Ecological Indicators 26 (2013) 49–60

Contents lists available at SciVerse ScienceDirect

Ecological Indicators

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

Original article

Addressing the complexity in non-linear evolution of vegetation phenological change with time-series of remote sensing images

E. Ivits*, M. Cherlet, S. Sommer, W. Mehl

EC Joint Research Centre, Institute for Environment and Sustainability, Land Resource Management Unit, Via E. Fermi 1, 21027 Ispra, Italy

ARTICLE INFO

Article history:
Received 23 February 2012
Received in revised form 9 October 2012
Accepted 16 October 2012

Keywords:
Non-linear phenology fluctuations
Stationarity
Linear regression

ABSTRACT

Earth observation based monitoring of change in vegetation phenology and productivity is an important and widely used approach to quantify degradation of ecosystems due to climatic or human influences. Most satellite based studies apply linear or polynomial regression methods for trend detections. In this paper it is argued that natural systems hardly react to human or natural influences in a linear or a polynomial manner. At shorter time-scales of few decades natural systems fluctuate to a certain extent in a non-systematic manner without necessarily changing equilibrium. Finding a systematic model that describes this behavior on large spatial scales is certainly a difficult challenge. Furthermore, the manner vegetation phenology reacts to climate and to socio-economic changes is also dependent on the land cover type and on the bioclimatic region. In addition to this, traditional parametric methods require the fulfillment of several statistical criteria. In case these criteria are violated confidence intervals and significance tests of the models may be biased, even misleading. This paper proposes an alternative approach termed the *Steadiness* to traditional trend analysis methods. Steadiness combines the direction or tendency of the change and the net change of the time-series over a selected time period. It is a non-parametric approach which can be used without violation of statistical criteria, it can be applied on short time-series as well and results are not dependent on the significance test or on thresholds. To demonstrate differences, a time-series of satellite derived Season Length images for 24 years is analyzed for the entire European continent using linear regression and the Steadiness approach. Spatial and temporal change patterns and sensitivity to pre-processing algorithms are compared between the two methods. We show that linear regression limits the possibilities of assessing fluctuating ecosystem changes whereas the non-parametric Steadiness index more consistently confirms the fluctuating phenological change patterns.

1. Introduction

To address complex global challenges, environmental managers need up-to-date information on the status and trends of land degradation, their causes and effects, and need to be offered routes for possible solutions. In particular, various aspects of vegetation dynamics, reflecting land cover/use transitions that can lead to land degradation, need to be considered in spatio-temporal context. Because of the large areal coverage and continuous temporal sampling, remotely sensed data provides a synoptic picture of vegetation dynamics in space and time and thus have a great potential for monitoring vegetation and ecosystem change from regional to global scales (Myneni et al., 1997). Using satellite based time-series imagery vegetation phenological metrics can provide a quantitative basis to monitor such changes. This has been one of the central features in global change research as it provides researchers with an independent measure on how ecosystems respond to external impacts, be it human induced or climate change (Fensholt et al., 2012; Linderholm, 2006; Parmesan, 2006; White et al., 2009).

Many studies applied time-series of remote sensing images to investigate the timing of recurring biological events and their connection to climate change. However, some of these studies present different magnitudes of phenological changes and report contradicting findings related to earlier/later or shorter/longer seasons (among others Hogda et al., 2001; Zhou et al., 2001; Stöckli and Vidale, 2004; Julien and Sobrino, 2009; Zhu et al., 2011; Jeong et al., 2011a,b). Contradicting results were also reported from studies using similar time series of satellite derived productivity data in terms of which areas show negative trends (Bai et al., 2008; Hein and de Ridder, 2006; Hellden and Tottrup, 2008; Prince et al., 2007; Wessels, 2009). Naturally, some of the contradictions arise from the different length of the time series analyzed and from the different satellite sensors used (deBeurs and Henebry, 2005; Jeong et al., 2011a), from differences in methods applied for detecting trends

* Corresponding author. Tel.: +39 0332 785315.
E-mail address: eva.ivits-wasser@ext.jrc.ec.europa.eu (E. Ivits).

1470-160X/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ecolind.2012.10.012

(Fensholt and Rasmussen, 2011) as well as from differences of the spatial scale of the studies.

Most of these studies rely on linear regression modeling of a monotonous trend in natural systems, which is disputable as it may not account well for the fluctuation inherent to natural systems. The present study refers to the fact that natural systems hardly change linearly or react to human or natural influences in a linear manner. Therefore, finding a systematic model that describes spatially distributed areas according to their different system stages in terms of their likeliness of undergoing ecosystem change or rather fluctuating in a relative equilibrium is a difficult challenge. Moreover, the linear regression as a quantitative parametric model that relies on four principal assumptions which must be met. If any of these assumptions are violated then confidence intervals and significance tests of the linear regression model may be misrepresentative: (1) linearity of the relationship between dependent and independent variables; (2) independence of the errors, i.e. no serial autocorrelation of the residuals; (3) stationarity (constant variance) of the errors versus the predictor; (4) normality of the error distribution. Some authors tried to get around these constraints by, e.g. adding a non-linear term in the regression equation (deBeurs and Henebry, 2004, 2005). However, it was also suggested by the same authors that depending on the latitude and land cover the non-linear function needs further adjustments (deBeurs and Henebry, 2005). Non-parametric trend measures as, e.g. the Theil–Sen's (Theil, 1950; Sen, 1968) and the Mann–Kendall tests (Mann, 1945; Kendall, 1975) are robust against non-normality of the distribution in the time-series (Yue and Pilon, 2004). However, non-parametric tests also require the setting of user-defined thresholds introducing subjectivity similarly to significance assessment based on the *t*-test. Zhou et al. (2001) calculated the persistency index from piecewise linear trends of the phenological time-series (denoted as $t(i)$), over increasing number of years but with the same starting year. In case $t(i+1) > 80\%$ of $t(i)$ a score of 1 was given, otherwise the score was zero. The sum of these scores is the index of persistent NDVI phenological change and is categorized into high and low persistency. However, the Persistence Index is also subject to setting thresholds that might jeopardize the consistency and repeatability of global scale studies. Local to regional scale studies may apply user-defined thresholds, continental to global scale assessments however cannot be based on subjective decisions because trends are land cover and bio-climate dependent (deBeurs and Henebry, 2005).

The aim of the present study is to address the methodological and conceptual constraints faced when using linear methods to assess the fluctuating nature of phenological changes. Using only one aspect of ecosystem change, for practical reasons we limit the study to time series analysis of the remote sensing derived vegetation growing season length (SL) over the European continent for the years 1982–2005 derived from NDVI of the GIMMS dataset. Tests and comparisons are made between a non-parametric linear trend analysis method and an alternative method termed "Steadiness", which combines the trend tendency and the net change of the time series. For the latter method, the slope of the linear trend (tendency) and the net change values are used as qualitative classifiers of the phenological fluctuation dynamics. Relying on a convergence of evidence rather than on significance values these classifiers are combined into a classification scheme that expresses the apparent long term direction into which the ecosystem appears to move in the given time window. The resulting classes represent the levels of Steadiness of change dynamics of the phenological metric. The different Steadiness classes characterize ecosystems ranging from strong negative to strong positive dynamics and also assign classes where the system can be assumed to fluctuate within the boundary condition of a steady equilibrium. In this study we argue that a not-significant test of the linear trend should not be taken indicative of a non-changing ecosystem state and that a simple measure like the here proposed Steadiness index might be better adapted to reflect fluctuating phenological dynamism of ecosystems.

To test and evaluate the two different approaches comparisons are made between the temporal profile and the spatial distribution of pixels under the different Steadiness classes and the temporal profile and spatial distribution of pixels with significant linear regression trends. The temporal profiles and spatial distribution of those pixels where the trend was not significant are also addressed and compared, respectively. In order to test the robustness of the Steadiness approach in relation to time series data normalization and data noise, we also considered two major pre-processing methods on the season length time-series frequently used in remote sensing change detection studies: *Z*-score normalization and moving average smoothing of the time-series. Subsequently, the Steadiness classes and the slope and significance statistics of linear trends were calculated from the unprocessed and from the pre-processed SL time-series. The manner phenology reacts to climate and to socio-economic changes is strongly dependent on the land cover type, land use and also on the bioclimatic region. Therefore, all comparisons were performed in a stratified way within the Northern, Central and Southern European Environmental Zones to account for the natural differences of phenological dynamics in major bioclimatic regions.

2. Materials and methods

2.1. Study area and data sets

The study area includes the entire European continent including Turkey and the Southern Mediterranean countries of Northern Africa and the Middle East extending between 10.8 degree West and 40.5 degree East and between 71.4 degree North and 27.8 degree South (Fig. 1). The bio-climatic zones covering the test region provided by the Environmental Classification of the world of Metzger et al. (2011) were used for the stratified analysis and comparison of the derived SL linear regression trends and Steadiness results, as shown in Fig. 1.

