

This discussion paper is/has been under review for the journal Biogeosciences (BG). Please refer to the corresponding final paper in BG if available.

Time series of vegetation indices and the modifiable temporal unit problem

R. de Jong^{1,2} and S. de Bruin¹

Received: 13 July 2011 - Accepted: 16 August 2011 - Published: 24 August 2011

Correspondence to: R. de Jong (r.dejong@wur.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Discussion Paper

Discussion Paper

Discussion Paper

Discussion Paper

BGD

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page

Abstract Introduction

Conclusions References

Tables Figures

I∢ ≯I

•

Back Close

Full Screen / Esc

Printer-friendly Version



¹Laboratory of Geo-Information and Remote Sensing, Wageningen University, The Netherlands

²ISRIC – World Soil Information, Wageningen, The Netherlands

Time series of vegetation indices (VI) derived from satellite imagery provide a consistent monitoring system for terrestrial plant systems. They enable detection and quantification of gradual changes within the time frame covered, which are of crucial importance in global change studies, for example. However, VI time series typically contain a strong seasonal signal which complicates change detection. Commonly, trends are quantified using linear regression methods, while the effect of serial autocorrelation is remediated by temporal aggregation over bins having a fixed width. Aggregating the data in this way produces temporal units which are modifiable. Analogous to the well-known Modifiable Area Unit Problem (MAUP), the way in which these temporal units are defined may influence the fitted model parameters and therefore the amount of change detected. This paper illustrates the effect of this Modifiable Temporal Unit Problem (MTUP) on a synthetic data set and a real VI data set. Large variation in detected changes was found for aggregation over bins that mismatched full lengths of vegetative cycles, which demonstrates that aperiodicity in the data may influence model results. Using 26 yr of VI data and aggregation over full-length periods, deviations in VI gains of less than 1% were found for annual periods, while deviations (with respect to seasonally adjusted data) increased up to 24% for aggregation windows of 5 yr. This demonstrates that temporal aggregation needs to be carried out with care in order to avoid spurious model results.

1 Introduction

Vegetation systems provide a quick and measurable response to many environmental changes at a wide range of spatial and temporal scales. The availability of historical time series from satellite observations with daily global coverage makes operational monitoring of vegetation condition a matter of detecting and interpreting changes within these datasets. Change detection, however, is often complicated by a number

Discussion Paper

R

Discussion Paper

Discussion Paper

Discussion Paper

BGD

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page

Abstract Introduction

Conclusions References

Tables Figures

l∢ ⊳l

•

Back Close

Full Screen / Esc

Printer-friendly Version



Time series of

R. de Jong and S. de Bruin

vegetation indices

BGD

8, 8545-8561, 2011

Title Page

Abstract Introduction

Conclusions References

Tables Figures

I◀

•

Close

Full Screen / Esc

Back

Printer-friendly Version

Interactive Discussion



of statistical preconditions that are intrinsic to time series of spectral vegetation indices with dense sampling intervals. A frequently used approach for detecting temporal trends is to fit linear regressions of a vegetation index (VI) against time, but this needs to be done with care in order to avoid spurious trends. The detected slope coefficient (or gain) can be used to calculate the amount of change, but it is not always tested for significant deviation from zero, nor are standard statistical assumptions always respected (de Beurs and Henebry, 2005). Seasonal variation, for instance, is an important cause for the data to violate assumptions like homogeneous variation and absence of serial correlation in the residuals. This is typically remediated using temporal aggregation, where the aggregation window (or bin size) corresponds to the length of a calendar year. The resulting bins can be regarded as temporal units, which, like spatial units, are modifiable (Taylor, 2010). In case of spatial units, it has been demonstrated that the size may influence the model results, which is known as the Modifiable Area Unit Problem (MAUP) (Openshaw and Taylor, 1979). This problem may affect a myriad of spatial studies in geography (Dark and Bram, 2007) and remote sensing (Marceau et al., 1994). Similarly, there is a Modifiable Temporal Unit Problem (MTUP) that is as troublesome as the MAUP. In analysis of time series of satellite vegetation indices this problem is often disregarded, although it may result in misjudgements of temporal trends in the data. Aspects of the problem include the starting phase of a time series or segment, its extent and the level of temporal aggregation. The aim of this paper is to demonstrate possible MTUP effects in analysis of time series of satellite imagery using both real and simulated VI data and to provide, in this sense, a framework for time series regression.

