistic over the 500 simulations.

Similarly, all the abovementioned measures for evaluating DBEST performance were computed using the simulated NDVI time series. As one exception, the average maximum-absolute-remainder (MAR) was computed instead of MAD, explicitly in order to assess any seasonal variations in fit accuracy.

The average execution time for DBEST was also measured. We ran DBEST (on an ordinary computer with 2.7 GHz Intel Core i7 CPU and 4 Gb RAM) in Matlab (version R2013a).

All the evaluation and method comparison computations (summarized in Fig. 3) were repeated for each noise level, in order to assess the models' robustness. Finally, we used DBEST to detect and map major changes in the Iraq data, as a practical application of the method.

4. Results and discussion

4.1. Validation and evaluation using the simulated data

4.1.1. Trend generalisation

DBEST's trend generalisation performance was tested using the five schemes described in Section 3.4. Fig. 4 shows results for one sample of the simulated trends for each site. The no-breakpoint ($m = 0$ or $\delta = 100\%$) scheme produced the most generalised fits (straight-lines) (Fig. 4a, f). The major events were successfully captured using the major breakpoint scheme (Fig. 4b, g), which represented the data by three significant segments, of which the central ones represented the major event (the sudden defoliation at site 1, and planting at site 2). By looking for magnitude changes of 0.1 NDVI (Fig. 4c, h), the number of detected breakpoints increased to $k = 4$ (site 1) and $k = 3$ (site 2), meaning that more segments (8 segments for site 1 and 6 segments for site 2) were generated. The all-inclusive breakpoints scheme generated the highest number of segments (14 segments for site 1 and 15 segments for site 2), and the least generalised fits ($\delta = 0$), while preserving details about minor events that carry valid information as confirmed by BIC (Fig. 4d, i). However, these minor events should be carefully interpreted as they may be artefacts produced by local variations (e.g. noise) or trend analysis in the presence of underlying environmental trends other than VI, as stated by Wessels et al. (2012).

As expected, the no-breakpoint scheme led to the lowest fit accuracies for all noise levels (average RMSE ~ 0.10–0.13 NDVI, and average MAD ~ 0.16–0.23 NDVI). The major breakpoint scheme was much more accurate (by about 76% for the average RMSE and 54% for the average MAD). Segmentation using the threshold scheme (0.1 NDVI) ranked second, closely after the all-inclusive breakpoints scheme that always ranked first (average RMSE < 0.01 and MAD < 0.03 NDVI). In all schemes, higher noise levels had a slightly negative impact on the fit accuracy (average RMSE and average MAD).

For all noise levels at both sites, the number of segments in the generalised trends decreased almost linearly when the generalisation-threshold (δ) increased from zero to 100% (results are not shown). However, the average number of segments was smaller for lower noise levels. The results demonstrated DBEST's efficiency in generalising trends, which carries the associated benefit of reducing the cost of further computations based on the generalised trend.

The results of evaluating different generalisation schemes using the simulated NDVI time series are not shown, since they are broadly similar to the results obtained using the simulated NDVI trends.

4.1.2. Detecting and characterizing trend change

Fig. 5 illustrates one sample of the simulated NDVI time series for each site, with the decomposed components (trend, seasonal and remainder), and three major changes detected by DBEST. The changes were statistically significant (α=0.05). Note that the simulated trends without noise (in green) which represent the plant life phases were

