

Sebastian Mader

Vom Fachbereich VI
(Geographie/Geowissenschaften)
der Universität Trier
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
genehmigte Dissertation

*A Framework for the Phenological Analysis of Hypertemporal
Remote Sensing Data Based on Polynomial Spline Models*

Betreuender:

Prof. Dr. Joachim Hill

Berichterstattende:

Prof. Dr. Joachim Hill

Prof. Dr. Thomas Udelhoven

Datum der wissenschaftlichen Aussprache:

8. November 2012

Trier, 2012

Zusammenfassung

Zeitreihen-Archive von Fernerkundungssatelliten zur Umweltbeobachtung bieten vielfältige Möglichkeiten, um die an der Erdoberfläche ablaufenden Prozesse in ihrer räumlichen und zeitlichen Dynamik zu erfassen. Am häufigsten beinhalten solche Archive Vegetationsindizes, die das charakteristische spektrale Reflexionsverhalten grüner Vegetation im roten und nah-infraroten Bereich des elektromagnetischen Spektrums ausnutzen, um ein Beobachtungsmerkmal zur Verfügung zu stellen, welches den Bedeckungsgrad und die Dichte der Vegetation an der Erdoberfläche in direkter Weise widerspiegelt. Die Aufnahmen hypertemporaler Erdbeobachtungssatelliten, zu kontinuierlichen Zeitreihen mit einer ein- bis zweiwöchigen Wiederholrate zusammengestellt, bieten einen guten Überblick über die verschiedenen Wachstumsphasen der Vegetation während einer Vegetationsperiode. Durch die Analyse solcher Zeitreihen können phänologische Parameter oder Metriken abgeleitet werden, wie zum Beispiel der Beginn und das Ende einer jährlichen Wachstumsperiode, sowie deren Länge. Obwohl diese phänologischen Parameter nicht genau den konventionell beobachteten phänologischen Ereignissen entsprechen, so bieten sie doch Anhaltspunkte für die in der Biosphäre ablaufenden dynamischen Prozesse. Für jedes Jahr eines längeren Zeitreihen-Archivs berechnet, bilden solche Metriken eine flächendeckende Basis für eine Vielzahl von Anwendungen, wie etwa Untersuchungen zum Klimawandel, oder zur Vegetationsüberwachung für die Ernährungssicherung einer rasch wachsenden Weltbevölkerung. Eine Herausforderung in Bezug auf phänologische Metriken ist die Entwicklung robuster Algorithmen zu deren Ableitung, denn Störungen durch Wolken, Absorption und Streuung in der Atmosphäre, sowie Sensorrauschen oder Sensorausfälle können dazu führen, dass Zeitreihen von Vegetationsindizes fehlerhafte Beobachtungen oder Datenlücken aufweisen. Daher werden für die Bestimmung der phänologischen Parameter aus Zeitreihen häufig mathematische Modelle und Verfahren verwendet, um Beobachtungslücken aufzufüllen oder den Einfluss fehlerhafter Beobachtungen auf das Endergebnis zu verringern. Die hierbei am häufigsten eingesetzten Verfahren sind digitale Filter und die Fourieranalyse von Zeitreihen. Polynomische Spline-Modelle sind Verfahren zur Kurvenanpassung, die eine beliebige Datenreihe in kontinuierlicher Art und Weise durch stückweise Polynome abbilden. Solche Spline-Modelle bieten eine vielversprechende Alternative zu bereits bestehenden Verfahren, da sie durch ihre spezifischen mathematischen Eigenschaften einerseits dazu beitragen können, bereits bekannte Probleme bei anderen Verfahren zu vermeiden, und andererseits vielfältige Möglichkeiten zur analytischen Auswertung einer Zeitreihe bieten, um phänologische Metriken abzuleiten. Zudem sind Spline-Modelle

sehr flexibel und können auf eine Vielzahl verschiedener Datensätze mit unterschiedlichen Charakteristika angewendet werden, ohne dass dafür eine umfangreiche Vorverarbeitung der Datensätze erforderlich ist. Im Rahmen der vorliegenden Arbeit wurde auf Basis verschiedener Spline-Modelle, die für die Analyse fernerkundlich erhobener Zeitreihen geeignet sind, ein Ansatz entwickelt und implementiert, der es erlaubt, diese Modelle für die Berechnung phänologischer Metriken zu nutzen. Anhand eines Beispieldatensatzes eines intensiv landwirtschaftlich genutzten Gebietes mit hoher Vegetationsdynamik wurde der neu entwickelte Ansatz verifiziert, indem ein Vergleich mit zwei alternativen, bereits etablierten Verfahren zur Berechnung phänologischer Parameter aus Fernerkundungsdaten durchgeführt wurde.

Contents

1	Introduction	4
1.1	Optical spectral properties of vegetation	5
1.2	Vegetation indices	5
1.3	Phenological descriptors derived from time series	7
1.4	An example: the Aksu oasis	8
1.5	Objectives	11
2	Models for remotely sensed time series	13
2.1	Digital filtering	15
2.2	Frequency domain methods	18
2.3	Fitting of exponential and logistic functions	23
2.4	Polynomial functions and splines	26
2.5	Wavelets	29
2.6	Linear decomposition	30
2.7	Summary: why splines?	31
3	A spline framework for remote sensing time series analysis	35
3.1	Piecewise polynomials	35
3.2	Spline representations and basis functions	39
3.3	B-splines	40
3.4	Fitting B-splines to observational data	45
3.5	Smoothing splines	48
3.6	Fitting periodic spline models	53
3.7	Knot number and knot placement	53
3.8	Derivatives of B-splines	54
3.9	Finding zeroes and inverse values	55
4	Estimation of phenological parameters	58
4.1	Choosing an appropriate spline	58
4.2	Deriving phenological descriptors from splines	62
4.3	Comparison to other methods and algorithms	65
4.4	Comparison sampling design	66
5	Results and discussion	70
5.1	Start of season	72

Contents	2
5.2 Length of season	73
5.3 Seasonal integrals	76
5.4 Trends	78
6 Conclusions and outlook	83
A Description of Computer Programs	86
A.1 splview	86
A.2 splfit	87
A.3 splcal	89
A.4 phencal	90
References	92

Acknowledgements

I am glad that this thesis found the interest and support of numerous people, whose concern and effort is gratefully acknowledged.

Prof. Dr. Joachim Hill, the supervisor of this thesis, gave me the opportunity to work on an interesting subject. I am grateful for his encouraging and supportive attitude, and most important, for his guidance, flexibility and understanding whenever things did not quite work out as expected.

Prof. Dr. Thomas Udelhoven provided the “TimeStats” software used to perform the trend analyses of the phenological parameters and accepted to be the second supervisor of this thesis.

Dr. Wolfgang Mehl wrote the “ts_phen_ind” digital filtering software for deriving phenological parameters from time series. He also kindly took the time and used his software to compute a set of phenological parameters for the Aksu test area presented in this work. The data set he provided was very useful as an independent reference to compare and verify the results of the phenological parameterisation approach developed within this thesis.

The MODIS vegetation index data sets (MOD13Q1 product, collection 5) used in this work were obtained through the online Data Pool at the NASA Land Processes Distributed Active Archive Data Center (LP DAAC), USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota¹.

I owe a lot to all my committed colleagues and the student assistants at Trier University’s department of Environmental Remote Sensing & Geoinformatics, especially Dr. Achim Röder, Dr. Marion Stellmes, Dr. Johannes Stoffels, Tobias Dick and Mandy Maas. My colleagues were always helpful and took extra teaching duties to cut me some spare time I could spend working on my thesis. Dr. Marion Stellmes took the trouble of carefully proof-reading a bigger part of this work’s manuscript, and the final version benefited a lot from her comments and suggestions. Tobias Dick, trusted whiz of Unices, backed me up by sharing some of my regular duties as an IT systems administrator.

Finally, I would like to express my gratitude to my parents Ursula and Josef Mader, and to my sister Dr. Heike Mader, for their care and support, and for carrying me through some harder times. For as long as I can remember, my parents always encouraged me to unfold and develop my own personality, imagination and abilities. I wouldn’t have come as far as writing a PhD thesis without them.

¹https://lpdaac.usgs.gov/get_data

1 Introduction

Time series of remotely sensed imagery have proven to be a valuable tool to study the dynamics of the Earth's surface. Semiautomatic data processing and compositing schemes allow the generation of continuous time series of vegetation indices and other vegetation related parameters such as the fraction of absorbed photosynthetically active radiation (FAPAR), with good quality and reliability (Justice et al., 1998; Yang et al., 2006). The data acquired by Earth observation satellite systems with high rates of repetitive coverage, or *hypertemporal* satellite systems, are compiled into consistent data archives with global coverage at weekly to bi-weekly intervals. Most of these global data fields produced from sensors such as AVHRR (Sellers et al., 1994; Tucker et al., 2005), MODIS (Huete et al., 2002; Yang et al., 2006) or SPOT VEGETATION are now available at no charge and provide an efficient and economic means to observe the environment. Typically, these data fields have a spatial resolution of 250 m to 1 km. In contrast to hypertemporal sensors, satellite systems designed for applications at a regional level such as the Landsat series provide a higher spatial resolution, but are not capable of realising the high repetitive coverage rates of hypertemporal systems (Jones and Vaughan, 2010; Vinciková et al., 2010).

Time series in hypertemporal satellite archives allow the spatially distributed assessment of the vegetation at the Earth's surface through the physical link of vegetation indices to vegetation density and abundance. The detailed information in the time domain provides a summary of the various stages and transitions of vegetation during a growing period. Thus, time series of vegetation indices can be analysed to extract phenological parameters (descriptors, metrics) for consecutive growing periods. Such phenological parameters are, for example, the start and end of a growing period, its length, and the location and value of a peak within a vegetation period. Although these phenological descriptors do not correspond directly to phenological events that are observed in situ, such as bud burst or leaf out, they are still indicators for vegetation and ecosystem dynamics. Phenological information derived from satellite imagery has been used in land cover classification to distinguish vegetation types (Peters et al., 1997; DeFries et al., 1995), in the assessment of land degradation dynamics (Hill et al., 2008), as well as to estimate phenological variables for ecological (Archibald and Scholes, 2007) and climate change (White et al., 1997, 2009) assessments.

The light reflected from the Earth's surface that is detected by satellite sensors is affected by various sources of error, including aerosols and clouds in the atmosphere, changes in the sun and sensor viewing angles, as well as data dropouts due to sensor failure. All these phenomena give rise to noise in the time series of satellite archives. A challenge in the phenological analysis of remotely sensed time series is to develop robust algorithms for the derivation of phenological parameters from real, noisy data.

1.1 Optical spectral properties of vegetation

The spectral reflectance of a vegetation canopy is primarily determined by the reflective and transmissive properties of its leaves (Curran, 1980). The controlling factors of leaf reflectance and transmittance are leaf pigments, cell structure and water content. Figure 1.1 shows the spectral reflectance of a vigorous and a senescent vegetation canopy in comparison to the spectrum of a soil. The data are taken from the ASTER spectral library (Balridge et al., 2009). The low reflectance of vegetation in the visible region of the electromagnetic spectrum from 0.4 to 0.7 μm is a consequence of the absorption of the leaf pigments. Higher plants contain four primary pigments, chlorophyll a, chlorophyll b, β carotene and xanthophyll, all of which absorb visible light for photosynthesis (Curran, 1980). Vegetation reflectance is lowest in the blue and red regions of the visible spectrum. These regions correspond to the absorption bands of the two most important leaf pigments, chlorophyll a (0.43 to 0.66 μm) and b (0.45 to 0.65 μm) (Curran, 1980; Jensen, 1983). The relative lack of absorption in the green visible wavelengths between the two chlorophyll absorption bands produces the green peak at approximately 0.54 μm . In the near infrared (NIR) spectral region between 0.75 and 1.1 μm , the cell structure of the leaves becomes the major controlling factor of canopy reflectance. In this wavelength range, healthy green vegetation is generally characterised by high reflectances and transmittances of 45 to 50 percent (Jensen, 1983). The sharp increase in vegetation reflectance at the end of the visible part of the electromagnetic spectrum (about 0.7 μm) is called red edge and is a result of the scattering of light at the interfaces of the cell walls of the leaves (Jensen, 1983). In the NIR region, a part of the electromagnetic radiation is transmitted through a leaf onto underlying leaves, and the reflected radiation in this spectral region becomes enriched as the number of reflections at leaf surfaces inside the canopy increases (Jones and Vaughan, 2010). As a result, the reflected energy in the NIR spectral region is correlated with plant biomass (Jensen, 1983). As can be seen from the dry grass spectral response in figure 1.1, the breakdown of plant pigments in senescent vegetation causes the reflectance in the visible wavelengths to rise, while the reflectance in the NIR does not significantly decrease (Curran, 1980).