The Season Length (SL) time-series was derived from the Global Inventory Modeling and Mapping Studies (GIMMS) NOAA-AVHRR 1982–2005 NDVI 10 days composite data set covering the entire Earth at 8 km × 8 km pixel sampling (Tucker et al., 2005). It is considered and described as an NDVI dataset that has been corrected for calibration, view geometry, volcanic aerosols, and other effects not related to vegetation change (Tucker et al., 2005). Full detail can be found at the website of the Global Land Cover Facility (www.landcover.org) of the University of Maryland. The yearly values of SL were derived for each image pixel of the study area using the algorithms of the "phenolo" package for the derivation of vegetation phenology and productivity metrics as outlined by Ivits et al. (2012). Using this method, the GIMMS NDVI data was overlaid with smoothed forward and backward lagged curves derived from the original time-series by means of a Moving Average (MA) algorithm (Fig. 2). The "phenolo" package is designed to account for the time series dynamism of each pixel individually and adopts an objective approach to compute the lag, expressed in days:

$$L = \frac{\sum_{i=1}^{N}(365 - 2STD_i)}{N} \tag{1}$$

where L is the lag (in days), N is the number of years, 365 is the number of days in the year and STD_i is the standard deviation from the barycentre of the area under the yearly NDVI curve (expressed in days). Thus, at the individual pixel level, the lag applied for shifting

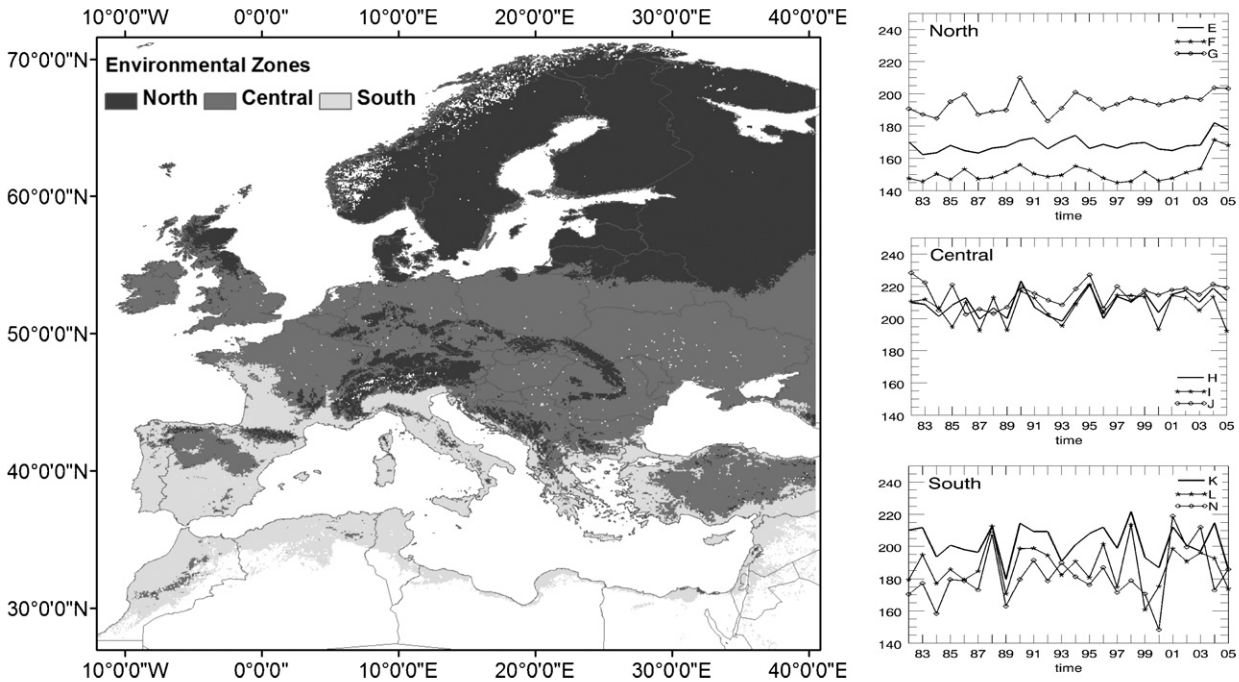


Fig. 1. Extent of study area and the average Season Length time series from 1982 to 2005 over the Northern, Central and Southern European Environmental Zones within the study area (*Y* axis: Season Length in days). The ecosystems are divided into the zones North, Central and South using the boundaries of the Environmental Classification (Metzger et al., 2011).

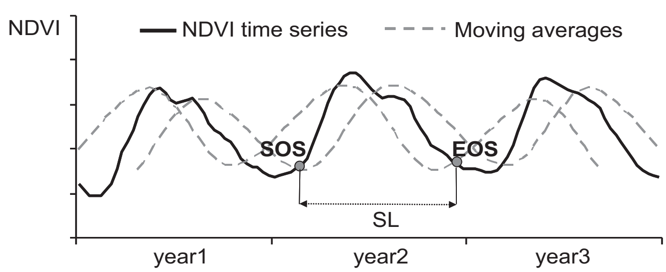


Fig. 2. Example for the calculation of Start of Season (SOS) and End of Season (EOS) points on the NDVI curve and the derived season length (SL).

the MA curves is constant throughout the time-series as determined by the per pixel average length of the non-growing season. In this way the phenolo approach, unlike the method of Reed et al. (1994), is independent of user defined determination of the forward and backward lag shift. Thus, the method adjusts to the individual characteristics of the NDVI time series curve under each pixel when processing the time series imagery accordingly and better reflects temporal and bioclimatic conditions, land-cover and management. The cross points of the original NDVI time series and the MA curves were used to approximate the start (SOS) and the end (EOS) of the vegetative growing season days for each year (Fig. 2). For each year, the distance between the SOS and EOS points was defined as the season length (SL).

Two main pre-processing algorithms were carried out on the phenological time-series which are often used in remote sensing change detection studies. Firstly, a normalization was carried out on the time-series by converting the series into *Z*-scores. The *Z*-score normalization gives insight into how typical a given observation is with regards to the entire time-series and enhances the comparison of the trend magnitudes because larger values will have larger slopes, which is inevitable along the North–South climatic gradient in Europe. As a result, this method provides a robust and valid estimate of the temporal rather than spatial and geographical transect driven trends (Hellden and Tottrup, 2008). The *Z*-scores were calculated according to:

$$Z = \frac{X_t - \bar{X}}{s} \tag{2}$$

where Z is the resulting normalized time series value, X_t is the original time series value for year t, $\bar{X}$ is the series mean and s is the standard deviation of the series. The *Z*-score normalization reduces the influence of outliers and of pixels with very high variations that result in extreme trend values and thus enables the comparison of the trend magnitudes along large spatial scales. Secondly, a moving average smoothing was performed over the time-series to remove high frequency variability using a 3-years window. Moving averages are used in time series analysis to smooth out short-term fluctuations (e.g. in Jeong et al., 2011b) thus highlighting longer-term trends or cycles. It has to be noted however, that the SL estimate is an already processed output and on shorter time-series a moving average filter might remove valid information rather than reducing additional noise. The filter was applied nevertheless to demonstrate the effect of pre-processing on the Steadiness index and on the significance test. Fig. 3 illustrates the effect of smoothing the SL time-series from 1982 to 2005 in the nine ecosystems (see Metzger et al., 2011) of Northern, Central and Southern Europe. When compared to Fig. 1 displaying the unsmoothed time series, the effect of smoothing in reducing the high fluctuations and thus the variance becomes obvious. However, in the Southern European ecosystems the residual variation in the time-series is still strongly indicating the difficulty of applying linear (or any polynomial functions for that matter) trend analysis methods over these areas.

2.2. Linear regression versus Steadiness – methodological considerations

When a linear model is fit to a time series that changes in a non-linear way, results tend to inaccurately over- or under estimate ecosystem change dynamics. This problem becomes more

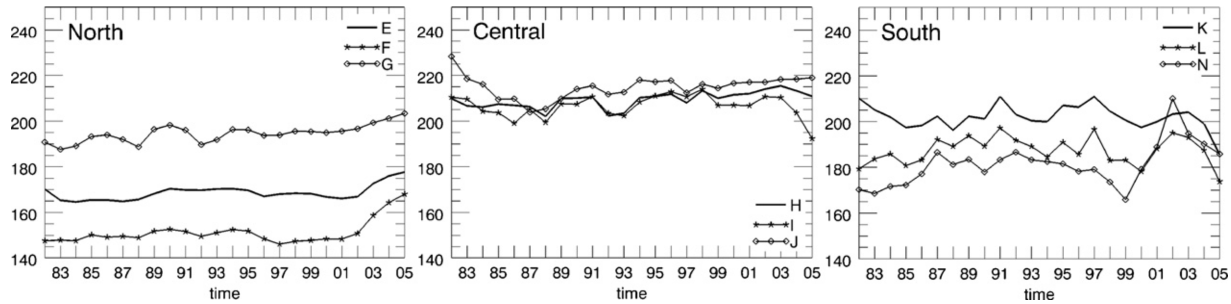


Fig. 3. Season Length time series over the Northern, Central and Southern European ecosystems (for the subdivisions see Metzger et al., 2011) after a moving average smoothing over the time series with a window size of 3.

serious when data are aggregated and results are extrapolated to larger spatial extents, as done with many remote sensing applications. Solutions to achieve linearity in the time-series include logarithmic transformations although the logarithm of, e.g. season length might be troublesome to understand. Adding a non-linear term in the regression equation in the form of a polynomial function is another solution. deBeurs and Henebry (2004) showed that in temperate regions quadratic models fit well for modeling the phenology of growing degree days over agricultural and herbaceous vegetation. The same authors however suggested that non-linear spherical models were fitting better for woody vegetation in Northern latitudes (due to a more rapid green-up and staying green for a relatively long period, deBeurs and Henebry, 2005). Although the value of these studies is unquestionable, important conclusions can be drawn: (1) phenological changes are rarely linear, (2) phenological changes are land cover dependent (deBeurs and Henebry, 2004), (3) phenological changes depend on the bioclimatic region and (4) traditional trend analysis methods must be adapted when applied on a continental or global level.