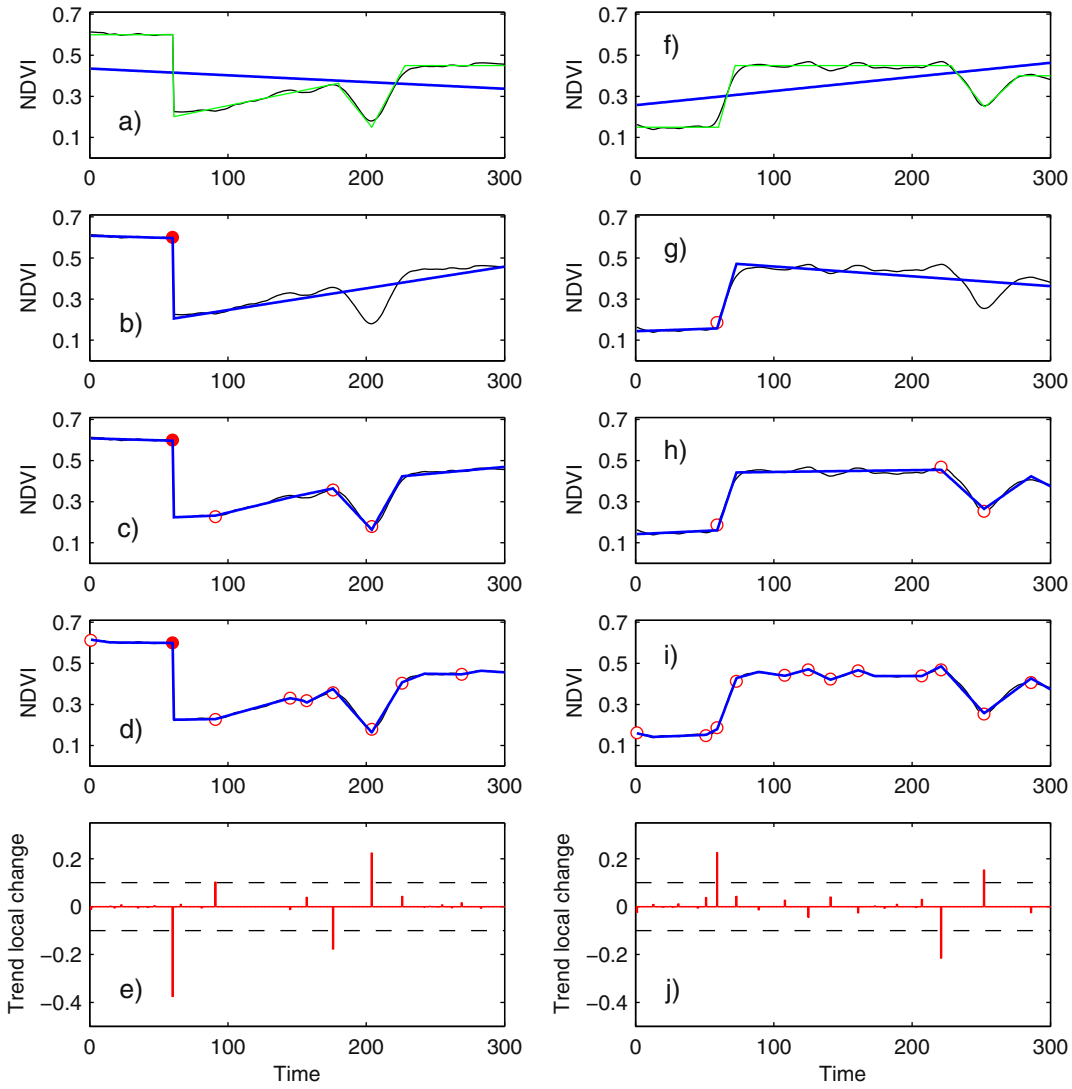

Fig. 4. Left-hand column shows results of DBEST's trend generalisation algorithm for site 1, right-hand column for site 2. The plots show simulated trends of NDVI at $\sigma = 0$ (in green) and $\sigma = 0.04$ noise levels (in black), and the model fits (in blue) obtained by applying DBEST. The DBEST generalisation schemes used were as follows: (a, f) no-breakpoint, (b, g) major breakpoint, (c, h) change magnitude-threshold $\beta = 0.1$ NDVI, and (d, i) all-inclusive breakpoints schemes at site 1 and site 2. Bar charts of the trend local change (i.e. h function) with 0.1 NDVI boundaries (dashed lines) are also shown (e, j). Every turning point has a trend local change value on the bar-charts, so the higher the absolute trend change value, the higher the probability that the point is a major breakpoint. Rings in red denote detected breakpoints in each scheme; solid rings represent level-shifts (abrupt breakpoint), and empty rings represent non-abrupt breakpoints. The simulated NDVI trend, shown for each site, is one of the 500 iterations of the simulation using $\sigma = 0.04$ noise level. The trend local change values are in NDVI units.

used as basis for simulating the NDVI time series (Fig. 5a, e). The seasonal components contained seasonal growing cycles with varying amplitudes for healthy and disturbed seasons (Fig. 5c, g). The remainders (unexplained variance) were, in general, small (Fig. 5d, h). Since DBEST's main purpose is currently for detecting changes in the trend component, we keep the focus on the trend results.

We chose to detect three breakpoints at each site. As intended, DBEST captured the processes and/or events that caused the greatest changes to the NDVI trend at both sites (the two defoliations and one post disturbance recovery process at site 1; and the plantation, drought, and post drought recovery at site 2); and, approximated them with segments sloped properly (either abrupt as the first sudden defoliation at site 1 or slower as the other changes) and matched adequately in shape to the trend (Fig. 5b, f).

Break dates were determined by DBEST with an average RMSE range of 2.7–3.9 time steps at site 1 (Fig. 6a), and 2.5–4.9 time steps at site 2 (the less accurate estimates correspond to higher noise levels) (Fig. 6b). Higher noise levels had a slightly more negative impact on the average RMSE of the time at site 2 than site 1. The more robust results obtained for site 1 might be because of the larger underlying changes, and hence higher signal-to-noise ratios, compared to site 2 (Table 2). Another reason may be due to the type of the first change (abrupt level-shift) at site 1; DBEST showed slightly more accurate results for level-shift changes. The change magnitude accuracies were ~0.04–0.05 NDVI units at both sites (Fig. 6a, b). The average RMSE for obtained fits was 0.02–0.06 (NDVI units) at both sites, and the average MAR was 0.09–0.18 (NDVI units) at both sites. The fit accuracy measures (average RMSE and average MAR) decreased slightly as the noise level increased (Fig. 6a, b).