1.2 Vegetation indices

Vegetation indices exploit the characteristically high reflectance difference between the red and near infrared spectral regions to summarise the amount and condition of vegetation present within a scene (Curran, 1980; Jackson and Huete, 1991). The most common vegetation index used in remote sensing data archives is the normalised difference vegetation index (NDVI). The NDVI is calculated as (e.g. Jones and Vaughan, 2010):

$$\text{NDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + \rho_{\text{red}}}$$

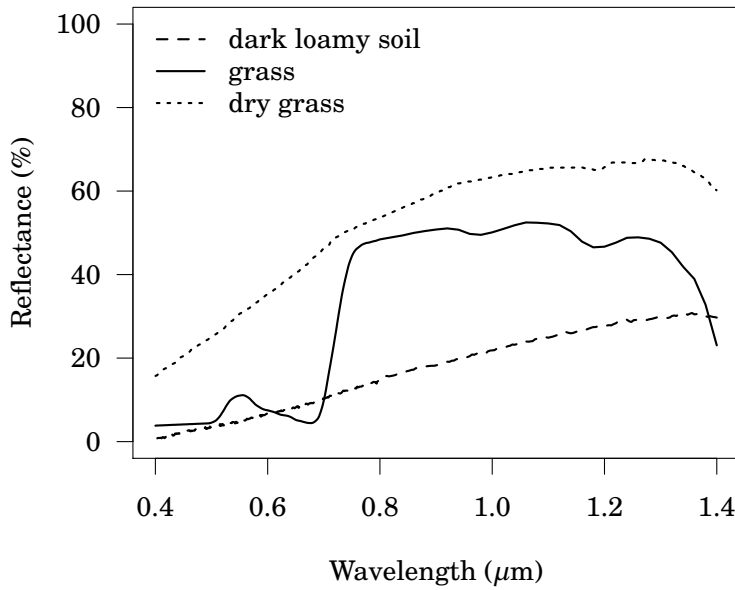


Figure 1.1: Spectra of vigorous and senescent vegetation in the visible and near infrared in comparison to a dark soil spectrum. Green vegetation is uniquely characterised by a sharp increase in reflectance around $0.7 \mu\text{m}$. The data are from the ASTER spectral library (Balridge et al., 2009).

where ρ_{NIR} is a reflectance or radiance value in a near infrared band, and ρ_{red} is a reflectance or radiance value in a spectral band corresponding to red visible light. The range of values of the NDVI is between -1 and 1. Higher NDVI values indicate a higher level of green vegetation density and coverage. Negative NDVI values indicate nonvegetated areas such as water, ice, snow and barren surfaces. However, any surface with a reflectance spectrum that exhibits a relatively higher reflectance in the near infrared than in the red spectral region, such as a typical soil spectrum (see figure 1.1), will yield a positive NDVI value. Price (1993) showed that bare soils could produce NDVI values as large as 0.3. This fact makes the NDVI sensitive to background reflectance at low vegetation cover. At high vegetation cover, saturation of the index may occur so that increasing vegetation densities do not cause a proportional increase in NDVI (Huete and Jackson, 1987; Jasinski, 1990). Therefore, alternative vegetation indices have been proposed such as the perpendicular vegetation index (PVI) of Richardson and Wiegand (1977) or the soil adjusted vegetation index (SAVI, Huete, 1988) or the tasseled cap transform (Crist and Kauth, 1986). The enhanced vegetation index (EVI) is routinely derived from MODIS imagery besides the NDVI. The EVI was developed to provide improved sensitivity in regions of high vegetation density (Huete et al., 2002). The equation of the EVI takes the form

$$\text{EVI} = G \frac{\rho_{\text{NIR}} - \rho_{\text{red}}}{\rho_{\text{NIR}} + C_1 \rho_{\text{red}} - C_2 \rho_{\text{blue}} + L}$$

where ρ_{NIR} and ρ_{red} are the surface reflectances in the near infrared and red spectral bands, L is a canopy background adjustment term, and C_1 and C_2 are the coefficients of an aerosol resistance term, which uses a spectral band in the visible blue region to correct for aerosol influences in the red spectral band (Huete et al., 2002). The data range of the EVI corresponds to that of the NDVI.

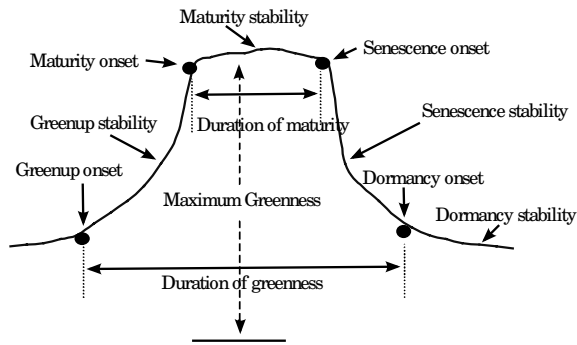


Figure 1.2: Phenological transition points and phases (after Zhang et al., 2001).

1.3 Phenological descriptors derived from time series

Vegetation phenology, the study of recurring vegetation cycles and their connection to environmental driving forces is an important subject in a wide variety of Earth science applications. The assessment of phenological parameters such as the seasonal timing of the onset of vegetation development and dormancy, the length of the growing season and the strength of annual or semi-annual vegetation cycles is crucial for understanding the structure and function of ecosystems, as well as their interaction with climate and human activities. (White et al., 1997, 2009; Zhang et al., 2001). Phenological parameters derived from remotely sensed time series of vegetation indices can be used as a substitute to biophysical phenological variables (Tucker et al., 1981; DeFries et al., 1995; White et al., 2009). Figure 1.3 schematically shows a typical vegetation index profile of a single growing period, illustrating the various key phenological transition points and phases, as well as the most important phenological parameters associated with them. The parameters describe the on- and offset of vegetation greenness, the details of vegetation greenup, maturity and senescence, e.g. the length of the growing period or the period of maturity. Figure 1.3 also illustrates intensity parameters associated with individual phenological phases, such as the maximum value of a vegetation index reached at the maturity stage, termed maximum greenness. The integral under the vegetation index curve from the onset of greenness to the onset of maturity may be used to estimate biomass (Tucker et al., 1981). Table 1.1 gives an overview over selected phenological descriptors that can be obtained from vegetation index time series and their biophysical equivalents. There are various approaches to derive the timing for the start of the season from vegetation index time series (de Jong et al., 2011). The simplest methods are based on a constant vegetation index threshold. However, according to White et al. (2009), these methods should be abandoned because they are not consistent across ecosystems. Jönsson and Eklundh (2002) used a variable threshold based on the amplitude of the vegetation cycle in a given year. They defined the start of season as the point along the time series where the vegetation index exceeds a value equivalent to ten percent of the annual amplitude. A special case of a variable threshold is the half-maximum method (White et al., 1997), where the threshold is set at half the maximum greenness value. It was argued by (Bradley et al.,

Metric	Biophysical interpretation
Maximum VI in year	Greenness of vegetation at peak of growing season
VI amplitude	Range in greenness of vegetation in year
VI threshold	Greenness of vegetation at beginning or end of growing season
Rate of VI increase during greenup	Rate of increase in photosynthetic activity per month of greenup
Rate of VI decrease during senescence	Rate of decrease in photosynthetic activity per month of senescence
Integrated VI over growing season	Total greenness of vegetation in growing season, biomass

Table 1.1: Overview over a selection of phenological metrics based on vegetation indices (VI) and their biophysical equivalents (after DeFries et al., 1995).

2007) that the half-maximum method is stable and consistent across ecosystems. Other researchers used derivative information such as inflection points (Moulin et al., 1997) or maximum curvature (Zhang et al., 2003) to determine the start of the season, while Reed et al. (2003) used delayed and forward-looking moving averages.

Not all of the phenological parameters presented in figure 1.3 and table 1.1 are equally well suited for different ecosystems or vegetation types. The profile in figure 1.3, for example, has a plateau-like shape during the stage of maturity, while other curves may exhibit a distinct phenological peak (see figure 1.4).

1.4 An example: the Aksu oasis

Remote sensing vegetation dynamics and phenology is particularly useful to monitor the effects of human interaction with the environment. An example of an area that has been strongly affected by human activities in the past decade is the Aksu oasis located in the Tarim river basin in the northwestern part of the People's Republic of China. (Jiang et al., 2005; Xu et al., 2005; Hao et al., 2008; Wang et al., 2010). The Aksu river basin is located at the northwestern edge of the Tarim basin, between $75^{\circ}35' - 82^{\circ}00'$ E and $40^{\circ}17' - 42^{\circ}27'$ N (Wang et al., 2010). The oasis region is characterised by a warm, continental arid climate. The annual mean temperature is about 10°C , and there are large temperature differences among days and years (Wang et al., 2010). There is only a sparse annual precipitation of about 60 mm, and the main water supply for the intensive agriculture in the region is through the Aksu and Tarim rivers, which are nourished by melting snow and glaciers from the mountains surrounding the Tarim basin (Xu et al., 2005; Wang et al., 2010).

Population growth and the improvement of socioeconomic conditions in modern times have led to the rapid expansion and intensification of cultivated land in the region (Hao et al., 2008). Most of the cultivators of land in the Tarim river basin are engaged in cotton

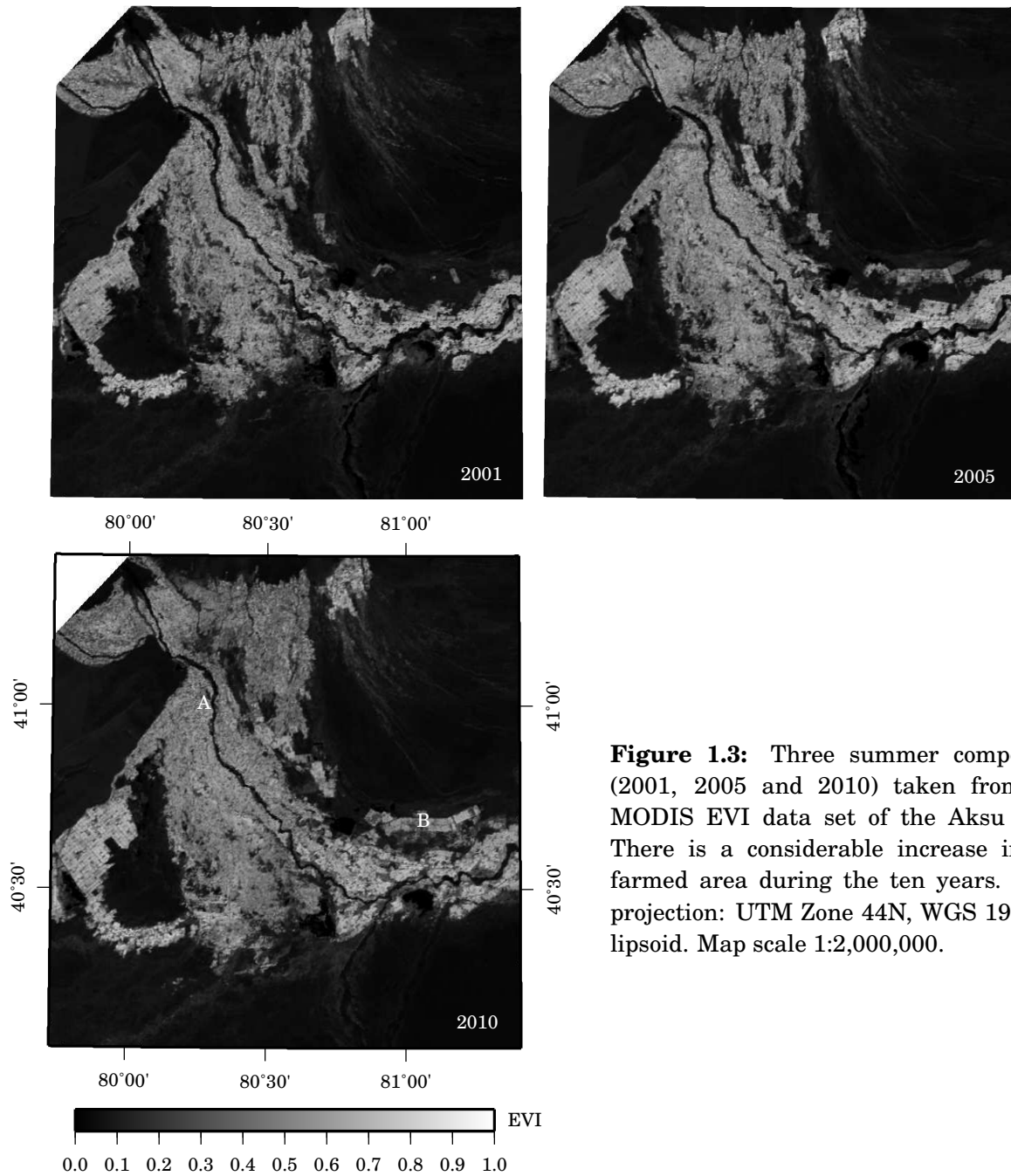


Figure 1.3: Three summer composites (2001, 2005 and 2010) taken from the MODIS EVI data set of the Aksu area. There is a considerable increase in the farmed area during the ten years. Map projection: UTM Zone 44N, WGS 1984 ellipsoid. Map scale 1:2,000,000.

production, and a large part of formerly unused land was converted into irrigated cotton fields (Xu et al., 2005; Jiang et al., 2005; Hao et al., 2008). The concomitant high demand for irrigation water in combination with the high evapotranspiration rates (Wang et al., 2010) lead to a number of environmental problems in the area, such as water shortage, salinisation of soils, and decrease of water quality in the Aksu river, e.g. through the discharge of water used in irrigation or to flush off the salt content of newly claimed agricultural land (Jiang et al., 2005).