Another problem is that estimating the linear regression coefficients of time-series over natural systems becomes problematic with the inherent complexity of ecosystems. The fitted line will be strongly dependent on the sample's variance and consequently the residuals will also strongly depend on the sample's variance: the larger the variance the larger the residuals and consequently a simple straight line might not represent the real nature of change dynamics. Furthermore, the standard error of the linear trend is strongly influenced by the residuals and by the variance of the time-series. As a result, the *t*-statistics that depends on the standard error will be strongly influenced in such ecosystems and the significance test becomes unreliable. As an example, consider Fig. 1 showing SL time-series (1982–2005) averaged for the Northern, Central and Southern European ecosystems. The circumpolar and Boreal regions in Northern Europe exhibit trends with low variances and with a monotonous change pattern where linear regression in fact could be considered good approximation of ecosystem change. In the Central and especially in the Southern European regions, however, ecosystems exhibit increasingly non-linear change dynamics and increasing variance in their time-series, which will influence the calculation of the standard error. Over these complex areas linear trend models will be strongly affected by large variances and by the violation of the linearity assumption and consequently, the dynamisms of the system will not be well described by the linear model.

Regarding the violation of the independent errors criteria, it mainly points out that seasonality of the time-series has not been properly accounted for and makes transformation of the variables necessary. Transformations may help to stationarize the variables but, as mentioned above, they strongly reduce the interpretability of the results. Regarding the stationarity criteria of linear methods, it is now widely understood that natural systems do not fluctuate within an unchanging envelope of variability, which, in fact has long been compromised by human disturbances (Milly et al., 2008). Furthermore, non-stationarity often arises due to growth that represents the real nature of the time series and therefore transforming the data would distort the results. However, violations of stationarity usually results in confidence intervals that are too wide or too narrow and may have the effect of giving too much weight to small subsets of the data and bias significance results. The *t*-test, used to conduct hypothesis tests on the regression coefficients obtained in the linear regression, can only be carried out if the random error term is normally and independently distributed. Yet natural systems hardly ever are normally distributed and the transformation of the variables introduces difficulties in the interpretation of results. Non-normality can also be caused by large outliers, but in many cases these outliers are genuine members of the natural system as e.g. productivity drop in case of forest fires or droughts. These extreme values are just the values we are interested in and accordingly should not be removed. Non-parametric trend measures as, e.g. the Theil-Sen's (Theil, 1950; Sen, 1968) and the Mann–Kendall tests (Mann, 1945; Kendall, 1975) are robust against non-normality of the distribution in the time-series (Yue and Pilon, 2004) and therefore are now increasingly used in studies instead of the *t*-test to assess trend significance. However, if all above mentioned statistical criteria are met and in case the natural system changes in a linear manner with low fluctuation one still faces the problem of setting a threshold for the non-parametric significance test that will only present results that fall in a given confidence interval.

The proposed Steadiness approach intends to address the above listed problems. The method, being non-parametric, does not have to comply with the assumptions of linear regression regarding independence of the errors, stationarity and normality and can be applied on any type of data. There is no reliance on statistical significance as a criterion of trend relevance in order to avoid the necessity of setting thresholds. Avoiding significance tests also enables the method to be applied on short time series where the calculation of statistical significance would not provide meaningful information due to the limited number of observations. Instead, the Steadiness index is based on a convergence of evidences that the ecosystem changes dynamics. This is reached by combining two simple measures:

(1) The tendency of the change, expressed in the slope of the trend. The tendency is calculated by fitting a linear trend over the time-series using the formula:

$$Y = \beta_0 + \beta_1 X \quad (3)$$

where β_0 is the intercept, β_1 is the slope of the fitted line, X is time and Y is the SL time series. The slope of the linear trend expresses the dominant tendency, positive or negative, toward which the system moves. In assessing this tendency no

Table 1
Summary of the four Steadiness classes.

Steadiness classes
Steadiness 1: *negative* slope and *negative* change. Represents pixels under strong and negative ecosystems dynamics, with a probability of changing equilibrium
Steadiness 2: *negative* slope and *positive* change. Represents pixels under moderate negative ecosystems dynamics but likely to remain in current equilibrium
Steadiness 3: *positive* slope and *negative* change. Represents pixels under moderate positive ecosystems dynamics but likely to remain in current equilibrium
Steadiness 4: *positive* slope and *positive* change. Represents pixels under strong and positive ecosystems dynamics with a probability of changing equilibrium

hypothesis test is performed but the raw values are used in a qualitative scheme independently of their significance.

(2) The net change, positive or negative, of the phenological metrics over the selected time period, supplying a second evidence of ecosystem dynamism. The net change is calculated by the Multi Temporal Image Differencing (MTID) method (Guo et al., 2008). MTID for the Season Length time-series over the period 1982–2005 of the present study is calculated as:

$$\mathrm{MTID} = \sum_{i=1982}^{2004} (D_{2005} - D_i) \tag{4}$$

where D_i equals the digital number of the Season Length value in the corresponding year. Although these two measures develop alike over large areas, due to the non-linear, fluctuating nature of Season Length values of ecosystems, a pixel with positive slope might express negative net change, and vice versa as shown in Fig. 4.

The combination of the tendency and change of the time-series provides a convergence of evidence for monotonous and dominant, upwards or downwards, change of the system and results in the classes of the Steadiness index. There are four potential combinations of the negative or positive trend and of the negative or positive change (Table 1 and Fig. 4) that represent the levels of equilibrium or possibly changing equilibrium of the phenological metric. The Steadiness 1 class indicates strong negative dynamics of the observed time series with monotonous downwards trends and negative net change, indicating that the ecosystem is developing toward a changing equilibrium. Steadiness 2 class shows areas where the trends are negative but the net change is positive. Here the system fluctuates but does not show a clear tendency toward a new equilibrium characterized by, e.g. systematically shorter season length. Hence, these areas are more likely to remain within the current equilibrium and have a chance to maintain the steady state. Steadiness 3 class shows areas where the time series trends are positive but the net changes are negative. Here the system fluctuates but is likely to remain in current equilibrium and positive trends will not necessarily result in long term improvement of the ecosystems. Steadiness 4 class shows positive trends with positive net change of the observed time series with a probability of changing equilibrium. These four classes indicate an apparent direction into which the phenology dynamics of each pixel is evolving over time, upward or downward, confirmed or enhanced by the net change of the metric expressing the prevailing dynamic fluctuation of the system. Furthermore, this qualitative approach precludes the requirement for statistical significance inherent to the short time window of the time series and avoids the need to introduce possibly subjective or solely locally valid thresholds for interpretation of strengths of trends. Relying on convergence of evidence from the slope of the linear regression and from the change (MTID) indicator of the time-series, the Steadiness replaces the significance test and assesses the general nature of the change without excluding pixels from the further analysis.

Furthermore, we note that the last value of the time series, which is used as the reference in the MTID index, might have a strong effect on both the calculated MTID value and on the slope value in case the least year is an outlier. This is not an entirely unwanted effect for two reasons. (1) In our study we search for convergence of evidence for similarity in the change of ecosystem dynamisms therefore it is desirable to use change measures which are affected by the same way. Using a change indicator of another nature would disable finding convergence of evidence because of showing ecosystem changes in a different manner. (2) At such early stage of a time-series analysis it is not possible to state weather the first or the last year is a real outlier or rather a measurement error, e.g. a sensor artifact. If extreme values in the first or last years are not outliers, these values form an important component of the time series, as an extremely wet or extremely dry year or significant land use change might have happened just in one of those years. Therefore, smoothing out the effect of these values at this stage of the analysis could be a mistake that might lead to severe information loss. Following this reasoning, in this study Eq. (4) is used for assessing the change while being well aware of the possible effect of the last year on the value and we argue that this affect can only be accounted for with extra information on land use change and climatic effects at a later stage of the analysis.

2.3. *Comparative case studies between Steadiness and linear regression*

We calculated the linear trend and the MTID change measure over the SL time series for the entire study area. Significant pixels were defined with the Mann–Kendall test setting a threshold of $p < 0.1$. In order to evaluate the potential and the robustness of the Steadiness approach compared to linear regression, the following comparative case studies were performed under the Northern, Central and Southern European Environmental Zones:

1. The area proportion of pixels under negative and positive trends (raw and significant) and under the four different Steadiness classes were computed and compared.
2. Comparisons were made between the temporal profile of pixels under the Steadiness 1 class with the temporal profile of significant negative trends, and to the temporal profile of those pixels where the negative trend was not significant, respectively.
3. Comparisons were made between the temporal profiles of pixels under the Steadiness 2 class with the temporal profile of pixels under not significant but negative trends.
4. Comparisons were made between the temporal profiles of pixels under the Steadiness 3 class with the temporal profile of pixels with not significant positive trends.
5. Comparisons were made between the temporal profile of pixels under the Steadiness 4 class with the temporal profile of significant positive trends, and to the temporal profile of those pixels where the positive trend was not significant, respectively.
6. Subsequently, we calculated the Steadiness classes on the smoothed, on the Z-score normalized and on the smoothed and Z-score normalized SL time-series and compared the results with the Steadiness classes calculated from the un-processed time-series. Where Steadiness is steady, after pre-processing the time-series most pixels should stay in the Steadiness class that is classified from the un-processed time-series.
7. Finally, we calculated the significance and slope statistics of the linear trends on the un-processed and on the pre-processed (as under point 6) SL time-series and compared the results.