The overall accuracy for the change type determination was in the range 95–100% at site 1 and 87–100% at site 2, with lower accuracies at higher noise levels (Fig. 7). Again, DBEST showed slightly higher accuracy at higher noise levels for site 1 (with level-shift change) than site 2. The results of evaluating DBEST for detecting and characterizing the three changes using the simulated trends are not shown, since they are generally similar to the results obtained using the simulated NDVI time series.

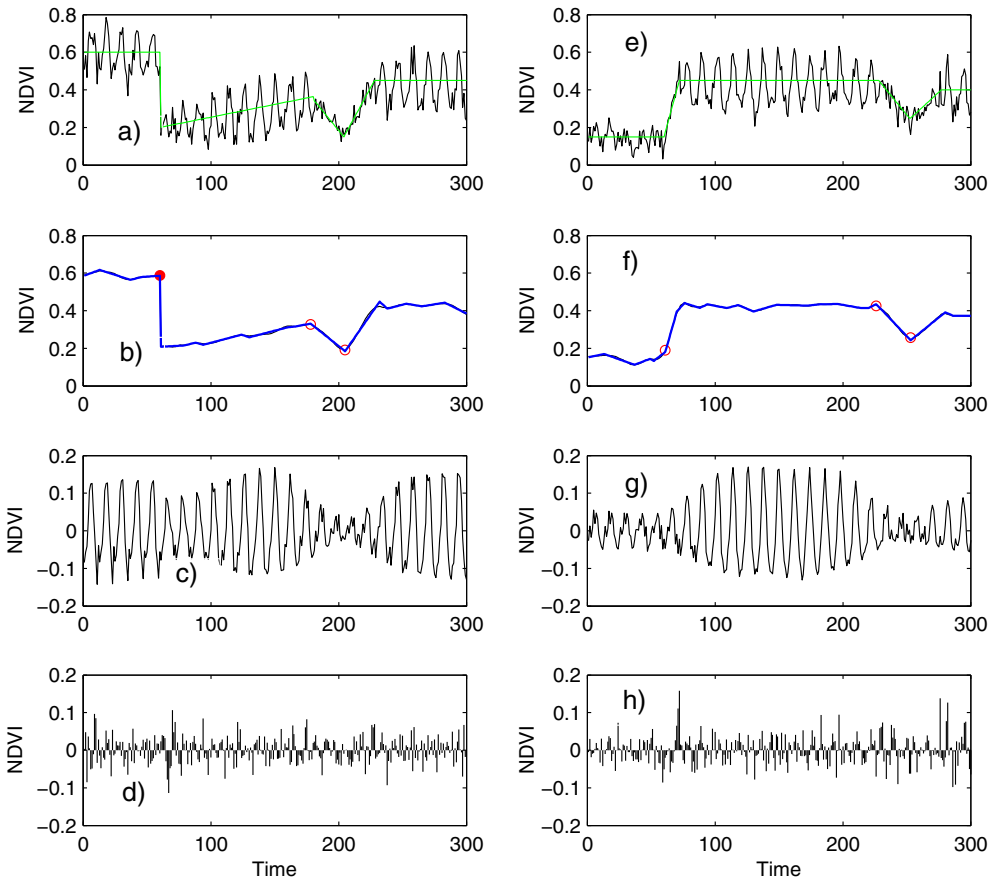

Fig. 5. The left- and right-hand columns show the results of DBEST's change detection algorithm for site 1 and site 2 respectively. (a,e) — The simulated trends of NDVI at noise level $\sigma = 0$ (in green) with simulated NDVI time series at noise level $\sigma = 0.04$ (in black). (b,f) — Decomposed trend components using STL (in black) and fit results obtained by DBEST (in blue) for detecting three major changes in the trend at each site. (c, g) — The decomposed seasonal, and (d, h) — remainder components. The detected breakpoints, corresponded to the three major changes, are marked with rings; a solid ring represents abrupt (level-shift) change, and an empty ring represents non-abrupt change. The simulated NDVI time series, shown for each site, is one of the 500 iterations of the simulation using a noise level of $\sigma = 0.04$.

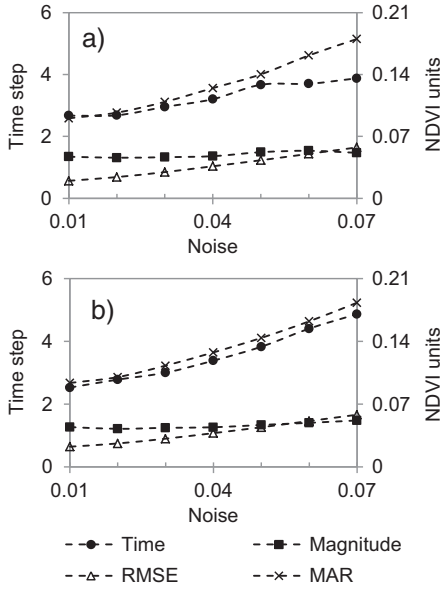

Fig. 6. Average RMSE for the time (left vertical axis), and magnitude (right vertical axis) estimates of three major changes within the simulated NDVI time series at (a) site 1 and (b) site 2. The average RMSE for the obtained fits, and average maximum-absolute-remainder (MAR) are also shown. The three major changes were detected by DBEST's change detection algorithm using the simulated NDVI time series with 500 iterations for different noise levels ($\sigma = 0.01$–0.07).