The increase in area of cultivated land is evident from a time series of MODIS EVI images of ten years from 2001 to 2010 for a section of the Aksu river basin. In figure 1.3, three individual composites are shown from the years 2001, 2005 and 2010. The spatial

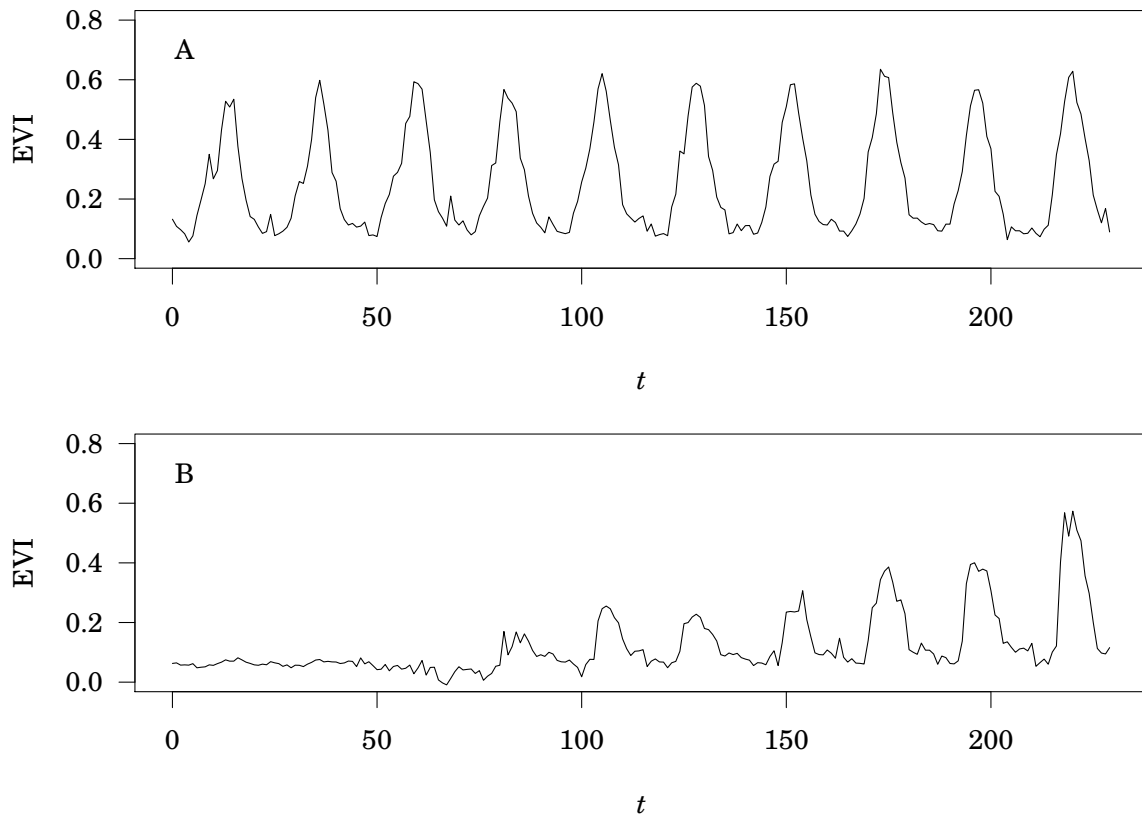


Figure 1.4: Two example EVI time series trajectories from the Aksu area. The spatial locations of the trajectories are marked in the lower left panel of figure 1.3

resolution of the MODIS EVI product is 250 m. From the comparison of the three images it can already be seen that the area of surfaces with high EVI values is constantly increasing over the period of ten years. This is most obvious in the eastern part of the images, where there is an area with initially low EVI values from which patches of higher EVI values emerge (marked “B” in the lower left panel of figure 1.3. In the course of time, the patches connect to form a contiguous area of high EVI values in 2010. The increase in EVI values indicates that agricultural fields were newly created in these areas from what was formerly a part of the desert surrounding the Aksu river. The dynamics of this development becomes fully apparent by looking at the entire time series trajectories from individual locations in the image. Figure 1.4 shows two example EVI trajectories from two selected locations in the MODIS EVI data set. The location marked “A” in the lower left panel of figure 1.3 is an example of a relatively stable location in the direct vicinity of the Aksu river that shows a strong annually recurring vegetation cycle with an amplitude that is approximately constant over the ten year period covered by the data set. In contrast, the location at “B” in figure 1.3 lies within an area that was newly cultivated from scratch. The corresponding time series record is shown in the lower part of figure 1.4. The time axis is labelled in terms of observation indices, where the zero index corresponds to the observation at the first day of year in 2001. The MODIS EVI product provides observations at a rate of 16 days, so that there are 23 observations in every year. The flat EVI profile at the beginning of the time series record in the lower part of figure

1.4 indicates that the area depicted by the corresponding MODIS pixel was not vegetated until 2004, when a weak annual vegetation cycle starts to emerge. Over the seven years from 2004 to 2010, the amplitude of the annual cycles continuously increases, indicating an intensification of the cultivation. By 2010, the amplitude of the annual vegetation cycle has reached a level similar to that of the permanently cultivated area “A” in the upper panel of figure 1.4.

By summarising the vegetation information contained in the EVI time series imagery on a per-year basis using phenological parameters, the dynamics of such land use changes can be studied in a spatial context. For example, a trend analysis of the amplitudes or seasonal integrals of the vegetation cycles of consecutive years may reveal the spatial extent of areas that are subject to change, as well as the rate of change in these areas. The results of such studies could have implications for management decisions to avoid or ease the serious environmental problems caused by the uncontrolled extension of agricultural areas by irrational human activities.

1.5 Objectives

Phenological descriptors can be derived from remotely sensed data in various ways using methods of time series analysis such as digital filtering, Fourier analysis and curve fitting methods. The ability to retrieve estimates of phenological parameters varies considerably between different methods and is generally ecosystem dependent (White et al., 2009). From a review of existing methods, which is presented in chapter 2, it may be concluded that polynomial splines have the potential to provide general, entirely data driven models that can be used to extract phenological descriptors from remotely sensed time series across different ecosystems. This potential stems from the versatility and flexibility of polynomial spline models, as well as their favourable mathematical properties which allow e.g. the analytical determination of derivatives to locate extreme values along a modelled time series. Informations about extreme values and derivatives can serve as an orientation to define the key phenological transition points and phases present in a given time series record. These advantages should improve the estimates of phenological parameters derived from splines as compared to other methods.

The objective of this work is to establish a framework for the use of spline models in the analysis of remotely sensed time series, with a focus on the use of these models to determine the vegetation cycles in the time series and estimate their phenological parameters. To this end, different spline models must be implemented in a way that they can be efficiently used with remotely sensed imagery.

Due to the coarse resolution of remotely sensed time series products, it is generally difficult to verify the results obtained from such products directly using ground truth data because of the scale gap between in situ and remotely sensed measurements (Liang et al.,

2002; Liang, 2004). Nevertheless, the capability of the spline models to determine phenological parameters can be tested by comparing the results to other existing methods for the determination of phenological descriptors.

Two well known alternative methods to derive phenological parameters are based on moving average filtering and fitting of Gaussian curves. To assess the validity of spline models in the determination of phenological metrics relative to the two alternative approaches, all three methods will be applied to determine the phenological parameters of the MODIS EVI time series of the Aksu oasis presented in chapter 1.4. The Aksu region provides an ideal test bed to compare the three methods, since it exhibits clear and strong phenological signals from agricultural areas, where some signals are steady and approximately constant (such as the example in the top panel of figure 1.4), while others exhibit strong temporal dynamics (such as the example in the bottom panel of figure 1.4). Because of these strong and distinct phenological signals, it is expected that the results of the three methods be consistent.

2 Models for remotely sensed time series

The variation in remotely sensed time series, such as vegetation index records from successive overpasses of earth observation satellites over a given area, may be assumed to originate from three different sources, namely cyclic variations, a general trend, and irregular fluctuations. Time series of vegetation observations typically exhibit cyclic variation that is annual in period. Apart from the annual fluctuations, additional cyclic variation may occur at other periods, e.g. in the tropics or in irrigated agricultural systems, where there may be a semiannual cyclic pattern. Trend may be broadly defined as a systematic long-term change in the mean level of the time series data. It is most important to have a clear understanding what “long-term” means in a given context. In the case of high temporal resolution earth observing satellites, the records go back until 1978, when the first AVHRR instrument on board the TIROS-N spacecraft became operational and started to provide imagery of the earth’s surface. Thus, the longest continuous time series of satellite observations span about 30 years. Having only three decades of data, a cyclic variation with a very long period of hundreds of years, for example caused by climatic influence factors, would appear as a trend. Although the true nature of cyclic components with periods longer than the record length cannot be assessed, it may still be meaningful to treat such long-term cycles as trend components in models tailored for shorter time series. The phrase “long-term” is always a subjective notion, depending on the number of observations available in a time series. Trend estimates from satellite observations are limited by a short record length and may be incompatible with analyses on longer time scales (Lambin, 1996; Myneni et al., 1997; White et al., 2009). When the variations due to the cyclic and trend components have been removed, there remains a time series of residuals that may be treated as random. It is common that the residual errors of time series models are autocorrelated (Bowman and Azzalini, 1997). When there is no systematic change in mean level with time (no trend) and no systematic change in variance in a time series, the time series is called *stationary* (Chatfield, 2004). Stationarity means that the statistical properties of the data in any one segment of the time series are comparable with those of any other segment. The concept of stationarity may be extended to higher order moments. Sometimes the term *weakly stationary* is used for time series that are stationary in mean and variance only¹.

There are three major tasks associated with time series analysis: prediction (also known as forecasting²), modelling and characterisation (Gershenfeld, 1999; Chatfield, 2004). Prediction deals with the estimation of future values of the time series based on past observations. A special case of prediction is noise reduction and the filling in of missing

¹The term *stationary* is used here to mean *weakly stationary*.

²Sometimes, the two terms “prediction” and “forecasting” are meant to distinguish between subjective and objective statistical methods, respectively (see e.g. Chatfield 2004). Here, both terms are used interchangeably.

values, where an underlying true series has to be estimated from a given erroneous time series, possibly with an incomplete record. Modelling involves the determination of a set of governing equations (the model) that describe a time series. A time series may be characterised by somehow specifying the kind of system that produced an observed signal. In the case of remotely sensed time series of vegetation indices, this may include the type of vegetation, the length of the growing period, estimating the amount of biomass, etc. The three tasks are closely related, but not identical: the most efficient way to characterise a time series may be by using a model, and a model may not work properly with data that is not noise reduced. The many methods of analysis of vegetation characteristics using time series comprise digital filters (Chen et al., 2004), exploratory methods such as principal components analysis (Eastman and Fulk, 1993; Lobo and Maisongrande, 2006), curve fitting methods (Fischer, 1994a,b; Jönsson and Eklundh, 2002; Zhang et al., 2003) as well as methods of harmonic analysis and wavelet analysis (Li and Kafatos, 2000; Sakamoto et al., 2005).

Models for the analysis of time series of remotely sensed vegetation indices may be constructed for various reasons, stressing different aspects of a given problem. The objectives for fitting models may be (Dierckx, 1993):

Parametric representation: If the model function $y = f(t)$ is meant to directly represent a certain physical process, part or all of the function's parameters have a specific physical meaning. The objective of modelling then is to estimate these parameters as accurately and precisely as possible from a given set of data.

Functional representation: A model may represent a discrete set of data points (t_i, y_i) by a function $y = f(t)$ that can be evaluated at any suitable value of t within the domain of representation. Functional representation is important if a model is to be used for further analysis and characterisation of a data set. A good functional representation should provide not only the possibility to evaluate the model at arbitrary values of t , but also to calculate derivatives, definite integrals, and to solve the inverse problem $t = f^{-1}(y)$.

Noise reduction: In a given dataset $\{(t_i, y_i)\}_{i=1}^n$, the measured values y_i may be accurate enough to define $f(t)$ as an interpolating function such that $y_i = f(t_i)$, $i = 1, \dots, n$. Most typically however, the values y_i will be subject to measurement errors ϵ_i that are not negligible in a certain context. In this case, $f(t)$ should be defined as an approximating function yielding conditional mean values of y that depend on t in order to reduce the noise in the observations. If $f(t)$ is a suitable model for a given set of data, the modelled values $f(t_i)$ will be more accurate than the original data y_i most of the time.