20

Discussion Paper

Discussion Paper

Discussion Paper

Discussion Paper

Abstract

Introduction

Conclusions

References

Tables

Figures

14

▶I

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Trend analysis and the modifiable temporal unit problem

Many trend analysis methods exist, including parametric and non-parametric approaches. The most common method to detect changes in cyclic time series is the use of a linear model (Eq. 1) obtained from ordinary least-squares (OLS) regression. The slope coefficient or gain β is used to calculate the change in Y as β times the number of bins. This number is determined by the aggregation level (or: number of observations per bin), which is equivalent to the sample interval. Given that the datasets consist of 24 observations per year, aggregation level 24 corresponds to yearly bins and so on.

 $E(Y|T) = \alpha + \beta \times t$ (1)

The dependent variable Y can be any kind of VI or cyclical environmental parameter in general. The most common spectral vegetation indices are based on the rapid change in reflectance of chlorophyll between the red and near infrared (NIR) ranges. Here, we used the Normalized Difference Vegetation Index (NDVI), which is a commonly used proxy for terrestrial photosynthetic activity. A decrease over time is referred to as browning, whereas an increase indicates greening (de Jong et al., 2011).

Given this, we used a three-step approach to demonstrate the MTUP effect:

1. The influence of starting phase and data extent is illustrated using a perfectly harmonic model, without trend or noise components. Using a sample size of 24 observations per cycle this implies that the model residuals are far from independent, which invalidates linear regression by OLS. However, it provides a theoretical scenario to demonstrate our point that spurious slopes can be detected from cyclical data.

- **BGD**
- 8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page Introduction **Abstract** Conclusions References

Tables

Figures

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



2. Next, the sample was aggregated into bins of fixed size. These bins represent different temporal units (or aggregation levels). The MTUP effect is demonstrated by calculating the change in NDVI from linear models (Eq. 1) fitted for different bin sizes. The minimum number of bins over the full length of the synthetic time series was set to 5.

3. Finally, step 2 was repeated using real data from Advanced Very High Resolution Radiometer (AVHRR) sensors. The detected changes were compared to the corresponding change obtained using seasonally adjusted data without temporal aggregation. The seasonal adjustment was carried out using a Fourier method (Roerink et al., 2000) with four components, following de Jong et al. (2011). The significance of the slopes in the seasonally adjusted data was assessed using generalized least-squares (GLS) in order to account for remaining short-lag serial correlation. A sample of 1000 pixels was used for calculation of deviations introduced by the MTUP and the state of Queensland in north-eastern Australia was used to illustrate possible spatial patterns introduced by different temporal aggregation schemes.

Provided that the level of serial autocorrelation can be disregarded after aggregation, the significance of the detected trends can be tested using analysis of variance (ANOVA) with the hypotheses H0: $\beta = 0$ and HA: $\beta \neq 0$. All analyses were performed using standard R functionality (R Development Core Team, 2011).

Time series data

15

Synthetic time series were used to illustrate the effect of cyclic data on regression analysis. For this purpose, model parameters were chosen in such a way that they approximate the AVHRR time series described below for a temperate (non-forest) environment with a single growing season. As such, the peak-to-peak amplitude (2A) was set to 0.6 – with a mean of 0.4 – and the number of observations per year to 24. NDVI (Y) was simulated using a cosine model with no underlying positive or negative

$$y = \bar{y} + a \times \cos(t + \Phi) \tag{2a}$$

$$t = x/24 \times 2\pi \tag{2b}$$

The longest run of NDVI measurements is available from AVHRR sensors on board a series of National Oceanic and Atmospheric Administration (NOAA) satellites. Acquisition started in 1981 and is still on-going, but pre-processed datasets are for now available until the year 2006. Pre-processing includes correction for orbital decay, satellite changes and several atmospheric effects. The resulting NDVI values were aggregated into biweekly composites with ~8 km spatial resolution. The full description of the dataset and the processing steps is provided by Tucker et al. (2005). A sample of 1000 pixels was used for the MTUP analysis. Other vegetation indices, including the Enhanced Vegetation Index (EVI) or the Soil Adjusted Vegetation Index (SAVI). have similar statistical characteristics and therefore the problem described here is not restricted to NDVI.