4.2. User-defined thresholds and parameters

The level-shift thresholds (θ_1 and θ_2) and the duration-threshold (ϕ), used for the detection of level-shift changes, are set based on the user's criteria for detecting abrupt changes of interest in the studied variable. We considered a change in NDVI as a significant level-shift when it was greater than 0.1 NDVI units ($\theta_1 = 0.1$), when it shifted the level (mean) of NDVI by more than 0.2 ($\theta_2 = 0.2$), the shift was valid for at least two years ($\phi = 24$ monthly time steps), and finally when it was

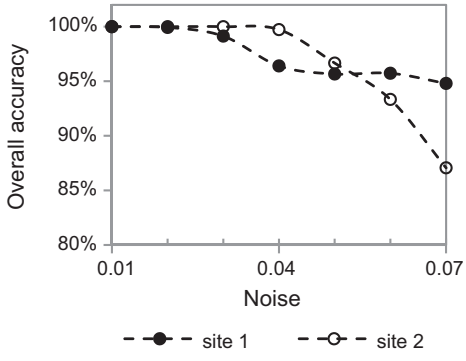

Fig. 7. The overall accuracy of the change type (abrupt, non-abrupt) of three major changes detected by DBEST's change detection algorithm using the simulated NDVI time series with 500 iterations for different noise levels ($\sigma = 0.01$–0.07) at site 1 and site 2.

a statistically significant breakpoint. The user can change these threshold parameters to suit their own data set, but it is recommended that $\theta_1 \leq \theta_2$. Setting larger level-shift thresholds (θ_1 and θ_2) and a longer duration-threshold (ϕ) is safer than smaller level-shift thresholds and shorter duration-threshold setting, because the latter may lead to false detection of NDVI seasonal variations. The abovementioned flexible setting for level-shift detection does not limit DBEST to detection of solely instantaneous discontinuities; an abrupt change (i.e. level-shift) detected by DBEST can be instantaneous or non-instantaneous dependent on whether the time duration of the change is one time-step or longer, respectively.

To check the effect on DBEST of the distance-threshold parameter (ε), we computed the accuracies (in terms of average RMSE over the 500 iterations) of the estimated times and magnitudes of the three major changes in the simulated trend time series at the two sites using a noise setting of 0.04, using the values $\varepsilon =$ 0–0.2 NDVI units (Fig. 8). As seen, the accuracies improve (RMSEs decrease) as the distance-threshold increases from 0.01 to around 0.02–0.03 for time and 0.06–0.07 for magnitude, but decrease (RMSEs increase except for the magnitude at site 1) before levelling off at around 0.1 NDVI units (this should be the maximum influential value of distance-threshold). For ε values up to 0.1, the time estimates showed higher sensitivity to higher values of ε whereas the magnitude estimates showed higher sensitivity to lower ε values. We then recomputed the accuracies using the ε values derived by Eq. (7) in DBEST. The derived value for the distance-threshold, averaged over the 500 simulations, was about 0.06 at site 1 and 0.04 at site 2 (markers in red in Fig. 8); they were close to the ε values for which the lowest RMSEs were obtained. The computed accuracies (using ε values derived by Eq. (7)) were also comparable with the highest accuracies achieved. Similar results were obtained for noise levels of $\sigma = 0.01$ and $\sigma =$ 0.07. The computations of the ε value and corresponding accuracies were repeated for different values for the constant C in Eq. (7). The optimum C value (i.e. the value for which the differences between the accuracies were the lowest for each level of noise) was about 3. These computations demonstrate that Eq. (7) works satisfactorily. The user can choose to keep the default value of the distance-threshold derived by Eq. (7) in DBEST, or change it to a value between zero and maximum absolute difference between data values of the given time series.

The generalisation-threshold (δ) parameter can be freely set by the user (between 0 and 100%) and no specific default is set in DBEST for this parameter. The same goes for the change-magnitude-threshold (β) and the number of change of interest (m) parameters. The only exception is that the m parameter can be no larger than the total number of breakpoints (s) determined by BIC.

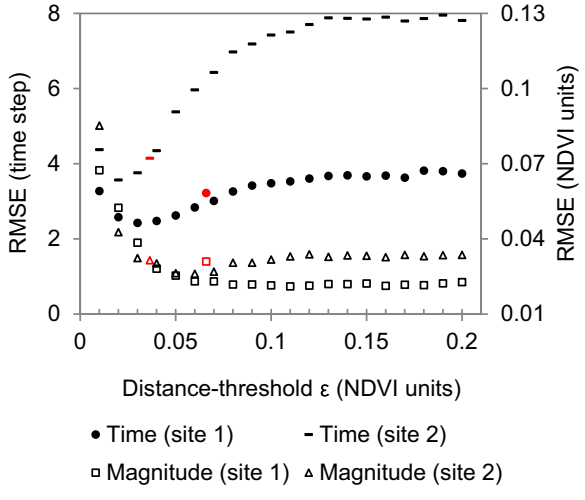

Fig. 8. Average RMSE for the estimated time (left vertical axis) and magnitude (right vertical axis) of three major changes found by DBEST's change detection algorithm for the simulated trend time series, as a function of the distance-threshold (ε) (on horizontal axis). The input data consisted of 500 iterations with the noise level $\sigma =$ 0.04 at site 1 and site 2. Markers in red denote the average RMSEs obtained for time and magnitude estimates using the distance-threshold, computed by Eq. (7), in DBEST.