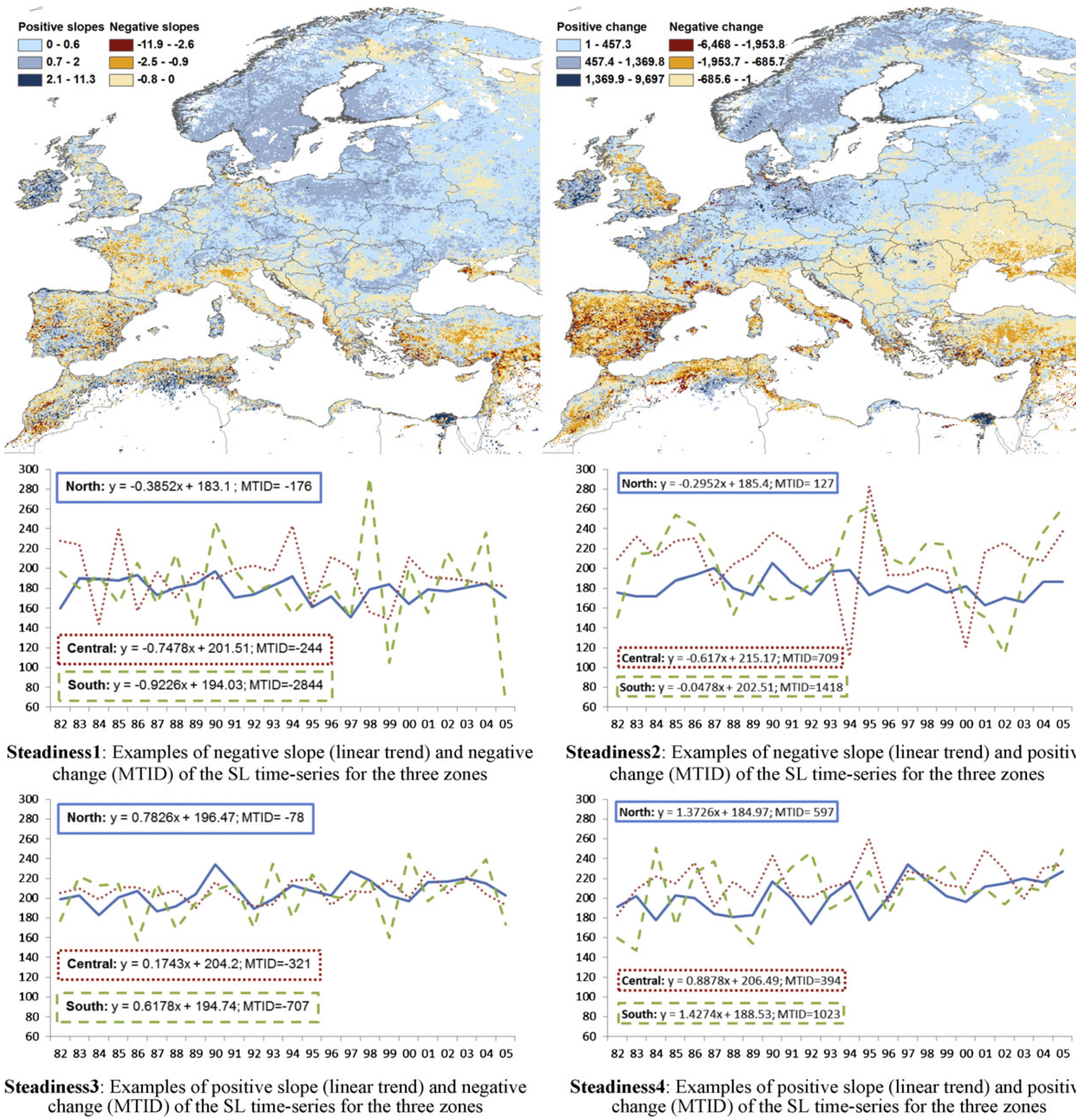


Fig. 4. Slope (upper left) and change (MTID, upper right) of the linear trend of the Season Length (SL) time-series for the years 1982–2005. Graphs: The SL time-series averaged within the zones North, Central and South as in Fig. 1 (*Y* axis: Season Length in days; *X* axis: calendar year). Boxes indicate the calculated linear trend equation respectively the MTID value of the presented SL time series.

3. Results

3.1. Spatial and temporal patterns of linear regression and Steadiness classes of the Season Length time series

The temporal profiles of the Northern, Central and Southern ecosystems indicate that significant negative trend pixels exhibit a strong downwards slope reacting a drop of SL values in 1989 and in 2000 (Fig. 5). The temporal profile of Steadiness 1 pixels, indicating a strong and downwards changing equilibrium of ecosystems, exhibit very similar pattern to pixels with significant negative trends and also react to the strong drops in SL values. The number of pixels with significant ($p < 0.1$) negative SL trend was less than 2% of the study area (Table 2), which drastically reduces the number of pixels to be used for further studies (Fig. 5). In comparison, the amount of pixels classified as Steadiness 1 class cover 18% of the study area with dominance in the Southern regions. These results indicate that while the temporal profiles of Steadiness 1 and of the significant negative trend pixels are very similar, the significance assessment excludes those areas from the further analysis where the system evolved toward changing equilibrium. Pixels where the fitted linear trend was negative, irrespective of their significance, exhibit a temporal profile similar to the Steadiness 1 class pixels but cover a much larger area (27.6% of the study area). It appears therefore that the Steadiness 1 class meaningfully refines the assessment of negative trends and is less restrictive than the assessment of significance of the linear trend model. Steadiness 2 groups pixels where the multi-annual residual net change of the variable value is positive despite negative linear slopes (Fig. 6). Over these pixels the system fluctuates therefore fitting a linear trend would not be representative for the overall resilience of such ecosystems. From the overall temporal profile it can be deducted that despite long term negative tendencies the dynamism of these ecosystems is not downwards with changing equilibrium but rather

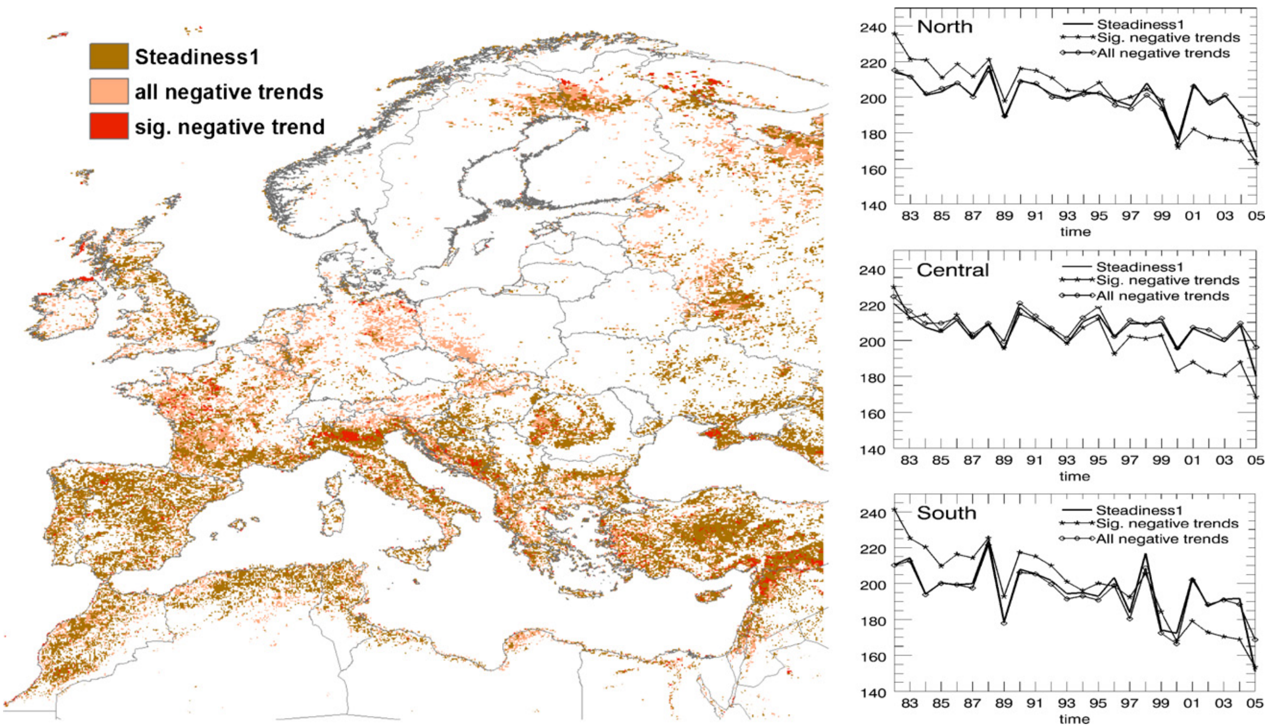


Fig. 5. Comparison of the spatial and temporal patterns of the Steadiness 1 class with the significance of the negative trends in the Northern, Central and Southern ecosystems. Note that the indicated areas overlap, where all negative values cover the largest and significant values cover the smallest areas. The *Y*-axes on the plots show the SL values in the ecosystems in days.