Results and discussion

Synthetic data

In case of seasonal data with a dense sampling interval, both starting phase and extent influence the linear model. Figure 1a demonstrates that a fit on a perfectly harmonic model without linear trend resulted in a range of slope coefficients – positive or negative, varying with the starting phase – but never zero. Zero slopes are obtained only if both sides of a minimum or a maximum are equally sampled. This might, however, result in over- or underestimation of the mean NDVI (\bar{Y}) from the model intercept (α) . The slope coefficient is linearly related to the amplitude used for the seasonal model

Discussion Paper

Discussion Paper

Discussion Paper

Discussion Paper

8, 8545-8561, 2011

BGD

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page

Introduction **Abstract**

Conclusions References

Tables

Figures

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



8550

BGD

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page Introduction **Abstract** Conclusions References **Tables Figures** 14

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



and is inversely related to the extent of the time series (or segments). The latter is illustrated in Fig. 1b, which shows 4 different extents (respectively 2, 4, 6 and 8 yr) and the associated linear models for a given starting phase.

The slope coefficient changed with the level of temporal aggregation, which influ-5 enced the detected change in NDVI. This is illustrated in Fig. 2 using a similar synthetic data set as above but now having a length of 26 full cycles, which is comparable to that of AVHRR data sets. The variation in detected NDVI change increased with the level of aggregation, which resulted in larger uncertainty in model predictions. The true change in NDVI (i.e. zero) is only obtained by aggregation over complete cyclic periods, i.e. 1 or more years. Given the perfect periodicity of this synthetic example, this is true by definition, but in reality calendar years may not fit the periodicity of VI time series because of shifts in vegetation phenology and variations in growing season length. For this reason, the same analysis was also performed on real AVHRR time series (see Sect. 3.2).

As mentioned, the starting phase and extent influence each segment of the linear analysis. The effect reduces with longer extents, but it appeared that the change for the longest available run of NDVI data (AVHRR) is of the same magnitude as changes found by trend studies which account for seasonality (e.g. de Jong et al., 2011; Wang et al., 2011; Zhou et al., 2001). Table 1 lists the calculated changes in NDVI for a perfectly periodic model without trend component (Fig. 1) for time spans of AVHRR (26 yr), MODIS (11 yr) and 1 yr. The seasonal peak-to-peak amplitude was set to 1 for ease of comparison to other amplitudes. Given that the order of maximum change found in literature is \sim 4.0 \times 10⁻² for AVHRR data, trends induced by phase shift might introduce errors varying from 10 % to 90 % if seasonality is not accounted for.

AVHRR data

An example of the MTUP effect for an AVHRR pixel with a significant (p < 0.05) trend component is shown in Fig. 3. The change in NDVI was found to vary among aggregation levels: slight differences with respect to seasonally adjusted data occurred for

Interactive Discussion



1 or 2-yr windows and substantial differences for all other cases. This illustrates that aperiodicity in the data or in the aggregation window might result in considerable deviations from the change in NDVI found using seasonal correction instead of temporal aggregation.

If temporal aggregation is used to account for seasonality, the characteristics of the commonly used harmonic functions dictate the use of whole periods as aggregation windows, in order to force the deviation in slope coefficient to zero (Fig. 2). Aiming for bins holding complete seasons, the MTUP analysis was carried out on AVHRR data for aggregation levels of exactly 1 to 5 calendar years. It appeared that the significant changes in NDVI (i.e. significant for all aggregation levels) varied considerably with respect to trend analysis with seasonal adjustment instead of temporal aggregation. Table 2 lists the mean deviation and mean absolute deviation for the AVHRR sample and for the example in Fig. 3. This table shows that the mean deviation and the variation for the sample are low (around 1%) for fine aggregation levels but considerably higher in case of coarser aggregation: a mean absolute deviation of 24 per cent was found for 5-yr aggregation. In few cases the detected change in NDVI switched between positive (greening) and negative (browning), although these trends could not be confirmed using significance tests. It was almost exclusively found that the amount of detected change increased with the level of aggregation, so greening and browning trends appeared stronger after aggregation. Figure 4 illustrates this by showing spatial patterns of deviations for the state of Queensland in Australia. Spatial patterns are hardly perceivable at fine temporal aggregation levels, but they become apparent at coarser levels.

Temporal aggregation will only lead to unbiased estimates of VI trends if the following requirements are met:

 The regarded time series is perfectly periodic. Any type of aperiodicity, including phase shift or incomplete periods at the start or end, may result in incorrect model parameters. BGD

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page

Abstract Introduction

Conclusions References

Tables Figures

I∢ ►I

Back Close

Full Screen / Esc

2. The aggregation level corresponds exactly to the period of the seasonal signal (often one calendar year). Aggregating over multiple periods increases the risk of MTUP effects.