4.3. Spatial application of the major change characterization

When detecting changes in Iraq, the major change detected in the NDVI trend was of the non-abrupt change type over the whole country. The exception was in small regions in the southeast close to the border with Iran, which showed abrupt (i.e. level-shift) positive changes (Fig. 9a). The non-abrupt changes were fairly small (≤0.1 NDVI units) over the most parts of the area (87% of unmasked area) (Fig. 9b), either positive (46%) mostly in west or negative (41%) scattered elsewhere, but mainly in east and northeast. Some scattered small areas with small positive changes were also found within the rest of the area that showed negative changes. Bigger positive changes (between 0.1 and 0.2 in NDVI units) were detected in north Iraq (5%). Surrounding them, there were also detected areas (5%) with considerable negative changes (between −0.1 and −0.3 in NDVI units) (Fig. 9b).

The small changes (<0.1 NDVI) started mainly in 1995–1998 and lasted for a 2–24 month time period (Fig. 9c, d). The northern greening started in 1984–1988 and lasted for 12–36 months in the most parts. The browning seen in the north and northeast, surrounding the greening, started in 1987 and lasted mainly for a 2–12 month period, or a few months longer. This gradual browning could reflect the impact of huge emigration in the 1990s (Sirkeci, 2005). The strong greening in NDVI (>0.2) in the southeast close to the border with Iran (~1%), identified mainly as abrupt change (e.g. Fig. 9e) with a one-month duration or some few non-abrupt changes with longer durations (2–12 months) in neighbouring pixels, started in 2004–2005. All the major changes detected in NDVI were significant (α=0.05). The exception was in small scattered areas (~2%) with insignificant small (≤0.1 NDVI units) changes.

Our findings are generally in agreement with Hamdan et al. (2010) and Richardson et al. (2005); they reported a similar vegetation response in the south-eastern wetlands due to a replenishment and releasing of the Tigris and Euphrates River waters, and another extensive re-flooding in March 2004. The uncontrolled replenishment was implemented by local residents following the fall of Iraq's government in spring 2003, after a long period of deliberate destruction and wetland reduction by Iraq's government during 1990.

Following this confirmation, the regions with large NDVI changes may be targeted for further investigation, using ground data and finer resolution imagery. This is because there is always a possibility that some of the observed changes are due to factors other than change in vegetation. One possible source of error is that the GIMMS data set comprises data from six NOAA satellites: orbital drift of the satellites can distort the calculated NDVI, leading to false change detection.

4.4. Strengths and weaknesses of DBEST

DBEST is independent of the sensors that provide its input data. It can be applied to image series of any vegetation index, on a pixel-by-pixel basis, regardless of spatial and temporal resolution. DBEST's main outcomes (i.e. fitted segments) are useful for estimating the characteristics (time and magnitude) of abrupt and long-duration changes, for comparing long-term vegetation trends for different locations, and for comparing NDVI trends observed by different sensors. DBEST also has the potential to be employed more generally in multiple breakpoint detection, an active research field in statistics and econometrics (Reeves et al., 2007).