Data compression: The purpose of a model may be to reconstruct a given data set from a number of parameters, where the set of fitted model parameters is considerably smaller than the original data set. An efficient model for data compression should represent a given data set as well as possible in as few parameters as possible.

The different aspects are generally related. While in a parametric representation the parameters themselves are the important part, in a functional representation the parameters are considered immaterial and the abstract mathematical properties of the function become important. Noise reduction is a common concern in all modelling applications. Physical parameters are less accurate if they are estimated from a set of noisy data, and if a functional representation of a data set is desired, it is expected to represent the true features of the data apart from noise. The same is true for a compressed version of a data set. In some models, noise reduction may be a part of the optimisation problem solved to obtain the model's parameters or coefficients, while other models do not include the capability of reducing noise in a data set. In the latter case, it is likely that the data will have to be preprocessed by a dedicated noise reduction method before they can be used for modelling.

2.1 Digital filtering

Many models for noise reduction of time dependent signals such as remotely sensed time series are based on the concept of digital filters. Filtering converts a given discrete time series $\{x_i\}$ into another discrete time series $\{y_i\}$ by an operation defined by the type and size of the filter. A linear digital filter, for example, is defined as a linear function of its input (Hamming, 1989). Fundamentally, there are two different types of digital filters, addressed by different nomenclature in different scientific disciplines.

The simplest filters are *nonrecursive* filters (Hamming, 1989); they are of the general form:

$$y_i = \sum_{k=-n}^n c_k x_{i-k} \quad (2.1)$$

The sequence of values from a usually short segment of a discrete signal that gives the current input of a filter is called the filter's *window*, n is called the *half-width* of the window. The filtered signal is generated by replacing the value y_i at the current position by a linear combination of the data points inside the filter window and the $2n + 1$ filter coefficients. The total width $2n + 1$ of the coefficient or filter window corresponds to the number of data points considered by the filter at any one time and is termed the *span* or *size* of the filter. A simple unweighted averaging filter would have constant coefficients $c_k = 1/(2n + 1)$, $k = 1, \dots, (2n + 1)$. The process of running the filter window over a discrete signal one sample at a time to compute a filtered output signal is termed *convolution* (Smith, 2003). Nonrecursive filters are called *moving-average* (MA) filters by statisticians, since the output y_i is a $(2n + 1)$ -point moving average of the input. Alternatively, they may be called *finite impulse response* (FIR) filters, because the output of the filter stops when the input stops. The moving average filter defined by (2.1) uses a symmetric filter window, where the anchor point, the data point that is replaced by the filter's output, is the central point of the filter window. In principle, any point inside the current

filter window may be used as the anchor point. In this general case, the left half width n_L of the filter is different from its right half width n_R , and the size of the filter is $n_L + n_R + 1$. For example, a time delay moving average filter has $n_R = 0$ and replaces the outmost sample to the end of the filter window by a weighted average of past signal values. Filters with $n_R = 0$ are termed *causal filters*. By comparing a raw signal with its time delayed moving average, Reed et al. (1994) and Archibald and Scholes (2007) were able to identify points along the time series where the signal values showed a rapid, steady increase. These points were then used to determine the time of onset of vegetation development.

While nonrecursive filters depend only on their current input signal, a *recursive filter* uses not only its input data x to compute its output y , but also previous values of its output. Its general form is:

$$y_i = \sum_{k=0}^n c_k x_{i-k} + \sum_{k=1}^m d_k y_{i-k} \quad (2.2)$$

The first sum in equation 2.2 is a moving average, the second sum is an *auto-regressive* (AR) filter in statistical terms, since the output is some regression of its m past values. Together, the two parts of equation 2.2 form an *auto-regressive moving-average* (ARMA) model. Autoregressive models alternatively are called *infinite impulse response* (IIR) filters in the literature, as their output can continue after the input has stopped. Although the filter may continue to respond indefinitely even when the input has ceased, the quality of its response is usually decaying rather quickly. If a time series is nonstationary, its trend can be removed by differencing the input to some order before using the model (2.2). Using first order differences will remove a linear trend, second order differences remove a quadratic polynomial trend, and so forth (Gershenfeld, 1999; Warner, 1998). Using such a differenced time series as input x to (2.2) gives an *auto-regressive integrated moving-average* (ARIMA) model. The regression methods provided by ARMA and ARIMA recursive filters are used in the statistical analysis of time series when the emphasis is on predicting future values for stationary and nonstationary stochastic processes over short terms (Shumway and Stoffer, 2011).

Nonrecursive filters are very commonly used to smooth out noise in time series generated from satellite observations. Moving window averaging with constant filter coefficients $c_k = 1/(2n + 1)$ is unbiased only up to and including the first moment (Press et al., 1992). For any signal with a nonzero second derivative, a bias is introduced by applying a moving average filter with constant coefficients to it. The Savitzky-Golay filter (Savitzky and Golay, 1964) is a linear filter designed to preserve higher moments. It approximates a given signal inside its window by a higher order polynomial. Applying a Savitzky-Golay filter of half width n and degree r is equivalent to least squares fitting an r th degree polynomial to the $2n + 1$ data points inside the filter window. Since a polynomial is a linear function in its coefficients, this operation can be expressed as a linear filter whose

coefficients are determined from a least squares fit of a polynomial to a dummy signal, usually a sequence $0, \dots, 0, 1, 0, \dots, 0$ of $2n$ zeros and a single 1. All other polynomials of the same degree can then be formed from linear combinations of these coefficients. Chen et al. (2004) used Savitzky-Golay filtering in an iterative procedure to smooth out noise in NDVI time series, following the upper envelope of the input NDVI signal. Their iterative use of the Savitzky-Golay filter is motivated by the assumption that higher NDVI values are generally more reliable than low values, since most of the relevant error sources tend to decrease NDVI values. Maximum value composites of NDVI observations (Holben, 1986) take advantage of the same concept. Chen et al. (2004) concluded that their method could outperform a nonlinear filter by Viovy et al. (1992) and a Fourier-based fitting method.

Among the nonlinear filters are the best index slope extraction (BISE) filter of Viovy et al. (1992) and smoothers that are based on running medians rather than running means as proposed by Tukey (1977) and Velleman (1980). Running median filters are motivated by the need for robust smoothing methods for data exhibiting occasional spikes or noise generated from long-tailed distributions. Nonlinear filters based on running medians may be more resistant if noise is characterised by extreme observations, without at the same time losing their ability to respond to rapid changes in the signal that are part of a true pattern. However, while removing extremely low observations that arise from a contamination of the signal e.g. by clouds, running median filtering may have the undesired effect of also lowering high NDVI values which may be presumed to be valid. Viovy et al.'s (1992) BISE filter was designed to retain high values and eliminate only extremely low observations.

Van Dijk (1987) concluded from a comparison of different running median filters that the "4253H, twice" filter of Velleman (1980) was the best smoother in two study cases dealing with ratio vegetation index data from agricultural areas. More recently, Hird and McDermid (2009) confirmed the good performance of this filter for normalised difference vegetation index data. From their comparison of different smoothing techniques, Hird and McDermid (2009) found that the ability of the Savitzky-Golay filter to reduce overall noise in an NDVI time series was greater in comparison to the "4253H, twice" filter. The "4253H, twice" filter, on the other hand, proved better able to maintain the integrity of the time series in terms of phenological features, a result that is consistent with the motivation and design considerations for running median filters.

Smoothing filters provide efficient means to mitigate the effect of residual noise in time series of vegetation indices, but at a cost of eventually blurring significant phenological features. The type and size of a filter along with the strength and character of the residual noise must be taken into account carefully if one intends to use the filtered time series to extract phenological metrics (Hird and McDermid, 2009).

2.2 Frequency domain methods

Instead of analysing the signal provided by a time series in the time domain, it can be represented by a number of frequency components to focus on the analysis of its cyclic behaviour. As phenologies typically exhibit strong seasonal cycles, frequency domain methods are used extensively and have long been established as a standard approach. Fourier analysis provides a consistent and replicable procedure to decompose a time-dependent signal into its frequency components by using the Fourier transform. The parameters calculated in Fourier analysis have a clearly defined mathematical and physical meaning, which facilitates further interpretation and analysis (Menenti et al., 1993). The usefulness of these methods in assessing vegetation dynamics has been pointed out by many researchers (Menenti et al., 1993; Olsson and Eklundh, 1994; Azzali and Menenti, 2000; Jakubauskas et al., 2001, 2002; Moody and Johnson, 2001). Parameters derived from frequency domain analysis of remotely sensed time series have been used in land cover classification (Andres et al., 1994; Moody and Johnson, 2001; Wagenseil and Samimi, 2006) as well as to describe extreme points and rates of change in phenological cycles using the derivative of the Fourier series (Olsson and Eklundh, 1994). Roerink et al. (2000) used interpolated NDVI values from a harmonic model to fill data gaps in a time series caused by cloud affected AVHRR observations over Europe. Jakubauskas et al. (2002) developed an index for landscape variability based on harmonic analysis of NDVI time series. Using a set of elaborate measures derived from a Fourier analysis of an NDVI time series, Evans and Geerken (2006) were able to infer the dominant vegetation type and its fractional coverage in an arid rangeland from remotely sensed observations. De Jong et al (2011) used harmonic models to detect trends in phenological parameters on a global scale. Analysis methods in the frequency domain can be roughly organised into *harmonic analysis*, *Fourier* or *periodogram analysis*, and *spectral analysis* (Warner, 1998; Bloomfield, 2000).

Harmonic analysis (Warner, 1998) is used if the true lengths of the cycles present in a time series are known or can be inferred from appropriate information. For example, there may be an obvious cycle length that can be seen by visual examination of a given time series, or the cycles and their lengths may be prescribed by the properties of certain types of vegetation under observation. A stationary time series with a single cycle of length τ can be represented by the following equation:

$$y_t = A + B \cos\left(\frac{2\pi t}{\tau}\right) + C \sin\left(\frac{2\pi t}{\tau}\right) + \epsilon_t \quad (2.3)$$

where the parameter A describes the mean level of the values y_t in the time series. The frequency of the sinusoid in radians ω is related to the cycle length, or *period*, by the relation $\omega = (2\pi/\tau)$. The modelled signal is obtained by evaluating the sine and cosine functions of period τ at suitable values of t . At any value of t , the harmonic model re-

produces the original signal with a residual error of ϵ_t . A sine and cosine function of the same frequency are orthogonal to each other and form a basis that can represent any sinusoidal waveform of a given frequency or period. Any sinusoidal waveform with a particular amplitude and phase can be generated by varying the coefficients B and C . The amplitude r of the generated waveform is a function of its coefficients:

$$r = \sqrt{B^2 + C^2} \quad (2.4)$$

The waveform's phase ϕ depends on the relative size of the coefficient C with respect to B :

$$\tan(\phi) = \frac{-C}{B}$$

Thus, an alternative way of formulating (2.3) is by characterising a sinusoid through its amplitude and phase

$$y_t = A + r \cos(\omega t + \phi) + \epsilon_t$$

The latter formulation is commonly used in the interpretation of waveforms, while (2.3) is used in most computations, since it explicitly contains the coefficients B and C associated with the orthogonal sine and cosine parts of the waveform. In (2.3), the resulting sinusoidal waveform is created as the sum of two components: a sine of period τ , and a cosine of period τ . Harmonic analysis can be seen as a type of regression analysis where the predictor variables are sines and cosines of varying frequencies, and the predictive coefficients are the linear coefficients B_k and C_k associated with each sinusoid (Warner, 1998). Harmonic analysis can be used to accommodate a number of m different cyclic components in a time series by using so called Fourier polynomials:

$$y_t = f_m(t) = A + \sum_{k=1}^m [B_k \cos(\omega_k t) + C_k \sin(\omega_k t)] + \epsilon_t \quad (2.5)$$

where the frequency of the k th sinusoid in radians is $\omega_k = (2\pi x_k)/\tau$. Typically, the k frequencies are chosen as multiples of the *fundamental frequency* $(2\pi)/\tau$ prescribed by the number of observations in the time series. This means that any of the m sinusoids completes a number of x_k cycles over a period τ , where the k are integers, $k = 1, \dots, m$. Likewise, the period of each sinusoid is a fraction of the signal's *fundamental period* τ . Sinusoids at frequencies that are integer multiples of a fundamental frequency are referred to as *harmonics* of the waveform at the fundamental frequency. Despite its name, harmonic analysis does not generally require the use of harmonics in a Fourier polynomial, one may also employ frequencies that are not integer multiples of the fundamental frequency (Hermance, 2007).

Since frequency domain methods deal only with cyclic components of a given signal, it is usually necessary to identify and remove any trends from a time series before it can be examined in the frequency domain. Trend removal may be done by differencing the

time series or by applying a suitable regression model (Warner, 1998). Frequency domain analysis is then carried out on the differenced time series or the residuals obtained from the regression model. It is also possible to simply incorporate harmonic components into a combined regression model that accounts for an overall trend as well as the cyclic components that characterise a signal (Jönsson and Eklundh, 2002; Hermance, 2007; Neuenschwander and Crews, 2008).