Other analysis methods may implicitly require similar conditions. For example, nonparametric trend tests such as the seasonal Mann-Kendall method (Hirsch and Slack, 1984) rely strongly on the absence of seasonal changes. In case of vegetation indices, this assumption often renders the method unsuitable for large spatial and temporal scales (de Jong et al., 2011). The results from this study indicate that similar effects may disturb linear regression. The effects, however, may not be as conspicuous because parametric models are less robust against this type of error than non-parametric models (McBride and Loftis, 1994). Seasonally adjusting the data using a decomposition method (e.g. Cleveland et al., 1990) provides another approach for eliminating serial autocorrelation and the MTUP. In that case, the seasonal model used should be appropriate for the growing regime and ideally should take possible seasonal changes into account. An example of such an approach is provided by Verbesselt et al. (2010).

Using OLS regression, trends may be found significant at a given aggregation level while not at another level. False positives (trend is found to be significant while it is not in reality) are likely to occur more frequently at low aggregation levels, but examples of the opposite case were found as well (not shown). If the significance of trends is not considered in the analysis, the MTUP may not only affect the gain but also the sign of detected changes in NDVI.

Conclusions

Ordinary least squares (OLS) time series regression can be used to quantify trends in cyclic data but temporal aggregation needs to be carried out carefully in order to avoid spurious results. The risk of artefacts is minimal at an aggregation level corresponding to a full period, for instance a calendar year. Coarser aggregation levels tend to

BGD

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page Introduction **Abstract** Conclusions References **Figures**

Tables

Back Close

Full Screen / Esc

Printer-friendly Version



overestimate the amount of change and result in higher variation in model predictions, especially from 3 periods onwards. However, the use of a full-period window may be impractical because VI time series are hardly ever free of changes in seasonality. Aperiodicity within long-term time series of vegetation indices is intrinsic to certain land cover types and may arise from variations in start and length of growing seasons as a result of variations in temperature and/or precipitation. The starting phase and the choice of aggregation level - or temporal unit - affect the estimation of model parameters, including the slope coefficient or gain. In this study, the amount of absolute change that was attributed to the modifiable temporal unit problem (MTUP) varied between 1 % and 24 % for full-period aggregation levels.

References

- Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I.: STL: A seasonal-trend decomposition procedure based on loess, Journal of Official Statistics, 6, 3-73, 1990.
- Dark, S. J. and Bram, D.: The modifiable areal unit problem (MAUP) in physical geography, Prog. Phys. Geog., 31, 471-479, 2007.
- de Beurs, K. M. and Henebry, G. M.: Land surface phenology and temperature variation in the International Geosphere-Biosphere Program high-latitude transects, Glob. Change Biol., 11, 779-790, 2005.
- de Jong, R., de Bruin, S., de Wit, A., Schaepman, M. E., and Dent, D. L.: Analysis of monotonic greening and browning trends from global NDVI time-series, Remote Sens. Environ., 115, 692-702, 2011.
- Hirsch, R. M. and Slack, J. R.: A nonparametric trend test for seasonal data with serial dependence, Water Resour. Res., 20, 727-732, 1984.
- Marceau, D. J., Howarth, P. J., and Gratton, D. J.: Remote sensing and the measurement of geographical entities in a forested environment. 1. The scale and spatial aggregation problem, Remote Sens. Environ., 49, 93-104, 1994.
- McBride, G. B. and Loftis, J. C.: The most important statistical aspects, Proceedings of the 1st international workshop Monitoring Tailor-Made (MTM): monitoring and assessment in water management, Beekbergen, The Netherlands, 153-161, 1994.

BGD

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page

Introduction **Abstract**

Conclusions References

Tables

Figures

Close

Back

Full Screen / Esc

Printer-friendly Version

- Discuss
- 8, 8545-8561, 2011

BGD

- Time series of vegetation indices
- R. de Jong and S. de Bruin
- Title Page

 Abstract Introduction

 Conclusions References

 Tables Figures
 - I4 ≯I
 - Back Close
 - Full Screen / Esc
 - Printer-friendly Version
 - Interactive Discussion
 - © BY