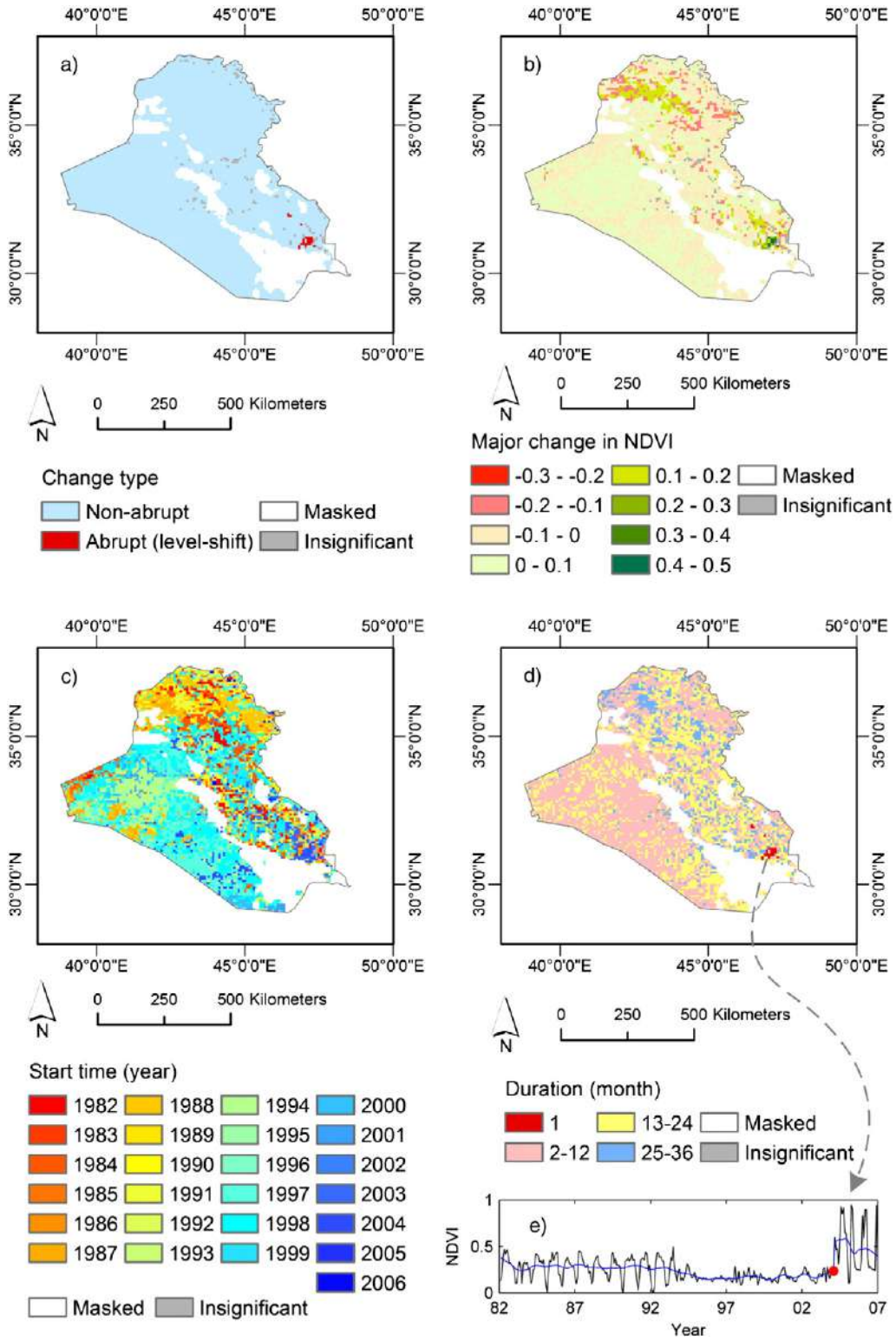

Fig. 9. DBEST estimated characteristics of major change in trend of monthly GIMMS–NDVI, 1982–2006, Iraq. The characteristics include: (a) change type, (b) change in NDVI, (c) start time, and (d) duration. Subfigure (e) represents the NDVI time series (in black), computed trend (in blue), and the detected major change (level-shift) (red dot) for one sample pixel located in areas with abrupt change, as shown in subfigure (a), in southeast Iraq. Pixels with at least 75% valid data points (i.e., quality flag value zero) through the whole time span were analysed. Areas with a mean monthly NDVI < 0.1 were masked out. All computations were done based on the monthly trend, but for an easier representation the start time is represented in years.

However, there are two issues that require attention when applying DBEST. First, if input data to DBEST are deseasonalized time series, they should not include high frequency variation (local roughness): this criterion should be fulfilled because the trend, by definition, is a slowly changing, smooth component of a time series. Otherwise, it is recommended to remove the variations using standard smoothing filters.

Second, DBEST works on a pixel-by-pixel basis and considers each pixel's time series data as an isolated entity in the change detection procedure; the spatial behaviour of adjacent areas is not used to improve robustness of detection. The use of neighbouring pixels for robust trend estimation of remotely sensed time series data is demonstrated by Bolin et al. (2009).

DBEST is based on simple mathematical principles and as a software product runs correspondingly quickly (0.07–0.18 s average execution time to detect change using the simulated trend and NDVI time series, respectively). The slower speed in the latter case is mainly due to the time it takes for the STL to run.

The DBEST user can control the trend generalisation and change detection processes through a number of thresholds and parameters, including the generalisation percentage, the number of required changes and the change threshold parameters.

With regard to a comparison of pros and cons of BFAST and DBEST, BFAST benefits from the advantage of computing confidence interval for the change time estimate and detecting changes in the seasonal component, whereas DBEST is not yet developed for these tasks. DBEST only indicates if the detected changes are statistically significant or not. However, DBEST's algorithm benefits from improved flexibility and manageability; it has a generalisation algorithm for flexible extraction of information at any scale, from details to main trend features. Moreover, discontinuities in DBEST can be modelled as one time-step or longer shifts depending on whether they are instantaneous are not, respectively. The latter advance in methodology may improve the change estimate precision. However, it should be mentioned that a quantitative-based comparison of the methods is not straightforward because they are based on different concepts and on fundamentally different strategies rendering the comparative analysis difficult.