Whenever the number of cycles and corresponding lengths is not known in advance, these informations have to be inferred from a periodogram or power spectrum of a time series (Roerink et al., 2000). A periodogram is obtained by computing the Discrete Fourier transform (DFT) of a signal. The DFT of any digital signal can be calculated efficiently by an algorithm called the fast Fourier transform, or FFT (Cooley and Tukey, 1965). The discrete Fourier transform of a time series of n observations, where n is even, is given by

$$y_i = \sum_{k=0}^{n/2} \left[B_k \cos\left(\frac{2\pi k i}{n}\right) + C_k \sin\left(\frac{2\pi k i}{n}\right) \right] \quad (2.6)$$

and it decomposes the time series into a constant term plus $(n/2)$ harmonic frequency components. For example, a sinusoid composed of a sine and cosine of frequency $(2\pi k)/n$ may be called the k th harmonic, or harmonic of k th order. The Fourier coefficients B_k and C_k , for the individual components can be computed explicitly (Smith, 2003). Thus, the DFT is a lossless transform: the time domain signal can be restored exactly from its Fourier coefficients by taking the inverse of the DFT. This is different from harmonic analysis, where the regression coefficients associated with each sinusoid have to be computed by solving a system of linear equations in the least squares sense, and the signal is represented by the regression model only up to a residual error (Roerink et al., 2000; Hermance, 2007). The periodogram is a plot of the amplitudes (2.4) of all the harmonic frequency components that make up a given signal. Since a periodogram has only a discrete number of $(n/2)$ ordinates or frequency bins, periodogram analysis implies that a time series sampled at equal intervals in time is composed of a number of frequency components with periods that are integer fractions of the fundamental period n . In a statistical sense, periodogram analysis is similar to an analysis of variance that partitions the total variance present in a signal among a set of sinusoids of different frequencies. It can be used as a means of investigating if a time series has frequency components that explain a large proportion of its variance, and if so, obtain estimates of the phase and amplitude of the sinusoids at these frequencies (Warner, 1998). However, there is the implication that the frequency content of the signal under investigation is strictly harmonic. If a signal is created from an underlying waveform that is not a harmonic with respect to the fundamental period of observation, *leakage* may occur. Since periodogram analysis can only describe the cyclic components of a signal in terms of integer multiples of a fundamental frequency, it is unable to accommodate cycles with frequencies

that are fractions of the fundamental frequency. Thus, the variance due to a fractional cycle that cannot be accounted for spills over, or leaks into adjacent frequency bins that correspond to harmonic frequencies. Leakage need not be considered a serious issue if a problem at hand allows for fairly wide frequency bands (Warner, 1998). If leakage poses a problem, the trick of *zero padding* (Smith, 2003) used to narrow the frequency bins by increasing the record length may mitigate this problem, but does not generally solve it. In zero padding, the length of a signal is increased artificially by appending a number of zero observations at its ends in order to modify the set of frequency bands to obtain a more favourable partitioning. Zero padding is a valid trick since the record length n of a time series that determines the fundamental period is arbitrary and does not reflect a property of the signal itself. Therefore, it may be altered without changing the nature of the signal. Modification of the Fourier frequencies by altering the fundamental period does not, however, alter the amount of information contained in each frequency band. Although the number of periodogram ordinates ($n/2$) increases with the record length n , the estimate of variance associated with each frequency band is still based on only two degrees of freedom (Warner, 1998).

Spectral analysis techniques represent an alternative approach to improve periodogram estimates, as opposed to merely increasing the record length artificially. A periodogram enhanced by spectral analysis techniques is called a *power spectrum* and is essentially a continuous version of a periodogram. Instead of partitioning the variance among a discrete number of frequency bins, it provides a statistical estimate of the power, or variance accounted for, across the entire frequency domain of a given signal. Under certain circumstances, a power spectrum gives a better accurate and more reliable impression of the distribution of power among the frequency components of a signal (Warner, 1998). A power spectrum can be obtained in one of many possible ways. Power spectrum estimators are usually expressed in the form of digital filters operating on a raw periodogram. One of the simplest estimators is the Daniell or boxcar spectral window, which is the integral average of the periodogram over a number of frequency bins corresponding to the window width (Bloomfield, 2000). Each estimator has its specific favourable and unfavourable properties that must be considered in its application to a given dataset. Most estimators are smoothing filters that form weighted averages of adjacent periodogram ordinates. In general, when the smoothing that is applied to the raw periodogram gets heavier and heavier, it becomes more and more difficult to distinguish the contributions of neighboring frequency bands to the overall signal and the spectrum does no longer correspond to a fair partition of the total variance present in a time series (Warner, 1998).

Usually, the aim of periodogram or power spectral analysis is to identify a range or set of frequencies that reflect some of the properties of the physical processes effective in the creation of a given time series. Subsequently, a harmonic model is formulated for the time series according to (2.5) for further analysis. For obvious reasons, harmonic models

are most suitable for signals with a strong cyclic behaviour, where most of the variance is concentrated in only a few frequency bands visible as distinct peaks in the signal's power spectrum. Time series with a flat power spectrum approaching that of white noise, where the variance is partitioned equally among all frequencies, do not lend themselves well to spectral analysis. To differentiate the frequency components that contribute a significant amount of variance to a signal from those that can be considered noise, statistical significance tests have been developed for periodogram ordinates (Fisher, 1929; Warner, 1998; Bloomfield, 2000). The individual frequencies or frequency bands to be incorporated into a harmonic model can be chosen specifically according to the amount of explained variance of the overall signal (e.g. Roerink et al., 2000).

Regarding Earth observation time series, an appropriate harmonic, periodogram or spectral analysis responds to the systematic changes in the temporal profile that are caused by changes in vegetation cover and vigour, while being relatively insensitive to random fluctuations of signal noise (Moody and Johnson, 2001). Since high frequency components of signals typically produced by noise are restricted to higher order terms in a Fourier series, a harmonic model with reduced sensitivity to noise may be built by dropping higher order harmonics from the series. In some applications, as much as two harmonics of an annual time series may contain sufficient phenological information (Olsson and Eklundh, 1994; Azzali and Menenti, 2000), while in other cases discarding even higher order terms might blur significant features present in the original data. Geerken et al. (2005) report that at least five harmonics of an NDVI time series of one year observed by a SPOT VEGETATION instrument are needed to account for the subtle phenological differences in semiarid rangeland vegetation. Generally, the inclusion of higher order terms is problematic, as fitting higher order Fourier coefficients tends to generate spurious oscillations in the modelled time series (Chen et al., 2004). As an alternative to simply truncating a Fourier series, the terms including noisy frequencies can be down-weighted or their amplitudes tapered to ameliorate erroneous contributions of individual harmonics to the overall signal (Hermance, 2007). Geerken et al. (2005) used frequency domain filtering of individual Fourier components to remove noise before the complete signal model was generated by summing the terms for the individual components. In either case, the ability of the modelled time series to resolve phenological features such as beginning, end and duration of multiple green cycles diminishes (Jönsson and Eklundh, 2002; Chen et al., 2004).

Artificial frequency components may also be found in lower harmonics, where they are triggered by different values at both ends of a recorded time series. An assumption implicit in Fourier analysis is that an observed time series record of finite length is only a part of a strictly periodic, ever repeating infinite signal. Any violation of the assumption of periodicity by different data values at the beginning and end of a finite observational record is compensated by the introduction of erroneous frequency components. In the

worst case, a significant trend in the time series that has not been eliminated prior to Fourier analysis is subsumed in a number of low order harmonics. Jönsson and Eklundh (2002) and Chen et al. (2004) argued that the dependence of Fourier analysis on symmetric sine and cosine functions may lead to large displacements between the original and modelled time series when the original data are dominated by asymmetric features, such as a steep incline followed by a distinct phenological peak and a slowly decaying shoulder. To improve regression estimates of higher order Fourier coefficients, Sellers et al. (1994) proposed a weighted least squares approach to fit a harmonic model. Their approach focuses on the usually negative bias of NDVI data and is similar to the iterative scheme Chen et al. (2004) later used to improve Savitzky-Golay filtering of NDVI time series. Sellers et al. (1994) showed that a more robust fit of a Fourier series in the presence of erroneous or missing data values can be achieved when the contribution of individual data points to the fit is weighted in favour of higher NDVI values.

Notwithstanding the many hindrances one may encounter, Fourier polynomials are very flexible and can provide good parametric and functional representations for remotely sensed time series of vegetation indices. The clearly defined meaning of harmonic model parameters in terms of amplitudes and phases of sinusoidal waveforms at different frequencies establishes a strong physical basis and facilitates the assessment and comparison of models of this kind. The coefficients of harmonic models are obtained in a straightforward way by linear least squares procedures or can even be calculated explicitly using the discrete Fourier transform. In addition, sinusoids are functional representations that are extremely useful in the further analysis of a signal, since, for example, the derivatives of sines and cosines are easily calculated (Olsson and Eklundh, 1994; Evans and Geerken, 2006).

2.3 Fitting of exponential and logistic functions

Not least to overcome the various difficulties encountered in modelling remotely sensed time series based on spectral analysis, it has been proposed to locally fit sigmoidal or asymmetric Gaussian functions to a time series as an alternative to using the Fourier transform. Sigmoidal forms have been used to describe simple phenologies with single maxima during the growing period and single minima during the dormant phase (Fischer, 1994a,b; Fisher et al., 2006) as well as phenologies with two annual maxima and minima (Zhang et al., 2003). Jönsson and Eklundh (2002) used piecewise asymmetric Gaussian functions that are combined to a global function able to follow a complex time series curve with multiple maxima and minima. Both classes of functions may reproduce asymmetric features more accurately than a low order Fourier series and can also be applied in the case of non-uniform sampling in time.

However, to assess complicated phenologies that possibly exhibit multiple peaks and min-

ima in photosynthetic activity during a single year, multiple sets of such local forms have to be assembled using nonlinear least squares fits with a large number of free parameters. Fitting of a model with a large number of degrees of freedom can easily become unstable when the quality of the data to be fitted is limited. Sophisticated nonlinear optimisation techniques such as the quasi-Newton method used by Jönsson and Eklundh (2002) are computationally involved (Nocedal and Wright, 2006). The number of costly nonlinear fits increases with the number of years, i.e. annual or semiannual seasons, covered by the data. Furthermore, the convergence properties of an iterative nonlinear optimiser depend on the initial parameters supplied to the method as a starting point for optimisation. If the set of start parameters is not reasonable, it may take a long time for the method to converge, or even worse, it does not converge at all. Jönsson and Eklundh (2002) address this problem by selecting parameters from a set of allowable combinations using a grid search procedure. To determine the center locations for the local Gaussian functions, a low order detrended Fourier series is fitted to the raw data to identify the approximate locations of the annual and possibly semiannual peaks. These locations are then used as the centers in a sequence of adjacent local Gaussian models. In a final step, the local models are merged into a global function to provide a smooth and continuous representation of the overall time series (Jönsson and Eklundh, 2002). Because a frequency domain model is used in the determination of local reference points for the Gaussians, a time series of at least three years is needed to properly fit a model composed of local Gaussian functions (Jönsson and Eklundh, 2004). Like Sellers et al. (1994) and Chen et al. (2004), Jönsson and Eklundh (2002) concentrate on the analysis of NDVI time series and propose an iterative scheme for the determination of the parameters to have the global model track the upper envelope of the time series data.

Why all this mathematically involved and computationally expensive nonlinear optimisation? Models based on logistic or Gaussian functions anticipate a certain shape for the signal present in a time series. Whenever this anticipated shape is well suited to represent the biophysical realities of a given phenological signal, then it is certainly worth every effort to find a set of coefficients that fit such a function to a set of data. It has been shown that, under certain assumptions, part of the parameters and properties of logistic functions can be assigned a specific physical meaning. The general form of a logistic function is

$$f(t) = \frac{a}{1 + be^{-ct}}$$

where the parameter a sets the function's asymptotic maximum, b is related to the asymptotic minimum and c affects the slope of the logistic curve. Using results from the theory of crop reflectance modelling, Badhwar (1984) argued that a single logistic function can be constructed with three parameters related to the date of crop emergence, rate of growth, and carrying capacity. Badhwar (1984) then fitted his model to a tasseled cap-like greenness index obtained from a number of Landsat MSS observations over the

midwestern United States during the early part of a growing season. In the course of this, the growth rate was replaced by the rate of change in greenness, and the asymptotic maximum of greenness became a substitute for the carrying capacity. Based on these two variables, Badhwar (1984) was able to construct a linear classifier to discriminate corn and soybean fields.