Table 2
Proportion (in %) of pixels within the four Steadiness classes and within the trend classes in the Northern, Central and Southern European ecosystems.

Ecosystems	Steadiness				Sig. negative trends, $p < 0.1$	All negative trends	Sig. positive trends, $p < 0.1$	All positive trends
	1	2	3	4				
% of North	8%	7%	10%	68%	0.7%	15.2%	28.7%	78.8%
% of Central	19%	9%	34%	38%	0.9%	27.9%	8.9%	72.0%
% of South	34%	13%	23%	30%	2.3%	46.9%	2.4%	52.8%
% of total	18.3%	9.2%	21.7%	48.1%	1.2%	27.6%	15.3%	69.9%

of a fluctuating nature. The temporal profile of Steadiness 2 pixels exhibit very similar pattern to pixels with not significant negative slopes without clear upwards or downwards changing pattern. We observed Steadiness 2 pixels over 9% of the study area mostly in the Central European and Southern regions (Table 2) whereas not significant negative trend pixels covered ca. 26% of the area (all negative - significant negative trends). The latter shows again that significance assessment would exclude a large area from further analyses whereas the Steadiness index assigns a value to this fluctuating ecosystem behavior.

Steadiness 3 groups pixels where the multiannual residual net change of the variable value is negative despite positive linear slopes (Fig. 7). Over these pixels the system fluctuates and fitting a linear trend would not represent the mere fluctuation around the equilibrium state of these ecosystems. The temporal profile of Steadiness 3 pixels exhibit very similar pattern to pixels with not significant positive slopes without clear upwards or downwards changing pattern. We observed Steadiness 3 pixels over ca. 22% of the study area mostly in the Central European and Southern regions whereas not significant positive trend pixels covered ca. 55% of the area (Table 2) excluding large territories from the analysis. The temporal profile of Steadiness 4 pixels, indicating a strong and upwards changing equilibrium of ecosystems, exhibit very similar pattern to pixels with significant positive trends (Fig. 8). The number of pixels with significant ($p < 0.1$) positive SL trend was only 15% of the study area (Table 2 and Fig. 5). In comparison, the amount of pixels classified as Steadiness 4 class covered 48% of the area. Pixels where the fitted linear trend was positive irrespective of their significance, exhibit an upwards temporal profile similar to pixels in the Steadiness 4 class but cover ca. 70% of the study area showing that Steadiness 4 meaningfully refines the assessment of positive trends. These results indicate that while the temporal profiles of the Steadiness 4 class and of the significant positive trend pixels are very similar, the significance assessment drastically reduces the number of pixels to be used for further studies and excludes those areas from the further analysis where the system evolved toward changing equilibrium.

3.2. Effects of data pre-processing on the Steadiness classes and on the significance and slope of the linear trend

Fig. 9 demonstrates the spatial distribution of the four Steadiness classes before and after applying various pre-processing algorithms on the SL time-series. All four images show almost identical results with respect to the spatial distribution of the four classes. Table 3 lists the proportion of pixels in the study area belonging to each of the Steadiness classes. The proportion of pixels classified in the different Steadiness classes remains similar independently

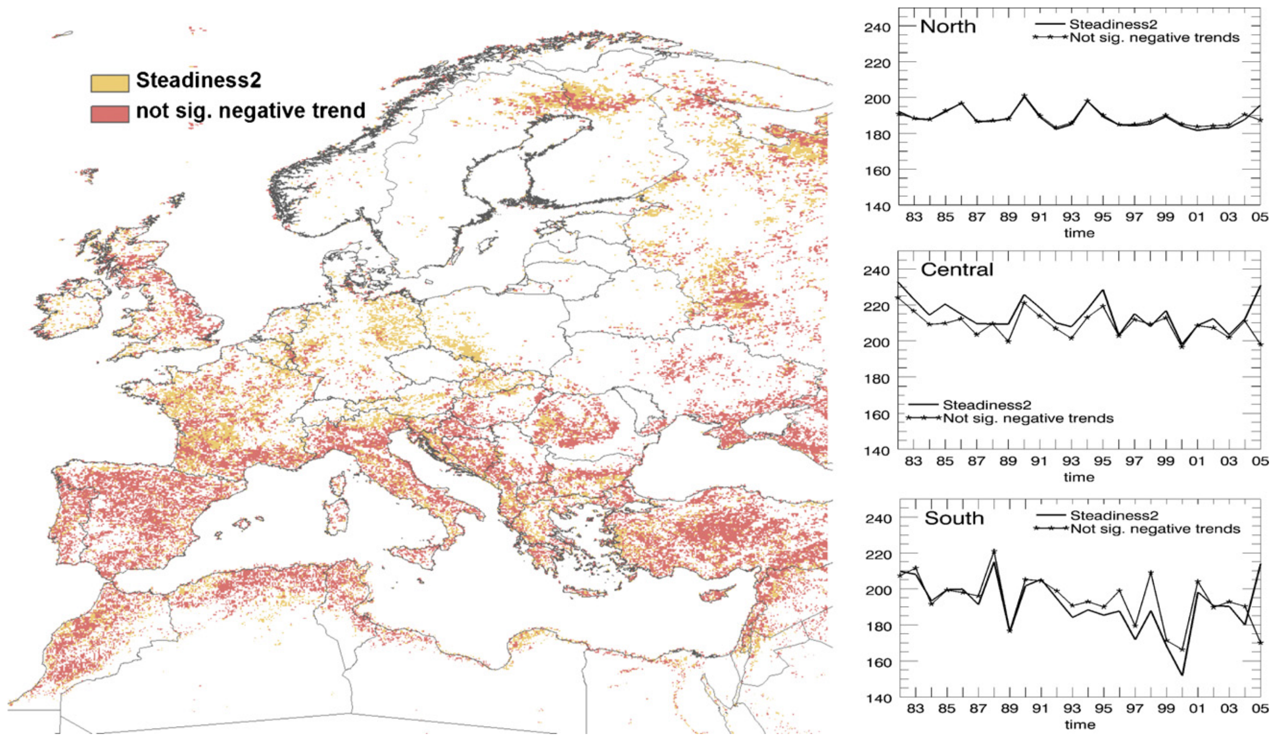


Fig. 6. Comparison of the spatial and temporal patterns of the Steadiness 2 class with the significance of the negative trends in the Northern, Central and Southern ecosystems. The *Y*-axes on the plots show the SL values in the ecosystems in days.

of pre-processing the SL time series and of the pre-processing method applied. In the Central European Region most pixels belong to Steadiness classes 3 and 4 whereas in the Southern zones the Steadiness classes 1 and 4 have the largest and most similar proportion of pixels. Moreover, the Northern European ecosystems are mostly characterized by the large, spatially continuous Steadiness 4 class whereas the Southern European regions are mostly fragmented belonging to one of the four Steadiness classes (Fig. 9).

Fig. 10 demonstrates the differences in spatial distribution of significant ($p < 0.1$) linear trends of the SL time-series calculated from the raw data and on the time-series after the pre-processing