4.5. Future work

Further development of DBEST will include testing it using remotely sensed satellite imagery for different types of changes in a wide range of ecosystems. It would be interesting, from the standpoint of ecosystem management or warning system development, to investigate DBEST's sensitivity to different trend types (e.g. monotonic, interrupted, reversal trends, or trends initiated from a steady-state equilibrium), and change magnitudes. In addition, validation under noisier conditions (caused by clouds or snow), using images with finer temporal and spatial resolution, still remains necessary. Moreover, the current version of DBEST can only analyse regularly spaced data, in the temporal sense. Development and testing DBEST using irregularly spaced data series (including time series with missing data), as well as the detection of changes in the seasonal component of a time series, are next in line as future works.

It is our intention to integrate DBEST with TIMESAT (Jönsson and Eklundh, 2004) to enable non-linear trend detection to be applied to remotely sensed seasonality parameters.

5. Conclusions

Our recently developed time series algorithm DBEST can properly simplify the temporal trend of a vegetation index time series into a flexible number of straight-line segments. The user can control the generalisation process in DBEST by setting the generalisation-threshold. This capability allows the user to capture a wide range of features, from fine details through to main features of the trend, over large areas. DBEST also detects trend changes, determines their type (abrupt or non-abrupt), and estimates their timing, magnitude, number, and direction. The user can set the number or magnitude of major changes for detection.

The validation and evaluation results demonstrated DBEST's efficiency in terms of providing results which were accurate, robust and cost effective. We used DBEST to detect major changes in NDVI trends over Iraq during 1982-2006; they were small (<0.1 NDVI) over large areas and in both forms of gradual greening (46% of the land area, mostly in the western regions) and browning (41% of the land area). Northern and eastern Iraq showed larger changes as both greening (0.1–0.2 NDVI) and browning (−0.1 to −0.3 NDVI) (5% each). Abrupt changes (>0.2 NDVI) leading to significant shifts in the level of NDVI were detected in relatively small areas (~1%) in southeast. All the major changes detected were statistically significant except small changes (<0.1 NDVI) in scattered areas (~2%).

DBEST can be applied to time series of different remotely sensed vegetation indices for detecting and mapping both short and long duration events at different spatial and temporal scales. DBEST also has the potential to be applied in other disciplines dealing with multiple trend change detection, such as hydrology and climatology. The DBEST code is developed in MATLAB with a graphical user interface (GUI) front end. The code together with test data and user instructions are available by contacting the authors.

Acknowledgements

The authors are very grateful for the constructive feedback from two anonymous reviewers that helped improve the quality of the article. The authors are also grateful to Dr. Jan Verbesselt for clarifications regarding the functionality of BFAST. They also thank Dr. Ford Cropley for proofreading the final version. The authors appreciate the NASA GIMMS group for providing the NDVI data at no cost. This work was supported by the European Union funding program, Erasmus Mundus "External Cooperation Window" (EMECW lot8). Partial support was also provided by the LUCID (www.lucid.lu.se), a Linnaeus centre of excellence at Lund University, from grant number 259-2008-1718 funded by the Swedish Research Council Formas (www.formas.se).

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22.

Bolin, D., Lindström, J., Eklundh, L., & Lindgren, F. (2009). Fast estimation of spatially dependent temporal vegetation trends using Gaussian Markov Random Fields. *Computational Statistics and Data Analysis*, *53*, 2885–2896.

Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, *62*, 241–252.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, *6*, 3–73.

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression — An approach to regression-analysis by local fitting. *Journal of the American Statistical Association*, *83*, 596–610.

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation. *Remote Sensing of Environment*, *114*, 2911–2924.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*, 35–46.

de Jong, R., Verbesselt, J., Schaepman, M. E., & de Bruin, S. (2012). Trend changes in global greening and browning: Contribution of short-term trends to longer-term change. *Global Change Biology*, *18*, 642–655.

de Jong, R., Verbesselt, J., Zeileis, A., & Schaepman, M. (2013). Shifts in global vegetation activity trends. *Remote Sensing*, *5*(3), 1117–1133, http://dx.doi.org/10.3390/rs5031117.

Draxler, R. R., Gillette, D. A., Kirkpatrick, J. S., & Heller, J. (2001). Estimating PM10 air concentrations from dust storms in Iraq, Kuwait and Saudi Arabia. *Atmospheric Environment*, *35*, 4315–4330.

Eklundh, L., & Olsson, L. (2003). Vegetation index trends for the African Sahel 1982–1999. *Geophysical Research Letters*, *30*(8), 1430, http://dx.doi.org/10.1029/2002GL016772.

Fan, L., Gao, Y., Brueck, H., & Bernhofer, C. (2009). Investigating the relationship between NDVI and LAI in semi-arid grassland in Inner Mongolia using in-situ measurements. *Theoretical and Applied Climatology*, *95*, 151–156.

Fensholt, R., Langanke, T., Rasmussen, K., Reenberg, A., Prince, S. D., Tucker, C., et al. (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.

Fensholt, R., Rasmussen, K., Nielsen, T. T., & Mbow, C. (2009). Evaluation of earth observation based long term vegetation trends — Intercomparing NDVI time series trend analysis consistency of Sahel from AVHRR GIMMS, Terra MODIS and SPOT VGT data. *Remote Sensing of Environment*, *113*, 1886–1898.