Fischer (1994a; 1994b) successfully employed double logistic functions to describe the temporal profiles of NDVI computed from one year of ground based radiometric measurements over homogeneous crop canopies in a mid-latitude agricultural area (Beauce, France). Based on the fitted curve, advances or delays in crop growth were determined by the dates of the inflection points of the ascending and descending legs of the double logistic function. The length of the growing period was defined by the width of the curve, which is the time difference between the two inflection points. Fischer's (1994a) double logistic model of five parameters is essentially a generalisation of Badhwar's (1984) approach to accommodate a full growth cycle. Since a single logistic curve can only represent a single growth profile between asymptotic levels of sustained biomass and crop maturity, two three-parameter logistic functions are used to model the ascending and descending profiles of a full growing period separately, where one of the six parameters is eliminated by the requirement that the levels of maturity of both functions must match to yield a continuous overall profile.

Zhang et al. (2003) used double logistic functions to model an annual phenological signal with four anticipated key phenological transitions of greenup, maturity, senescence and dormancy. They found that the dates for these transitions corresponded well to the points on the time axis where the double logistic function showed its maximum rates of curvature. These results were then used to describe the annual vegetation dynamics over New England from a timeseries of one year of EVI derived from MODIS reflectance data at a spatial resolution of 1 km.

A favourable property of such *parametric* or *semi-parametric* (Fischer, 1994a) models is that, whenever some a priori knowledge is available, it is easily incorporated. If a clear physical meaning can be attributed to a number of variables involved in a curve fit, it is always possible to assign fixed values to specific parameters according to ancillary information, e.g. sowing dates (Fischer, 1994a), thereby reducing the number of variables to optimise and the uncertainty in the fitted model. When a priori knowledge is exploited in a given problem, it can increase the accuracy of the remaining parameters that are determined by curve fitting (Fischer, 1994a).

The employment of a priori information to narrow down the choice of parameters in an optimisation problem is a common feature of all methods using logistic or Gaussian basis functions. In fact, such information is absolutely required to achieve a meaningful curve fit. If such knowledge cannot be obtained beforehand from an independent source as in

Fischer's (1994a) work, it has to be inferred from a structural description of the time series itself. Zhang et al. (2003) used a moving window average filter to divide the signal into level parts between segments of increasing or decreasing EVI values. The segments were taken to identify the number of growth cycles in a given annual time series. Only then could they assemble a model for the entire time series from single logistic functions fitted to each segment of EVI increase or decrease, subject to the constraint of continuity of the curve between adjacent segments. Similarly, Jönsson and Eklundh (2002) used a Fourier-based decomposition to identify the number of growth cycles and associated local maxima present in a time series. The approximate local maxima associated with each growing period then marked the points along the curve where the individual Gaussian functions were blended together in order to obtain a smooth and continuous overall curve.

Except for the Gaussian and logistic curve fitting methods implemented in the computer program TIMESAT (Jönsson and Eklundh, 2004), which is intended for the analysis longer time series, logistic functions have been used mainly to analyse short time series covering only one or two growing periods of specific types of vegetation. Jönsson and Eklundh's (2004) TIMESAT software is intended for general use on remotely sensed time series of vegetation indices. However, TIMESAT does not use its model parameters directly to obtain phenological descriptors such as start, end and length of growing seasons. Instead, Jönsson and Eklundh (2004) apply moving window averaging to the modelled curve to obtain a consistent set of maxima and minima that represent cycles of vegetation growth and senescence. The fitted model is merely used to provide a smooth continuous curve that can be sampled at arbitrary positions, and the samples analysed by different methods of signal processing and time series analysis. As a consequence, the Gaussian and logistic functions employed in TIMESAT lose their significance as physically based parametric representations of remotely sensed time series, which is the very motivation for using these types of functions in the first place (Badhwar, 1984; Fischer, 1994a,b; Zhang et al., 2003).

2.4 Polynomial functions and splines

Polynomials are widely used to interpolate or approximate functions because of their favourable functional properties. A polynomial of order k in one variable is defined by k polynomial coefficients as follows:

$$y_t = f_k(t, \mathbf{c}) = c_1 + c_2 t + c_3 t^2 + \dots + c_k t^{k-1} \quad (2.7)$$

where $\mathbf{c} = (c_1, c_2, c_3, \dots, c_k)$ is a vector of polynomial coefficients. Alternatively, a polynomial may be specified by its degree ($k - 1$) which is the exponent of the largest term in the polynomial. The coefficient c_k associated with the largest term is called the *leading coefficient*. Among other things, polynomial models allow the general calculation of derivatives

and integrals of a modelled curve directly from the model itself. In addition, a polynomial is a linear function in its coefficients and can be fitted to a set of data straightforwardly by solving a system of linear equations, applying least squares methods when necessary. Savitzky and Golay (1964) used these properties in order to express smoothing based on local polynomials as a linear filtering operation, as discussed in chapter 2.1. Although a polynomial of a high enough order can in principle accommodate any curve shape, single polynomial functions are not suited for modelling time series of arbitrary length, since they tend to oscillate (Epperson, 2007), which leads to unacceptably large displacements between the modelled curve and the actual data. Nevertheless, polynomial functions of low order, or more generally, regression models including polynomial terms, may be used to describe nonlinear trends in time series (Wood, 2006; Hermance, 2007).

The idea of polynomial splines is to avoid displacements in polynomial models by fitting single polynomials to smaller pieces of a set of data, and join the polynomial pieces to a continuous curve at certain locations called *knots*. In general, some of the models discussed in chapter 2.3 may be understood as Gaussian or logistic splines, since they consist of pieces of Gaussian or logistic functions that are combined under certain continuity conditions to form an overall curve. In a narrower sense, the term *spline* refers to piecewise curves composed of polynomials, where the continuity conditions at the knots require not only the continuity of the curve itself, but also the continuity of a number of its derivatives (Wold, 1974; Dierckx, 1993; de Boor, 2001).

A (polynomial) spline is characterised by the degree of its polynomial pieces and the number of its knots. The number of polynomial pieces in a spline is one less than the number of its knots. Splines can be defined as *interpolating splines* or *approximating splines*. An interpolating spline is a “join the dots” type of model (Faraway, 2006) where the knots of the spline correspond to the data points in a given dataset, so that the spline’s coefficients can be determined in order to interpolate the given values at all points. *Regression splines* and *smoothing splines*, in contrast, are approximating spline models that are not required to reproduce all given values in a dataset. Regression splines differ from smoothing splines in the number of knots used. For regression splines, the number of knots is usually much smaller than the number of data points. The level of detail captured by a regression spline is controlled by the number of piecewise fits to the data. A lower number of knots means less piecewise fits and thus less detail in the fitted curve. Varying the grid of knots, i.e. the number of knots and their placement, is an essential concept to gain control over the fit of a regression spline to a given dataset. One may start, for example, with a fine grid of knots and then drop knots sequentially until the resulting curve provides an acceptable model for the data. In a smoothing spline, the number of knots is roughly of the same order as the number of data points, and the number of degrees of freedom of the spline model is controlled by a regularisation term that penalises large values for the spline coefficients (Yandell, 1993; Faraway, 2006; Takezawa, 2006; Hastie

et al., 2009). The degrees of freedom in a given spline model may also be controlled by varying the number of knots and applying a smoothing penalty at the same time, e.g. by choosing the number of knots slightly larger than needed and use regularisation to exert fine control on the level of detail in the fitted spline (Eilers and Marx, 1996). A spline model where both the number of knots and a penalty term are used to adjust the degrees of freedom may be called a *penalised regression spline* (Ruppert, 2002). All these spline models may be employed in the analysis of remotely sensed time series.

Hermance et al. (2007) together with Bradley et al. (2007) used spline models to assess the phenology of prairie vegetation in the Great Basin on the basis of NDVI time series derived from AVHRR data. They used splines of an order of 11-14 with knots placed at the junctions of annual subintervals in a time series of multiple years. Using very high order polynomials over the annual subintervals, Bradley et al. (2007) had to use an iterative weighting scheme that resembles those of Sellers et al. (1994), Chen et al. (2004) and Jönsson and Eklundh (2002) to achieve stability of fit. Initial weights were computed using an exponential weighing function with respect to an average annual curve determined by a polynomial regression model including harmonic components (Hermance, 2007; Hermance et al., 2007). By adjusting the initial weights in a number of recursive least squares fits, detailed high order spline models could be fitted to accommodate the upper envelope of multiyear NDVI time series (Hermance et al., 2007; Bradley et al., 2007). Typically, splines of order 4-6 are used in most applications and the number of knots is used to control the level of fit (Wold, 1974; Stone, 1986; Dierckx, 1993; Ruppert, 2002; Ruppert et al., 2003; Takezawa, 2006; Keele, 2008). Instead of using lower order splines and varying the number of knots to control the level of detail, Hermance et al. (2007) and Bradley et al. (2007) decided to cope with the numerical difficulties inherent in high order polynomials (Epperson, 2007) and increased the spline order to attain greater levels of detail while keeping the number of knots constant. Despite the fact that the behaviour of regression spline models tends to depend strongly on the locations chosen for the knots (Wood, 2006), knots are an abstract concept and the knot locations have no physical significance whatsoever. There is no point in insisting on knot locations at the splits of annual subintervals.

An advantage of the spline is that it is a entirely data driven curve. Unlike the logistic and Gaussian functions in chapter 2.3, splines are nonparametric³ models and their fit is not limited by any constraints imposed by the shape of a certain model function. A spline model can be designed specifically to meet the requirements associated with the characteristics, amount and quality of the observations in a given dataset. If a time series originates from a system with a simple phenology and contains a strong signal, for

³That is to say, the parameters of a spline have no physical interpretation. But strictly, spline models *are* defined by parameters, and once a certain number of knots is chosen, the shape of the curve is constrained from the set of all possible splines to a subset of splines with a given number of knots at the specified positions. Nevertheless, knots are immaterial and their number and positions may be chosen freely.

example a homogeneous agricultural area, a regression spline with only a few knots may be sufficient. In contrast, the intimate mixture of vegetation types in savanna ecosystems and their very different response to the variable rainfall creates subtle phenological signals that call for more complicated spline models with a rich set of knots and a smoothing penalty to regulate the sensitivity of the model. Likewise, an interpolating spline may capture every detail in a small, high quality dataset. For a larger dataset with erroneous observations on the other hand, a regression or smoothing spline to estimate conditional mean values may be more appropriate. Whatever the appropriate spline model may be, once its parameters are determined, the representations for all these splines can be arranged to be the same (de Boor, 2001), although their individual properties may be quite different. Spline models are good functional representations for a wide range of data analysis problems. Derivatives and integrals of polynomials are easily calculated, and any algorithm for their differentiation and integration immediately applies to all spline models, since any spline may be represented in a piecewise polynomial form (de Boor, 2001). Thus, algorithms for phenological analysis can use a spline representation of a time series as a universal data structure to identify phenological cycles and assess their properties using extreme values and curvature obtained by direct evaluation of the spline model. A class of spline models termed interpolating B-splines even have representations as linear digital filters, in much the same way as interpolating polynomials can be expressed as Savitzky-Golay filters (Unser et al., 1993b; Unser, 1999).

A disadvantage of splines may be their susceptibility to overfitting a given dataset, since a spline with a large number of parameters can, in principle, accommodate any given shape. The spline-like models in chapter 2.3 composed of Gaussian and logistic functions, for example, use a fixed number of model parameters per phenological cycle. For a spline, the number of knots in a given segment of a time series may be chosen arbitrarily large, within certain limits set by the amount of data available. The number of knots covering a phenological cycle may even be different between individual inter-annual cycles. In order to avoid overfitting, care must be taken when choosing the number of knots, their positions, and the order of the polynomial pieces in a spline model.

2.5 Wavelets

The Fourier models presented in chapter 2.2 often suffer from the disadvantage that the sine and cosine functions used to describe the signal do not provide compact support. This means that the fit of each sine and cosine function that is a part of a harmonic model depends on the whole data set. The lack of compact support is one reason why spurious oscillations may be introduced into Fourier models when the recorded time series is not strictly periodic. Contrary to sinusoids, which are local in frequency only, wavelets (Strang, 1994; Aboufadel and Schlicker, 1999; Fugal, 2009) are functions that are localised in frequency and in time. Sinusoids are regular and particularly suited to

describe constant-frequency (stationary) signals (Fugal, 2009). Wavelets are irregular functions with compact support, i.e. of limited duration, and are better suited to analyse signals containing breaks and other nonstationary phenomena. There are many different kinds of wavelets designed for different purposes. Among the simplest wavelets is the Haar wavelet, which is good for detecting breaks in a signal (Fugal, 2009), while the Morlet wavelet is better suitable for analysing a signal's cyclic behaviour (Fligge et al., 1999). Another family of wavelets called Coiflets are adequate for fractal time series with a high degree of self-similarity (Fugal, 2009). An additional advantage of wavelets is their ability to look at time series, at multiple resolutions, which allows to focus on grosser features of the data without neglecting finer details when they are important (Percival and Walden, 2000). Wavelet transform methods have been used in remote sensing time series analysis for example to separate cyclic components of signals (Li and Kafatos, 2000), or to remove noise in time series (Sakamoto et al., 2005; Galford et al., 2008).