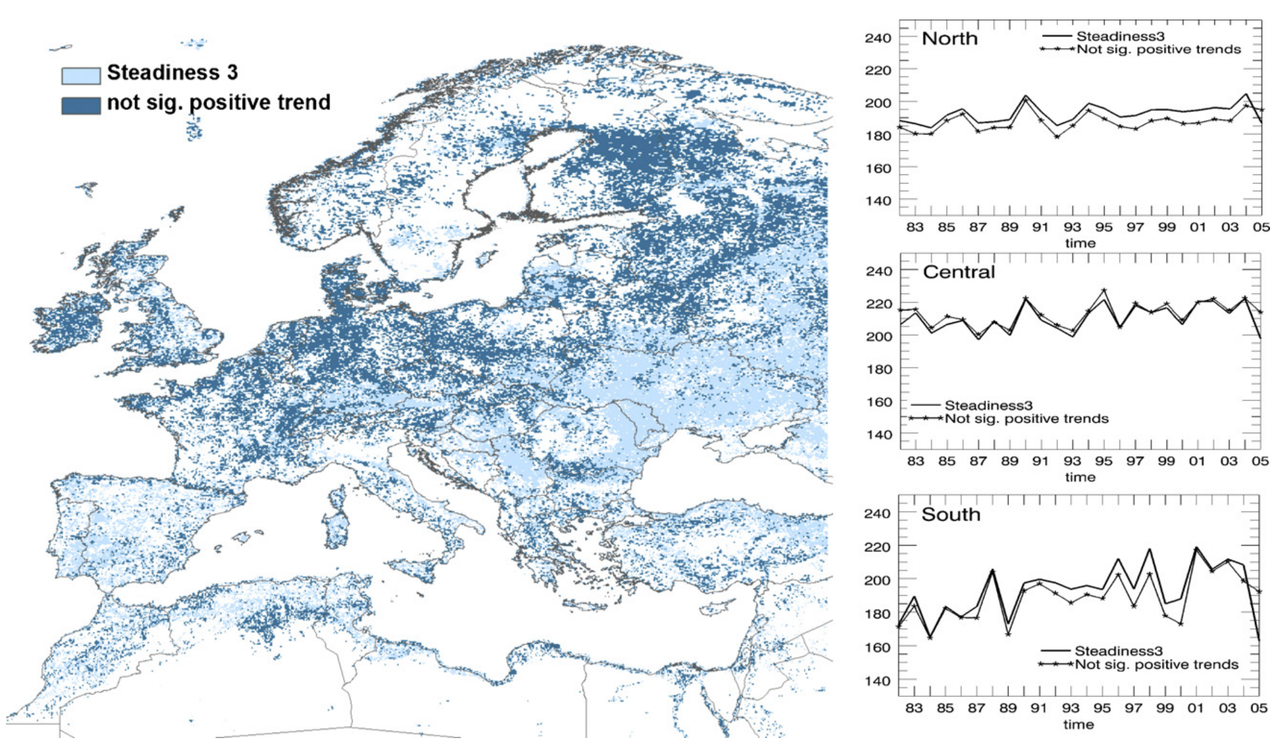


Fig. 7. Comparison of the spatial and temporal patterns of the Steadiness 3 class with the significance of the positive trends in the Northern, Central and Southern ecosystems. The *Y*-axes on the plots show the average SL values in the ecosystems in days.

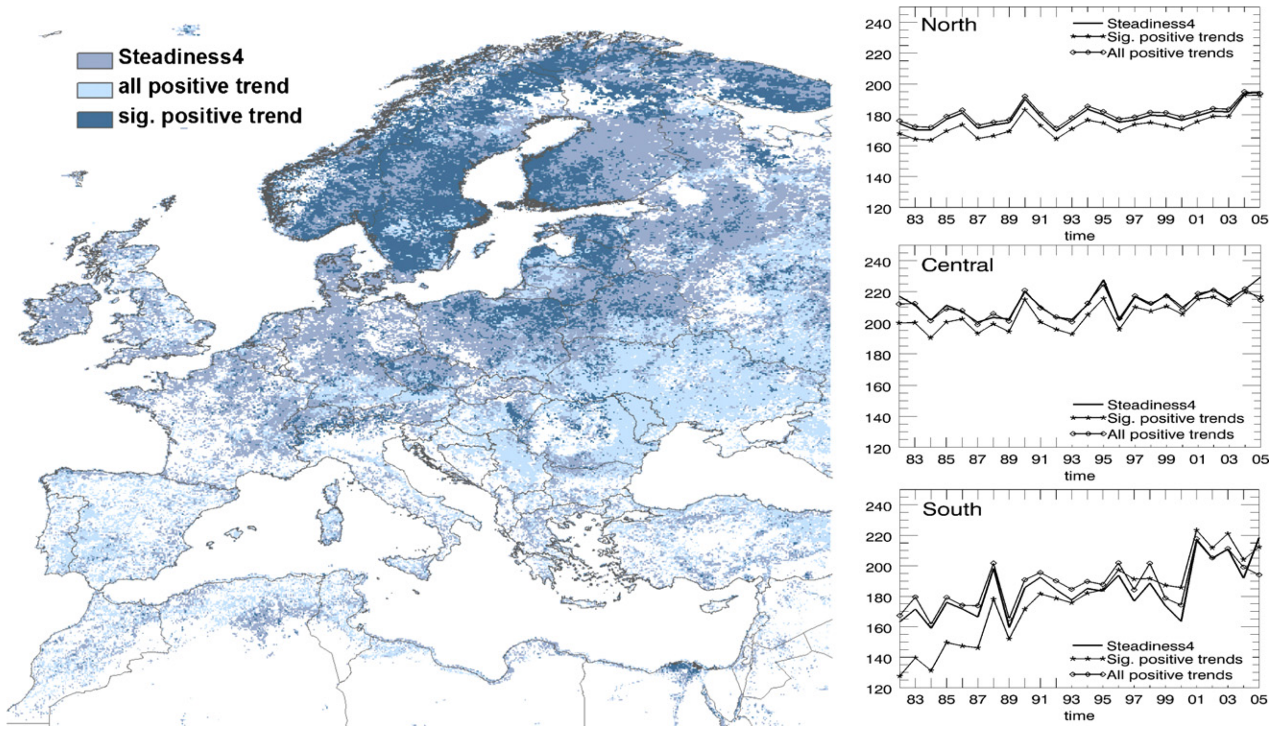


Fig. 8. Comparison of the spatial and temporal patterns of the Steadiness 4 class with the significance of the positive trends in the Northern, Central and Southern ecosystems. The *Y*-axes on the plots show the average SL values in the ecosystems in days.

Table 3
Proportion (in %) of pixels within the four Steadiness classes before and after pre-processing the time-series in the Northern, Central and Southern European ecosystems.

	Raw time-series				Smoothed time-series				*Z*-score normalized time-series				*Z*-score normalized, smoothed time-series			
	ST1	ST2	ST3	ST4	ST1	ST2	ST3	ST4	ST1	ST2	ST3	ST4	ST1	ST2	ST3	ST4
North	8%	7%	10%	68%	9%	10%	10%	65%	8%	7%	11%	68%	9%	10%	10%	65%
Central	19%	9%	34%	38%	16%	9%	37%	37%	19%	9%	34%	38%	16%	9%	38%	37%
South	34%	13%	23%	30%	31%	14%	26%	28%	34%	13%	23%	30%	30%	14%	27%	29%

Table 4
Proportion (in %) of pixels under significant ($p < 0.1$) negative and positive trends before and after pre-processing the SL time-series in the Northern, Central and Southern ecosystems.

	Raw time-series		Smoothed time-series		*Z*-score normalized time-series		*Z*-score normalized, smoothed time-series	
	Sig. neg.	Sig. pos.	Sig. neg.	Sig. pos.	Sig. neg.	Sig. pos.	Sig. neg.	Sig. pos.
North	0.164%	4.925%	1.164%	7.940%	0.164%	4.925%	0.956%	7.486%
Central	0.183%	1.623%	1.299%	5.985%	0.183%	1.623%	1.003%	5.496%
South	0.311%	0.386%	1.812%	2.467%	0.311%	0.386%	1.457%	2.067%

runs. The spatial distribution of the significant trend pixels exhibit strong changes after applying the smoothing filter either without or after the *Z*-score normalization of the time series (Table 4). The *Z*-score normalization alone did not affect the proportion of pixels under negative or positive trends. Smoothing the *Z*-score normalized time-series decreases the proportion of negative trends compared to the effect of smoothing the raw time-series whereas the proportion of positive trends was more stable. The difference in trend proportion was more significant in the Central European and Southern regions. Here, large areas which did not exhibit significant trends in the raw time-series became significant after smoothing was applied. In the Southern regions negative trends increased from 0.3% to 1.8% of the area whereas positive trends showed even higher differences with an increase from 0.4% to 2.5%. As shown in Fig. 10, the different pre-processing algorithms have a strong influence on the strength of the linear trends as well. In particular the *Z*-score normalized time-series changes the distribution of positive and negative slopes along the North–South gradient of Europe largely. Over Northern Europe for example the increase of strong positive trends can be observed which in turn was only moderate in the raw time-series. These results show the enormous impact of simple pre-processing algorithms on the outcome of linear regression, which in turn may strongly change the findings of research studies concerning phenological changes of ecosystems.

4. Discussion

Significance assessment of the linear trend might perform well where the established statistical criteria correspond to situations characterized by gradual system change. We argue however, that