- Openshaw, S. and Taylor, P. J.: A million or so correlation coefficients: three experiments on the modifiable areal unit problem, in: Statistical applications in the spatial sciences, edited by: Wrigley, N., Pion, London, 127–144, 1979.
- Roerink, G. J., Menenti, M., and Verhoef, W.: Reconstructing cloudfree NDVI composites using Fourier analysis of time series, Int. J. Remote Sens., 21, 1911–1917, 2000.
- Taylor, R.: Communities, Crime, and Reactions to Crime Multilevel Models: Accomplishments and Meta-Challenges, J. Quant. Criminol., 26, 455–466, 2010.
- Tucker, C., Pinzon, J., Brown, M., Slayback, D., Pak, E., Mahoney, R., Vermote, E., and El Saleous, N.: An extended AVHRR 8km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, Int. J. Remote Sens., 26, 4485–4498, 2005.
- Verbesselt, J., Hyndman, R., Zeileis, A., and Culvenor, D.: Phenological change detection while accounting for abrupt and gradual trends in satellite image time series, Remote Sens. Environ., 114, 2970–2980, 2010.
- Wang, X., Piao, S., Ciais, P., Li, J., Friedlingstein, P., Koven, C., and Chen, A.: Spring temperature change and its implication in the change of vegetation growth in North America from 1982 to 2006, P. Natl. Acad. Sci., 108, 1240–1245, 2011.
- Zhou, L., Tucker, C. J., Kaufmann, R. K., Slayback, D. A., Shabanov, N. V., and Myneni, R. B.: Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999, J. Geophys. Res., 106, 20269–20283, 2001.



Table 1. Detected changes in NDVI using linear regression on synthetic seasonal data for different starting phases and extents. Change values were calculated as the slope coefficient β (Eq. 2) times the length of the time series and were multiplied by 100 for numerical convenience. The listed lengths of 26 yr and 11 yr correspond to the length of AVHRR and MODIS time series respectively. The applied harmonic model had peak-to-peak amplitude 1 and therefore change values can be multiplied by the actual amplitude.

phase shift (months)	change in NDVI * 100							
	26 yr	11 yr	1 yr					
0	-0.48	-1.14	-12.52					
1.5	2.24	5.30	58.40					
3	3.65	8.63	95.11					
4.5	2.92	6.91	76.11					
6	0.48	1.14	12.52					
7.5	-2.24	-5.30	-58.40					
9	-3.65	-8.63	-95.11					
10.5	-2.92	-6.91	-76.11					

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page

Abstract Introduction

Conclusions References

> **Figures Tables**

14 **▶**I

Full Screen / Esc

Table 2. The MTUP effect for a sample of 1000 AVHHR pixels and aggeregation levels (AL) of 1 to 5 yr. Detected changes in NDVI (dNDVlagg) were compared to those found using seasonally adjusted data (dNDVIref) and listed as dNDVlagg–dNDVIref (dNDVI), dNDVlagg/dNDVIref (pct) and square root of the sample variation of pct (sd). The last columns list the detected changes in NDVI for the pixel in Fig. 3. In all cases, NDVI values were multiplied by 100 for numerical convenience.

AL (yr)	mean deviation			mean absolute deviation			Fig. 3	
	dNDVI	pct	sd	dNDVI	pct	sd	NDVI	pct
0							2.75	
1	0.015	100.5%	1.6%	0.044	101.0%	1.3%	2.77	2.0%
2	0.019	100.6%	3.0%	0.084	102.2%	2.2%	2.89	14.0%
3	-0.016	102.4%	9.0%	0.264	107.2%	5.8%	2.96	21.0%
4	0.022	105.4%	15.7%	0.469	112.9%	10.5%	3.47	72.0%
5	0.066	106.6%	31.4%	0.950	124.1%	21.1 %	2.28	-47.0%

BGD

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I₹











Full Screen / Esc

Printer-friendly Version





8, 8545-8561, 2011

BGD

Time series of vegetation indices

R. de Jong and S. de Bruin

Introduction

References

Figures

▶I

Close



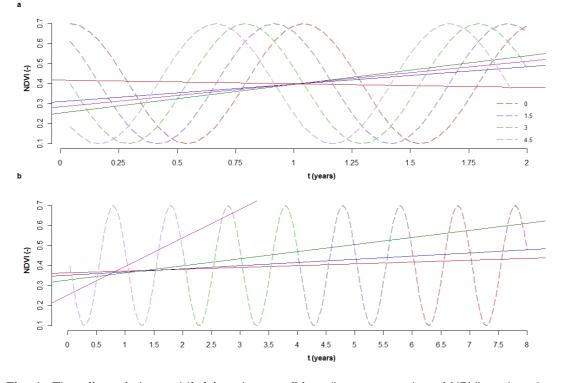


Fig. 1. The effect of phase shift (a) and extent (b) on linear regression of NDVI against time for seasonal data. The legend in (a) shows the phase shift in months. In (b), the colors each represent an additional extent of 2 yr. Accordingly, magenta, green, blue and red refer to extents of 2, 4, 6 and 8 yr respectively.