2.6 Linear decomposition

There exist a number of methods for the decomposition of a time series into a trend, cyclic components and a remainder that may be treated as residual error. In the analysis of remotely sensed time series of vegetation indices, such methods have been used to analyse phenology in ecosystems where the vegetation can be categorised into perennial and annual types associated with the trend and cyclic components, respectively. Roderick et al. (1999) and Lu et al. (2003) used linear decompositions of NDVI time series to assess Australian woody and herbaceous vegetation cover on a continental scale. Lu et al.'s (2003) method is based on a linear decomposition procedure called STL (Cleveland et al., 1990), where the trend is determined using a method of locally weighted regression, often shortened to lowess, or loess (Cleveland, 1994). Verbesselt et al. (2010) developed a similar method that uses a piecewise linear approximation of the trend component to accommodate certain breaks and points of change in trend. The trend and cyclic components present in remotely sensed time series may be described in a number of different ways, and any of the modelling approaches discussed above may be used as a part of a linear decomposition procedure. Thus, linear decompositions may be seen as a general framework for linear time series analysis rather than a group of models by themselves. Generalised additive models (Wood, 2006) may provide another framework for analysis of trend, seasonal and random components of time series. The expected value of the response of a generalised additive model is some monotonic function of a linear predictor specified as a sum of functions of a number of predictor variables. The (detrended) harmonic models discussed in chapter 2.2 as well as the spline models of chapter 2.4 may be placed in this category of models.

2.7 Summary: why splines?

As discussed at the beginning of this chapter, the four general objectives of curve fitting are the parametric or functional representation of discrete data sets, noise reduction and data compression. The nonrecursive filters discussed in chapter 2.1 are methods designed for reducing noise. A running mean or median filter simply replaces each and every value in a data set by a smoothed, noise reduced equivalent, without changing the general representation of the data as a set of points. A data set functionally represented by a spline may or may not be noise reduced, depending on whether the spline is an interpolating or an approximating spline. An interpolating spline evaluated at its knots yields the original data values, while an approximating spline smooths a data set. Most uses of logistic functions to assess short vegetation index time series over one or two growth cycles (Badhwar, 1984; Fischer, 1994a,b; Zhang et al., 2003; Fisher et al., 2006) are examples of the use of models as parametric representations of a data set, where all or part of the model parameters have a direct physical interpretation. Fisher et al. (2006) for example used the two inflection points of a double logistic function to determine the timing of the onset of greenness in the spring caused by leaf expansion and grass greenup, and the the offset of greenness in the fall due to leaf abscission and grass brown-off. Fourier based models and polynomial splines are examples of functional representations of remotely sensed time series. These models allow an arbitrary resampling of the original data set based on the model functions, as well as the analysis of the dynamics of the data based on the derivatives of the model functions. Because of their multiresolution properties, wavelets are efficient models when the problem involves compression of the data (Aboufadel and Schlicker, 1999).

Broadly speaking, parametric representations take a problem-specific perspective, while the perspective of functional representations is data-centred. Using a problem specific parametric model that anticipates a certain shape of the phenological curve makes most sense if the phenological behaviour can be assumed as known, e.g. for applications in monocultural farming, where the measured signal originates mostly from one crop, with only negligible other contributions (Fischer, 1994a,b). Double logistic functions seem particularly suited to model phenologies with long growing seasons where there is no distinct maximum in the observed phenological curve, but a plateau-like shape instead, enclosed by a relatively steep incline and decline at the beginning and end of the growing season, as well as a distinct dormant phase where the curve is flat and low. If a recorded satellite signal has significant contributions from multiple crop, seminatural or natural vegetation types, the shape of the phenological curve cannot generally be determined, since it depends on the strength of the contributions from the individual components. In such cases it may be more appropriate to adopt a data centred approach to curve fitting where the degrees of freedom of the fit can be adjusted according to the level of detail present in the signal.

In Fourier based models, where the signal is represented as a sum of sinusoids of different frequencies, both perspectives are amalgamated. Depending on the context, a Fourier series may be regarded as a data-centred functional representation as well as a problem-specific parametric representation of a time series. Provided the number of harmonics is large enough, a Fourier series can represent any given curve without restrictions on its shape (Smith, 2003). Furthermore, the amplitudes and phases of the sinusoids may also be interpreted as physical parameters, for example if the frequency of a waveform corresponds to a recurring vegetation cycle. Wagenseil and Samimi (2006) found from their analysis of NDVI time series of two consecutive growing periods in an area with high rainfall variability (Etosha National Park, Namibia), that Fourier analysis was useful, but did not quantify all relevant aspects of plant phenology and that the parameters of the Fourier series were altered considerably by the different rainfall situation in the two years. They concluded that amplitude and phase could well be interpreted in their physical relation to vegetation phenology, but the spatial and temporal variation caused by the differences in precipitation between the two years exceeded the intra-annual variability. This leads to limitations in the separability of vegetation entities based on Fourier coefficients, since information on intra-annual variability is used to discriminate the different vegetation types. Rather than using the absolute values of parameters derived from a time series, Geerken et al. (2005) and Evans and Geerken (2006) proposed methods for vegetation discrimination based on Fourier analysis that quantify the similarity between the annual cycle of an individual signal with a reference cycle, i.e. the average annual cycle of a signal known to originate from a specific vegetation type. The similarity measure of Evans and Geerken (2006), which is based solely on a number of Fourier coefficients, is explicitly designed to consider the various dynamics in a time series that could be caused by altered growing conditions between different years or growing periods, most important amplitude scaling, which would be the case if an identical vegetation type was growing with greater vigor or a more complete coverage, e.g. due to more favorable rainfall conditions, and amplitude translation, which would occur e.g. due to an earlier or later onset of a rainy season. Moreover, the algorithm allows the estimation of the fractional coverage for the vegetation type identified by the similarity measure on the basis of the derivative of the Fourier series. This allows Evans and Geerken (2006) to deal with the mixed pixel problem that typically arises from the low spatial resolution of satellite observations that otherwise provide the high temporal coverage required to discriminate vegetation types based on their phenology. The work of Evans and Geerken (2006) shows that a data centred perspective does not mean that problem specific methods cannot be developed. It is by virtue of the functional representation properties of Fourier models that Evans and Geerken (2006) were able to develop a problem oriented approach to semiquantitative land cover mapping.

Jönsson and Eklundh's (2002) motivation for using logistic and Gaussian functions is

not parametric representation, but providing alternative representations for phenological curves that are not well represented by Fourier based methods. Fourier based representations are less suitable for phenological curves that approach a rectangular shape, like the plateau-shaped phenologies encountered by Fischer (1994b) in the Beauce agricultural area, as well as asymmetric growth profiles, where the ascending and descending parts of the growth curve differ in shape (Fischer, 1994a; Jönsson and Eklundh, 2002; Chen et al., 2004). Box shaped or asymmetric features in a signal correspond to high frequency information in terms of a Fourier transform, and are generally accommodated by higher order harmonics that are susceptible to noise. If high frequencies contribute significantly to a signal, it is not possible to simply remove noise by truncating its Fourier series representation and it becomes difficult to separate the true frequency content from the frequencies that constitute noise. This renders the Fourier transform inefficient for these kinds of signals and they may be better represented by locally fitting logistic or Gaussian functions (Jönsson and Eklundh, 2002). However, logistic or Gaussian functions cannot be fitted to a time series without ancillary information on its structure. To determine approximate annual or semiannual maxima as the points of continuity where the local logistic or Gaussian fits are blended together, Jönsson and Eklundh (2002) still use a low order Fourier based model. Since Fourier models are regular in frequency, this implies that the middle of a growing period, represented by an individual local function, occurs at the same time every year. This is not always a realistic assumption, as for example in semi-arid regions, where plant growth may be driven spontaneously by rainfall (Tan et al., 2011). In addition, the assemblies of logistic or Gaussian-type models are not easily analysed, and yet other signal processing techniques such as moving window averaging have to be used to find a consistent set of minima and maxima as a basis for the determination of the phenological metrics (Jönsson and Eklundh, 2002, 2004).

Many of the reviewed methods for noise reduction and subsequent extraction of phenological parameters focus on normalised difference vegetation index time series and try to maintain the upper envelope of values to account for the negatively-biased noise in NDVI observations (Sellers et al., 1994; Chen et al., 2004; Jönsson and Eklundh, 2004; Hermance, 2007; Hermance et al., 2007). However, several alternatives to the NDVI have emerged over the years, such as the enhanced vegetation index (EVI, Huete et al., 2002) or the sensor independent fraction of absorbed photosynthetically active radiation (FAPAR, Fensholt et al., 2004). Both products are routinely derived e.g. from MODIS observations. EVI and FAPAR are less influenced by atmospheric and background effects than the NDVI, and may have less negatively-biased noise and more erroneous spikes compared to NDVI data, in which case denoising techniques maintaining the upper envelope of values may not be the most effective choice (Hird and McDermid, 2009).

A method for noise reduction, curve fitting and phenological characterisation of remotely sensed time series that is generally applicable across different ecosystems and to differ-

ent observed vegetation parameters (NDVI, EVI, FAPAR, etc.) must adopt a data-centred perspective. Polynomial splines are entirely data driven functional representations of ordered data sets such as time series. Splines avoid some of the troubles with other methods. Compared with methods based on fitting logistic or Gaussian functions, the knots of a spline may be positioned without ancilliary information about curve structure. The spline adjusts to the data just by a linear least squares fit, no positioning of local functions and no nonlinear search for parameters is required. Furthermore, in a spline model rapid changes in the curve may not be confused with noise as in a Fourier series, since a spline model is aware of the time domain only. Nevertheless, the Fourier coefficients of a modelled signal may be recovered from its spline representation (Dierckx, 1993), so if an analysis of the cyclic behaviour of a signal needs to be done, it can be done using the spline model rather than the signal itself to save preprocessing of the signal e.g. by filtering. Splines can be defined so as to possess the desirable property of compact support in the time domain, which means they have good smoothing and noise reduction capabilities and their use does not generally require preprocessing of the data. In addition to their noise reducing properties, splines are also excellent models for functional representation. Since the derivatives and integrals of spline curves are splines themselves, a spline curve can be analysed completely by computing higher derivatives or integrals recursively. This makes it straightforward to determine crucial points such as minima and maxima of phenological cycles, or to calculate definite integrals over growing periods directly from the model and independently of the sampling resolution of a data set. All these features render the spline a robust, general, self-contained and yet flexible model that can be used for noise reduction as well as characterisation of all kinds of remotely sensed time series (Hermance et al., 2007; Bradley et al., 2007).

3 A spline framework for remote sensing time series analysis

Polynomial splines (Wahba, 1990; Dierckx, 1993; de Boor, 2001) have been used to interpolate or approximate functions in one or more variables in computer graphics and computer aided geometric design (Foley et al., 1997; Rogers, 2001; Farin, 2002; Lengyel, 2004; Mosier, 2009), signal and image processing (Unser et al., 1993a,b) as well as in statistics and data mining (Ruppert et al., 2003; Takezawa, 2006; Wood, 2006; Keele, 2008). Splines are named after a tool used for drawing and fabricating smooth shapes and surfaces. The actual drawing tool called a spline is shown schematically in figure 3.1. It consists of a set of weights or *ducks* that are placed in order to force a flexible beam into a prescribed smooth shape. An abstract mathematical spline is a construct that imitates the workings of the tool that is called a spline by craftsmen. A confusing aspect of mathematical splines is that there exist many different types, such as regression splines, smoothing splines, B-splines, P-splines. The different types stem from different optimisation criteria, end point constraints and representations for spline models of curves and surfaces. A spline as a functional representation of a time series may be defined as a function in one variable, $y = f(t)$, where y is the value of a remotely sensed parameter, e.g. NDVI, at time t . An intuitive representation of splines is through piecewise polynomial functions.