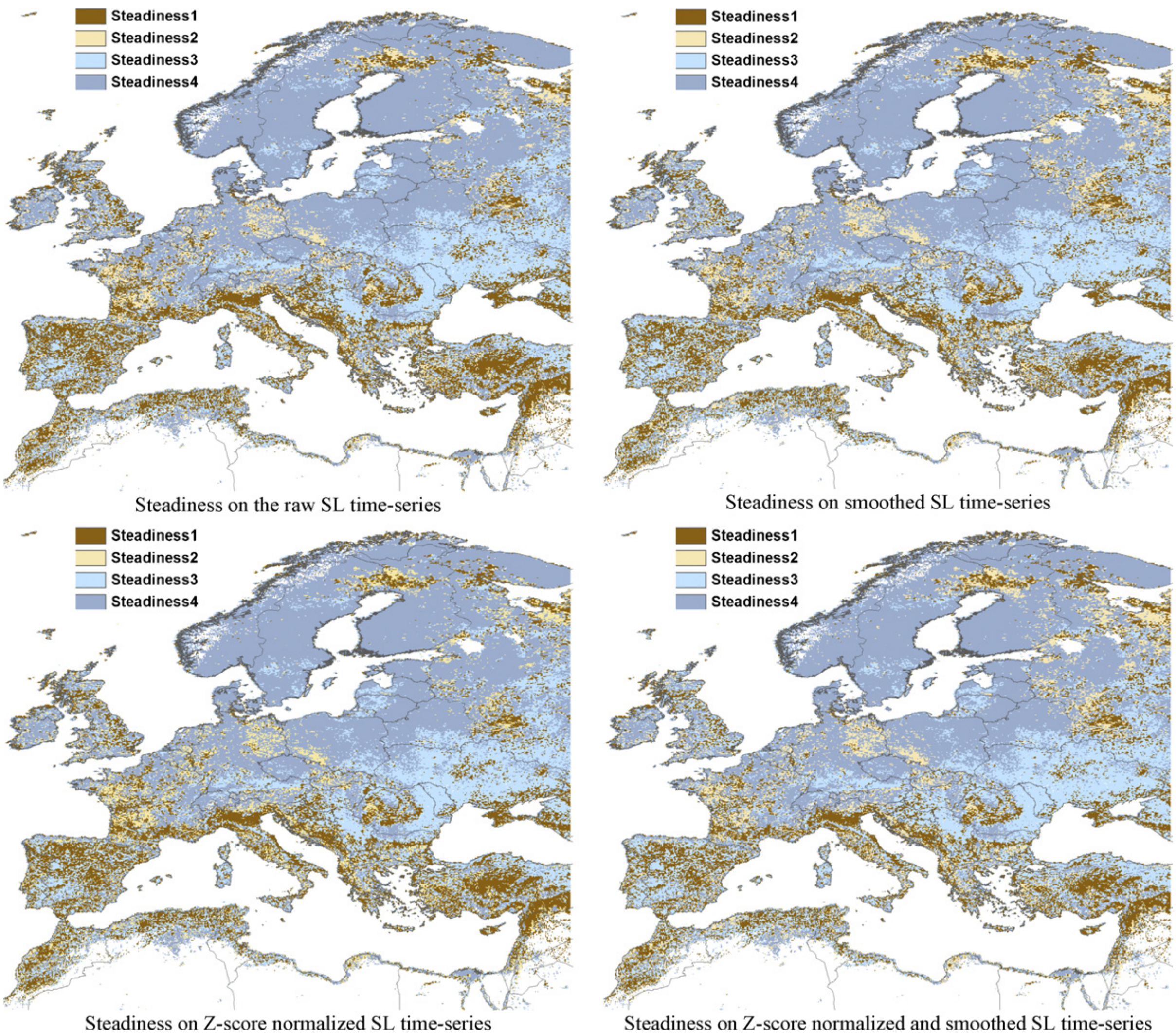
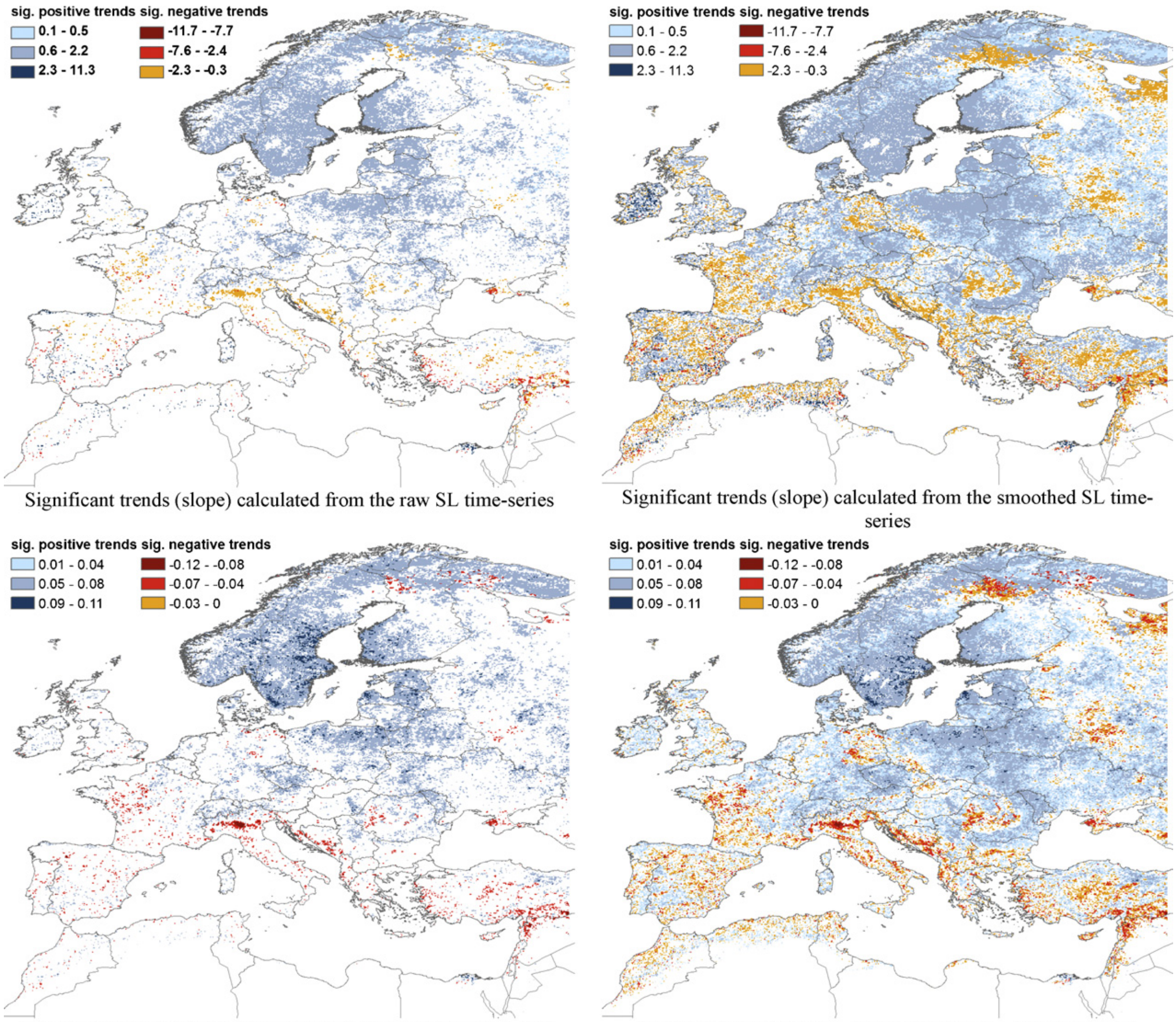


Fig. 9. Spatial distribution of the four Steadiness classes before and after pre-processing the Season Length time-series.

linear trend assessment does not optimally reflect the way vegetation phenology varies as a function of environmental change and human impact. At shorter time-scales of few decades natural systems appear to fluctuate in a steady but non-systematic manner even without necessarily changing equilibrium. As shown for the remote sensing derived Season Length index in this study, significance thresholds of linear trends exclude most of the study area from further assessment not confirming to the inherent, non-linear way ecosystems adapt to disruptions and their predisposition to resilience where and whenever possible. We argue therefore that a not significant trend assessment should not be considered indicative for non-changing ecosystem conditions. This study showed that after the assessment of significance less than 2% of the study area showed negative Season Length trends and most of the trends displayed a spatially scattered pattern. Such a decreased sample size and a spatially incoherent pattern would not supply a spatially and statistically representative sample for further comprehensive environmental monitoring and assessment. Furthermore, the size of the area unassigned depends on the significance level set which introduces an element of subjectivity to objective research studies. The study indicates that the Steadiness approach resulted in a more coherent spatial pattern of negative (and positive) dynamics. The Steadiness approach assigns the entire study area into classes that make distinction between probable equilibrium change and apparent variations that however stay within the limits of the natural fluctuation of the ecosystems. In this way the entire spatially continuous dataset is maintained and addressed in an objective manner. This provides a more complete basis to further analyze the characteristics of the ecosystem changes in different land cover and/or in different bioclimatic regions.

The postulated ecosystem fluctuation with varying tendencies toward short term change, which insinuates trends of improving or worsening stages, depends widely on the resilience toward land management and changing environmental factors. Instances where an area does not exhibit statistically significant trends does not necessarily mean that the ecosystems over this area is not undergoing change processes over the observed period, thus excluding pixels from further studies based on significant tests would impose a risk of misleading statements by giving-up the overall spatial context. Assigning instead all trends to relevant change irrespective of their significance increases the probability of overestimating negative and positive changes of the study area by including pixels where the system does not change but fluctuates in the limits of steady equilibrium. Results of this study suggest that the Steadiness approach is more effective in describing the inherent way ecosystems behave than traditional linear regression methods because it better reflects the fluctuating, non-stationary manner of constant ecosystem changes. Combining negative or positive trends with the net residual change provides convergence of evidence of changing equilibrium and indicates the apparent direction into which



Significant trends (slope) calculated from the raw SL time-series

Significant trends (slope) calculated from the smoothed SL time-series

Significant trends (slope) calculated from the normalized SL time series

Significant trends (slope) calculated from the normalized and smoothed SL time series

Fig. 10. Differences in spatial distribution of the significant ($p < 0.1$) slopes before and after pre-processing the SL time-series.

the dynamics of each pixel is evolving in the given time interval. The Steadiness approach is suggested as a simple, transparent and complementary way for interpretation of the change or fluctuating state of all ecosystems. Additionally, the Steadiness method being non-parametric, does not need to comply with the assumptions of linear regression, can be applied to any type of data and to short time-series where the calculation of statistical significance would not provide meaningful information due to the limited number of observations.