Printer-friendly Version



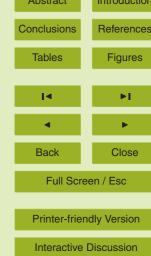
8, 8545-8561, 2011

Time series of vegetation indices

BGD

R. de Jong and S. de Bruin





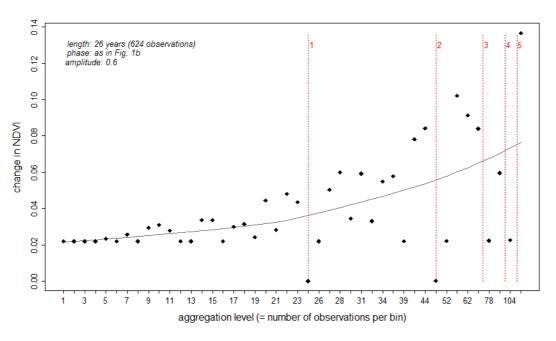


Fig. 2. Change in NDVI detected using linear models depending on the temporal aggregation level. The plot illustrates the MTUP effect for a harmonic model of 26 yr (length of common AVHRR datasets) and an amplitude of 0.6 NDVI units. The points shown are the ones that resulted in a change in the number of bins, while increasing the aggregation level. The vertical dotted lines indicate aggregation levels corresponding to 1, 2, 3, 4 and 5 yr respectively and the grey line shows a LOESS curve indicating an average trend.



8, 8545-8561, 2011

Time series of vegetation indices

BGD

R. de Jong and S. de Bruin

Title Page



Printer-friendly Version



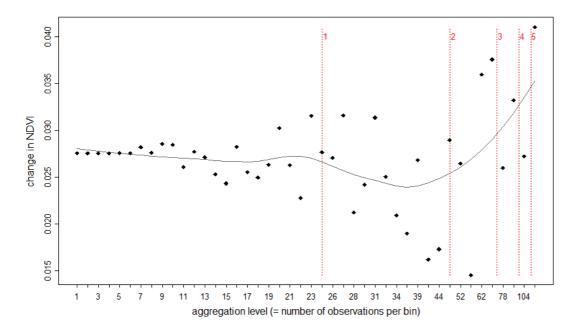


Fig. 3. Change in NDVI detected using linear models depending on the temporal aggregation level. The plot illustrates the MTUP effect for an AVHRR pixel with significant (p < 0.05) trend component (location: 46.21° N/110.65° E, Montana, USA). See the caption of Fig. 2 for additional information and Table 2 for the NDVI values corresponding to full-year aggregation levels.

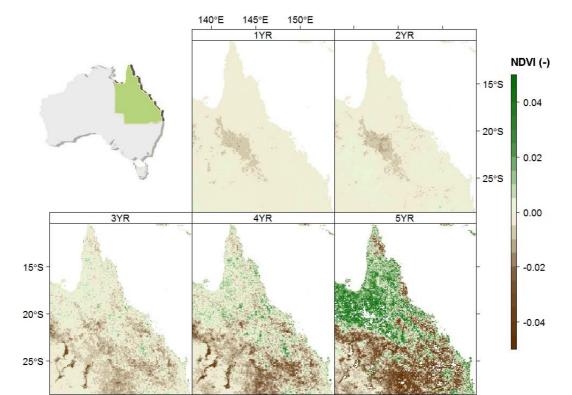


Fig. 4. Detected changes in NDVI for different levels of temporal aggregation – example for Queensland (Australia). The top-left map shows the location of Queensland and the 5 panels show the change in NDVI with respect to changes found using seasonal adjustedment without temporal aggregation. The aggregation levels correspond to 1, 2, 3, 4 and 5 yr, respectively.

140°E

145°E

150°E

145°E

140°E

150°E

BGD

8, 8545-8561, 2011

Time series of vegetation indices

R. de Jong and S. de Bruin

Title Page

Abstract Introduction

Conclusions References

Tables Figures

Close

Full Screen / Esc

Back

Printer-friendly Version