Fuller, D. O. (1998). Trends in NDVI time series and their relation to rangeland and crop production in Senegal, 1987–1993. *International Journal of Remote Sensing*, *19*, 2013–2018.

Hamdan, M., Asada, T., Hassan, F., Warner, B., Douabul, A., Al-Hilli, M., et al. (2010). Vegetation response to re-flooding in the Mesopotamian Wetlands, Southern Iraq. *Wetlands*, *30*, 177–188.

Heumann, B. W., Seaquist, J. W., Eklundh, L., & Jönsson, P. (2007). AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982–2005. *Remote Sensing of Environment*, *108*, 385–392.

Higgins, P. A. T., Mastrandrea, M.D., & Schneider, S. H. (2002). Dynamics of climate and ecosystem coupling: Abrupt changes and multiple equilibria. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *357*, 647–655.

Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, *7*, 1417–1434.

Jacquin, A., Sheeren, D., & Lacombe, J. P. (2010). Vegetation cover degradation assessment in Madagascar savanna based on trend analysis of MODIS NDVI time series. *International Journal of Applied Earth Observation and Geoinformation*, *12S*, S3–S10.

Jamali, S., Seaquist, J., Eklundh, L., & Ardö, J. (2014). Automated mapping of vegetation trends with polynomials using NDVI imagery over the Sahel. *Remote Sensing of Environment*, *141*, 79–89.

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

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

Kennedy, R. E., Cohen, W. B., & Schroeder, T. A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, *110*, 370–386.

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sensing of Environment*, *114*, 2897–2910.

Keogh, E., Chu, S., Hart, D., & Pazzani, M. (2004). *Segmenting time series: A survey and novel approach. Data mining in time series databases*, *57*. (pp. 1–22), 1–22.

Lambin, E. F., & Ehrlich, D. (1997). Land-cover changes in sub-Saharan Africa (1982–1991): Application of a change index based on remotely sensed surface temperature and vegetation indices at a continental scale. *Remote Sensing of Environment*, *61*, 181–200.

Lambin, E. F., Geist, H. J., & Lepers, E. (2003). Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources*, *28*, 205–241.

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.

Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S.C., Tucker, C. J., et al. (2003). Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, *300*, 1560–1563.

Peng, J., Liu, Z., Liu, Y., Wu, J., & Han, Y. (2012). Trend analysis of vegetation dynamics in Qinghai–Tibet Plateau using Hurst Exponent. *Ecological Indicators*, *14*, 28–39.

Prentice, I. C., Heimann, M., & Sitch, S. (2000). The carbon balance of the terrestrial biosphere: Ecosystem models and atmospheric observations. *Ecological Applications*, *10*, 1553–1573.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1990). *Numerical recipes*. Cambridge University Press.

Reeves, J., Chen, J., Wang, X. L., Lund, R., & Lu, Q. (2007). A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology*, *46*, 900–915.

Richardson, C. J., Reiss, P., Hussain, N. A., Alwash, A. J., & Pool, D. J. (2005). The restoration potential of the Mesopotamian marshes of Iraq. *Science*, *307*, 1307–1311.

Roderick, M., Smith, R., & Cridland, S. (1996). The precision of the NDVI derived from AVHRR observations. *Remote Sensing of Environment*, *56*, 57–65.

Rouse, J. W., Hass, R. H., Schell, J. A., & Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. *Third ERTS Symposium, NASA SP-351 I* (pp. 309–317).

Rousseeuw, P., & Leroy, A. (1987). *Robust regression and outlier detection*. John Wiley and Sons (329 pp.).

Schnepf, R. (2004). Iraq agriculture and food supply: Background and issues. *CRS Report for Congress, Order Code RL32093*. Congressional Research Service, The Library of Congress.

Schwarz, Gideon E. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.

Sirkeci, I. (2005). War in Iraq: Environment of insecurity and international migration. *International Migration*, *43*, 197–214.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., et al. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*, 161–185.

Tucker, C. J., Pinzon, J. E., & Brown, M. E. (2004). Global inventory modeling and mapping studies. *Global land cover facility*. College Park, Maryland: University of Maryland.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., et al. (2005). An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing*, *26*, 4485–4498.

UNEP (2001). The Mesopotamian marshlands: Demise of an ecosystem. *Division of early warning and assessment*. Nairobi, Kenya: United Nations Environment Programme.

Venables, W. N., & Ripley, B.D. (2002). *Modern applied statistics with S* (4th edition). Springer-Verlag, 156–163.

Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010a). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment*, *114*, 106–115.

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

Verbesselt, J., Robinson, A., Stone, C., & Culvenor, D. (2009). Forecasting tree mortality using change metrics derived from MODIS satellite data. *Forest Ecology and Management*, *258*, 1166–1173.

Wessels, K. J., van den Bergh, F., & Scholes, R. J. (2012). Limits to detectability of land degradation by trend analysis of vegetation index data. *Remote Sensing of Environment*, *125*, 10–22.

Zeileis, A., Kleiber, C., Kramer, W., & Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, *44*, 109–123.

Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2002). strucchange: An R package for testing for structural change in linear regression models. *Journal of Statistical Software*, *7*(2), 1–38.