The analysis also showed that regardless of the pre-processing algorithm used, the spatial distribution of the Steadiness classes and the proportion of pixels belonging to each class remained stable. Applying simple smoothing algorithms on the Season Length time series largely changed the spatial distribution of the significant Season Length trend pixels. Large areas that did not exhibit significant trends in the raw time series became significant after smoothing was applied. When smoothing was applied on the *Z*-score normalized time series the proportion of significant negative and positive trend pixels changed in a drastic manner. The *Z*-score normalization also influenced the spatial distribution of strong negative and strong positive trends. These results may conform to statistical criteria, but the mere level of dissimilarity undermines further consistent analysis and interpretation of ecosystem changes. Non-parametric trend measures as, e.g. the Theil–Sen's (Theil, 1950; Sen, 1968) and the Mann–Kendall tests (Mann, 1945; Kendall, 1975) are robust against non-normality of the distribution and missing values in the time-series (Yue and Pilon, 2004) overcoming the need to conform strict statistical criteria. However, this study showed that pre-processing the time-series impacts the spatial distribution of non-parametric significant trends as well, influencing the findings of research studies concerning phenological changes of ecosystems. Additionally, non-parametric tests also require the setting of user-defined thresholds introducing subjectivity in the study of ecosystem change dynamism. Local to regional scale studies may apply user-defined thresholds, continental to global scale assessments however cannot be based on subjective decisions and on unique thresholds because trends are land cover and bio-climate dependent (deBeurs and Henebry, 2004).

This paper illustrates that the Steadiness approach, that combines the evidence of long term state fluctuation and residual change, is a valid complementary approach to address the fuzzy nature of ecosystem variability. Traditional trend analysis techniques cannot jointly address fluctuation and change and are unable to deliver this supplementary information, as they only classify ecosystems into hard boundaries of changing, either significantly or not. The results presented here illustrate the applicability and document the logic of the Steadiness approach for ecosystem change analysis, an important asset considering needs for environmental impact evaluation. The strength of the Steadiness method lays in the simplicity, robustness and straightforward applicability without setting user-defined subjective thresholds. It is proposed that for full thematic analysis the Steadiness should be calculated for a wider range of phenological and productivity variables derived from time-series images. More specific thematic analysis of environmental change and its cause–effect relationship of human and natural drives could then be performed as, e.g. proposed by Hill et al. (2008). In this context, Steadiness would have a strong potential to extend this type of approach to larger variable sets in a stratified way, e.g. combined analysis of selected variables e.g. season length, start/end of season, overall annual productivity and timing of vegetation maximum.

References

Bai, Z.G., Dent, D.L., Olsson, L., Schaepman, M.E., 2008. Proxy global assessment of land degradation. Soil Use Manage. 24, 223–234.

deBeurs, K.M., Henebry, G.M., 2004. Land surface phenology, climatic variation, and institutional change: analyzing agricultural land cover change in Kazakhstan. Remote Sens. Environ. 89, 497–509.

deBeurs, K.M., Henebry, G.M., 2005. Land surface phenology and temperature variation in the International Geosphere–Biosphere Program high latitudes. Global Change Biol. 11, 779–790.

Fensholt, R., Langanke, T., Rasmussen, K., Reenberg, A., Prince, S.D., Tucker, C.J., Scholes, R.J., Le, Q.G., Bondeau, A., Eastman, E., Epstein, H., Gaughan, A.E., Hellden, U., Mbow, C., Olsson, L., Paruelo, J., Schweitzer, C., Seaquist, J., Wessels, K., 2012. Greenness in semi-arid areas across the globe 1981–2007 – an Earth Observing Satellite based analysis of trends and drivers. Remote Sens. Environ. 121, 144–158.

Fensholt, R., Rasmussen, K., 2011. Analysis of trends in the Sahelian 'rain-use efficiency' using GIMMS NDVI, RFE and GPCP rainfall data. Remote Sens. Environ. 115 (2), 438–451.

Guo, W.Q., Yang, T.B., Dai, J.G., Shi, L., Lu, Z.Y., 2008. Vegetation cover changes and their relationship to climate variation in the source region of the Yellow River, China, 1990–2000. Int. J. Remote Sens. 29 (7), 2085–2103.

Hein, L., de Ridder, N., 2006. Desertification in the Sahel: a reinterpretation. Global Change Biol. 12 (5), 751–758.

Hellden, U., Tottrup, C., 2008. Regional desertification: a global synthesis. Global Planet. Change 64, 169–176.

Hill, J., Stellmes, M., Udelhoven, Th., Röder, A., Sommer, S., 2008. Mediterranean desertification and land degradation: mapping related land use change syndromes based on satellite observations. Global Planet. Change 64, 146–157.

Hogda, K.A., Karlsen, S.R., Solheim, I., 2001. Climatic change impact on growing season in Fennoscandia studied by a time series of NOAA AVHRR NDVI data. In: Proceedings of IGARSS, 9–13 July, Sydney, Australia, ISBN: 0-7803-7033-3.

Ivits, E., Cherlet, M., Tóth, G., Sommer, S., Mehl, W., Vogt, J., Micale, F., 2012. Combining satellite derived phenology with climate data for climate change impact assessment. Global Planet. Change 88–89, 85–97.

Jeong, S.J., Ho, C.H. Kim, Molly, E. Brown, 2011a. Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982–2008. Global Change Biol. 17, 2385–2399, doi:10.1111/j.1365-2486.2011.02397.x.

Julien, Y., Sobrino, J.A., 2009. Global land surface phenology trends from GIMMS database. Int. J. Remote Sens. 30 (13), 3495–3513.

Jeong, S.J., Ho, C.H., Park, T.W., Kim, J., Levis, S., 2011b. Impact of vegetation feedback on the temperature and its diurnal range over the Northern Hemisphere during summer in a 2× CO_2 climate. Clim. Dyn., doi:10.1007./s00382-010-0827-x.

Kendall, M.G., 1975. Rank Correlation Methods. Griffin, London, UK.

Linderholm, H.W., 2006. Growing season changes in the last century. Agric. Forest Meteorol. 137, 1–14.

Mann, H.B., 1945. Nonparametric tests against trend. Econometrica 13, 245–259.

Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G, 2011. Top-level tiers for Global Ecosystem Classification and Mapping Initiative (GEOSS Task ED-06-02). EBONE-WP3 Deliverable report D3.1. www.ebone.wur.nl

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Stationarity is dead: whither water management? Science, 319.

Myneni, R.B., Keeling, C.D., Tucker, C.J., Asrar, G., Nemani, R.R., 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. Nature 386, 698–702.

Parmesan, C., 2006. Ecological and evolutionary responses to recent climate change. Ann. Rev. Ecol. Evol. Syst. 37, 637–669.

Stöckli, R., Vidale, P.L., 2004. European plant phenology and climate as seen in a 20-year AVHRR land-surface parameter dataset. Int. J. Remote Sens. 25, 3303–3330.

Prince, S.D., Wessels, K.J., Tucker, C.J., Nicholson, S.E., 2007. Desertification in the Sahel: a reinterpretation of a reinterpretation. Global Change Biol. 13 (7), 1308–1313.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., Ohlen, D.O., 1994. Measuring phenological variability from satellite imagery. J. Veg. Sci. 5, 703–714.

Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. J. Am. Stat. Assoc. 63, 1379–1389.

Theil, H., 1950. A rank-invariant method of linear and polynomial regression analysis (parts 1–3). Nederl. Akad. Wetensch. Proc. 53, 386–392, 521–525, 1397–1412.

Tucker, C.J., Pinzon, J.E., Brown, M.E., Slayback, D., Pak, E.W., Mahoney, R., Vermote, E., Saleous, N.El., 2005. An extended AVHRR 8-km NDVI data set compatible with MODIS and SPOT vegetation NDVI data. Int. J. Remote Sens. 26 (20), 4485–5598.

Wessels, K.J., 2009. Comments on 'Proxy global assessment of land degradation' by Bai et al. (2008). Soil Use Manage. 25 (1), 91–92.

White, M.A., DeBeurs, K., Didan, K., Innouyes, D., Richardson, A.D., Jensen, O., O'Keefe, J., Zhang, G., Nemani, R.R., VanLeeuwen, W.J.D., Brown, J.F., DeWit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A.S., Kimball, J., Schwartz, M.D., Baldocchi, D.D., Lee, J.T., Lauenroth, W.K., 2009. Intercomparison, interpretation and assessment of spring phenology in North America estimated from remote sensing from 1982 to 2006. Global Change Biol. 15, 2335–2359.

Zhou, L., Tucker, C.J., Kaufmann, R.K., Slayback, D., Shabanov, N.V., Myneni, R.B., 2001. Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981–1999. J. Geophys. Res. 106 (D17), 20,069–20,083.

Zhu, W., Tian, H., Xu, X., Pan, Y., Chen, G., Lin, W., 2011. Extension of the growing season due to delayed autumn over mid and high latitudes in North America during 1982–2006. Global Ecol. Biogeogr., doi:10.1111/j.1466-8238.2011.00675.x.

Yue, S., Pilon, P., 2004. A comparison of the power of the *t* test, Mann–Kendall and bootstrap tests for trend detection. Hydrol. Sci. J. 49 (1).