3.1 Piecewise polynomials

A spline is a piecewise polynomial function defined on a finite interval from a to b . The abscissae along the curve where two polynomial pieces join are called *knots*. There always are at least two knots at a and b called *exterior* or *end knots*. The remaining knots are the *interior knots* of the spline. If there are only exterior and no interior knots, the spline is just a single polynomial. A spline of order k is assembled out of a number of polynomial

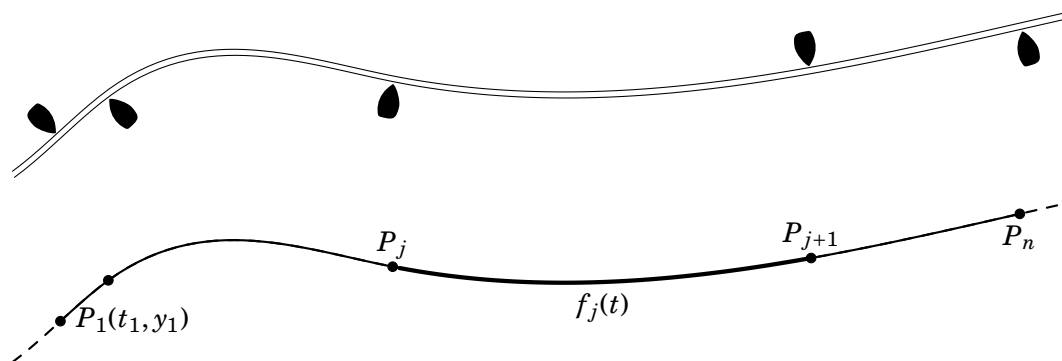


Figure 3.1: A craftsman's spline and its abstract equivalent, the natural cubic spline. The tool used to craft smooth continuous shapes consists of a flexible plastic strip and a number of massive weights, the ducks. The flexible beam is forced to pass along the points marked by the ducks. An abstract spline that closely imitates the tool's mechanics is called a natural spline.

pieces of degree $k - 1$. The individual pieces are blended together so as to make the resulting spline curve continuous. A spline of order k is continuous up to and including its $(k - 2)$ derivative. Let the knots of a spline be given by

$$a = x_1 \leq x_2 \leq \dots \leq x_n = b$$

then the piecewise polynomial spline function defined for this set of knots may be written as

$$f(t) = \begin{cases} f_1(t) & \text{if } x_1 \leq t \leq x_2 \\ f_2(t) & \text{if } x_2 \leq t \leq x_3 \\ \vdots & \vdots \\ f_{n-1}(t) & \text{if } x_{n-1} \leq t \leq x_n \end{cases} \quad (3.1)$$

where the f_j , $j = 1, \dots, (n - 1)$, are single polynomials of equal degree. In principle, the first and last polynomial pieces of $f(t)$ can be extended when $f(t)$ is to be defined for the entire set of real numbers (de Boor, 2001):

$$f(t) = \begin{cases} f_1(t) & \text{if } t < x_1 \\ \vdots & \vdots \\ f_{n-1}(t) & \text{if } t > x_{n-1} \end{cases}$$

Extending a piecewise polynomial in this way amounts to extrapolation. As with any extrapolation technique, spline extrapolation may reflect a modelled function only badly away from the original interval $[a, b]$ on which it was defined.

To construct a piecewise polynomial that interpolates a set of data points $\{(x_j, y_j)\}_{j=1}^n$, n knots are needed at locations x_j , $j = 1, \dots, n$, where $x_1 = a$ and $x_n = b$ are the exterior knots; with corresponding ordinates at the knots of

$$y_1 = f(a), y_2, \dots, y_n = f(b)$$

In general, for an interpolating spline of order k with n knots, $k(n - 1)$ polynomial coefficients have to be found. If for example, the individual polynomials are of degree 3 (order 4), the piecewise polynomial is a cubic spline and is determined by a total of $4n - 4$ polynomial coefficients, since every single fourth-order polynomial has four coefficients, and the spline is composed of $n - 1$ single polynomial pieces. The individual pieces are given by

$$f_j(t) = A_j^*(t - x_j)^3 + B_j^*(t - x_j)^2 + C_j^*(t - x_j) + D_j^* \quad (3.2)$$

where $x_j \leq t \leq x_{j+1}$ and $j = 1, 2, \dots, n - 1$. An interpolating spline must match the values of the abscissae at each interior knot. This is enforced by $k - 2$ interpolation equations of

the form

$$D_j^* = y_j \quad \text{for } j = 2, \dots, n-1 \quad (3.3)$$

In order for the cubic spline curve to be continuous up to the second derivative, another set of $3n - 6$ equations is required to enforce continuity across the interior knots. Let

$$h_j = x_{j+1} - x_j \quad \text{for } j = 1, \dots, n-1$$

then the equations ensuring continuity across the interior knots are

$$A_j^* h_j^3 + B_j^* h_j^2 + C_j^* h_j + D_j^* = D_{j+1}^* \quad (3.4a)$$

$$3A_j^* h_j^2 + 2B_j^* h_j + C_j^* = C_{j+1}^* \quad (3.4b)$$

$$3A_j^* h_j + B_j^* = B_{j+1}^* \quad (3.4c)$$

Every interior knot of the spline curve must satisfy these three equations. An additional four equations (k in general) are needed to solve for all the unknowns. These equations specify the boundary conditions at the two exterior knots. To interpolate not only the data values at the interior knots, but also the values at both ends, the following two boundary conditions are required:

$$D_1^* = f(a) \quad (3.5a)$$

$$A_{n-1}^* h_{n-1}^3 + B_{n-1}^* h_{n-1}^2 + C_{n-1}^* h_{n-1} + D_{n-1}^* = f(b) \quad (3.5b)$$

To be able to solve for all the $4k - 4$ unknowns, two more constraints must be put on both end knots by specifying two more equations. To be *natural spline*, a piecewise polynomial of order k , where k is even, must have polynomial pieces at the beginning and end with derivatives that are zero up to the r th derivative, where $k = 2r$ (Mosier, 2009). To define a natural cubic spline, the two additional constraints are specified to have the second derivatives at the end knots equal zero:

$$f''(a) = f''(b) = 0$$

This means that a natural spline is forced to behave linearly beyond its end knots. The spline is called natural because its behaviour at the ends most closely mimics the mechanics of a craftsman's spline (figure 3.1). If the straight plastic beam of an actual spline is relieved from the strain put to it by the ducks, it flexes back to a straight shape. Beyond the first and last duck, the beam would thus prescribe a linear shape just like its abstract counterpart, the natural spline does. Thus, the additional two constraints for a

natural cubic spline are given by the equations

$$B_1^* = 0 \quad (3.6a)$$

$$3A_{n-1}^* h_{n-1} + B_{n-1}^* = 0 \quad (3.6b)$$

A natural spline is always a spline of even order (Dierckx, 1993; Mosier, 2009). As discussed above, a cubic spline with n knots is defined by $4n - 4$ polynomial coefficients. If the spline is quadratic, i.e. of order three, then the number of unknowns is $3n - 3$. As with the cubic spline, the number of conditions needed for the interpolation of the values at the interior knots is $n - 2$, leaving $2n - 1$ unknowns to determine. In addition, there are $2n - 4$ conditions ensuring continuity of the spline function and its first derivative at the interior knots, leaving only $(2n - 1) - (2n - 4) = 3$ equations to specify the spline's behaviour at the end knots and beyond. Since the number of equations for the boundary conditions is odd, one of the ends inevitably has to go short by one condition. Since a natural spline requires to have an appropriate number of zero derivatives at both ends, it cannot be of even order. In regard of the fact that the end knots of an even order spline obey different boundary conditions, some authors recommend using only splines of even order (odd degree) in curve fitting applications (Dierckx, 1993; Mosier, 2009).

There are other possible formulations for the end point constraints of an interpolating cubic spline. Another common end constraint besides the natural spline condition is to specify the slope of the spline curve at the ends. If the slope of the curve to be modelled by a spline is known at both ends or can reasonably be estimated, then a cubic spline can be constructed whose first derivatives $f'(t)$ at the end points match the prescribed values. A boundary constraint that may be of particular interest in the analysis of remotely sensed time series is the constraint of a *periodic* spline. Time series of vegetation indices most often have a seasonal component that makes the underlying signal periodic in nature, so it may be desirable to make the spline used to model the signal periodic too. For a periodic cubic spline, the spline's values have to agree at both end knots up to their second derivative: $f(a) = f(b)$, $f'(a) = f'(b)$, $f''(a) = f''(b)$. A cubic piecewise polynomial as described above can be made periodic by substituting equations (3.6) with the following:

$$3A_1^* h_1^2 + 2B_1^* + C_1^* = 3A_{n-1}^* h_{n-1}^2 + 2B_{n-1}^* + C_{n-1}^* \quad (3.7a)$$

$$3A_1^* h + B_1^* = 3A_{n-1}^* h_{n-1} + B_{n-1}^* \quad (3.7b)$$

Note that $y_1 = y_n$ must be given in order for an interpolating piecewise polynomial to be periodic. If the data values y_1 and y_n at both ends do not match (3.5) cannot be satisfied and the interpolating spline cannot be periodic. If $y_1 \neq y_n$, the problem can only be solved in a mean sense by using an approximating spline that is not required to pass through all its knots.

3.2 Spline representations and basis functions

The three major decisions to ponder on when choosing a spline model are concerned with the degree of the spline's polynomial pieces, the locations of the knots where the individual polynomials are blended together, and what constraints to impose at the exterior knots. Along with these primary options, the secondary choice of a suitable representation for the spline curve must be considered as well. The form of representation does not generally influence the shape of the resulting spline curve. Representing splines in their piecewise polynomial form as a set $\{(A_j^*, B_j^*, C_j^*, D_j^*)\}_{j=1}^{n-1}$ of $n - 1$ tuples of polynomial coefficients is only one of a number of possible representations for splines. It is a useful representation since its coefficients can be directly interpreted as the coefficients of the $n - 1$ single polynomial pieces. Nevertheless, it is inconvenient for the formulation of curve fitting problems that do not have an exact solution and can only be solved in a mean sense. Another possibility is to use a set of *basis functions*. The motivation behind basis functions is to imitate single polynomials for the sake of easier curve fitting. Instead of using a set of coefficients associated with a number of monomials as in (2.7) to define a single polynomial, each coefficient is associated with a particular basis function:

$$f(t) = c_1 B_1(t) + c_2 B_2(t) + \cdots + c_m B_m(t)$$

In this way, a piecewise polynomial can be assembled as a linear combination of basis functions just like a single polynomial is a linear combination of monomials. Basis functions also avoid the need to specify the continuity conditions at the interior knots explicitly. The basis functions themselves are defined so as to ensure continuity of the spline function across the set of interior knots. Other constraints, such as the spline's behaviour at the end points, are fundamentally part of the spline model, but are subsumed as a characteristic of one or more members of a specific set of basis functions used to construct the spline model (Ruppert et al., 2003).

The type of representation for the spline used in computing the spline's coefficients and the representation used for working with the fitted spline model may be chosen independently of each other. The best representations for fitting, interpreting and evaluating a given spline model are not necessarily identical. In the piecewise polynomial representation, the interpretation of the model's coefficients is particularly clear and easy, but the model is fitted more efficiently using a different form of representation. The flexibility of spline models is partly a consequence of the fact that it is possible to convert between the different representations as necessary (de Boor, 2001).

When choosing a representation to be used in computing the spline's coefficients, the main concern is numerical stability – the basis functions should be chosen so as to reduce numerical errors as much as possible. Unlike, for example, the sine and cosine functions that form the basis functions in a Fourier transform, spline bases are not orthogo-

nal, which means that there are correlations among the different basis functions. These correlations can make it difficult to estimate the coefficients associated with each basis functions, as the contributions of the individual basis functions to the overall signal are blurred by the collinearity effects. There exist basis functions for spline models known as B-spline basis functions that are especially designed to reduce collinearity among the individual functions that form the basis (de Boor, 2001; Keele, 2008). Spline models resulting from this kind of basis are more numerically stable than, for example, models that use truncated power functions as basis functions. This is of importance especially if a spline model has a large number of knots and the least squares principle is used to solve for a set of spline coefficients in a mean sense. Both is true in modelling remotely sensed time series: to reduce noise present in the observed data, a model is usually fitted in a mean sense using least squares, and a considerable number of knots is required to obtain a good fit if the time series record is long. Moreover, unlike truncated power function representations, B-splines possess the desirable property of compact support. The influence of a single B-spline basis function extends only over a part of the entire spline curve. This allows to take into account specifically the local properties of the data when fitting a model. B-splines are thus the basis functions of choice for fitting spline models to remotely sensed time series. Once the coefficients for the model are obtained, it may be converted to whatever representation is most useful to tackle a given problem.

3.3 B-splines

Using B-splines, a spline function is defined as a continuous function that is a linear combination of basis functions $B_{j,k}$, where $B_{j,k}$ is the j th B-spline basis function of order k :

$$f(t) = \sum_{j=1}^m c_j B_{j,k}(t) \quad (3.8)$$

where $a \leq t \leq b$ is the supported domain and $2 \leq k \leq m$. The basis functions are defined by the following recursion formulas (de Boor, 1972). The recursion starts with the first order basis function which is either 1 if its argument is inside the j th knot interval, or 0 otherwise:

$$B_{j,1}(t) = \begin{cases} 1 & \text{if } x_j \leq t < x_{j+1} \\ 0 & \text{otherwise} \end{cases} \quad (3.9)$$

Higher order basis functions are computed as

$$B_{j,k}(t) = \frac{(t - x_j)B_{j,k-1}(t)}{x_{j+k-1} - x_j} + \frac{(x_{j+k} - t)B_{j+1,k-1}(t)}{x_{j+k} - x_{j+1}} \quad (3.10)$$

The x_j are elements of a knot vector. The sequence of knot vector elements must be nondecreasing, i.e. $x_j \leq x_{j+1}$. The recurrence relation in (3.10) allows computing higher order basis functions from the lower order ones. To compute a B-spline of order k , two

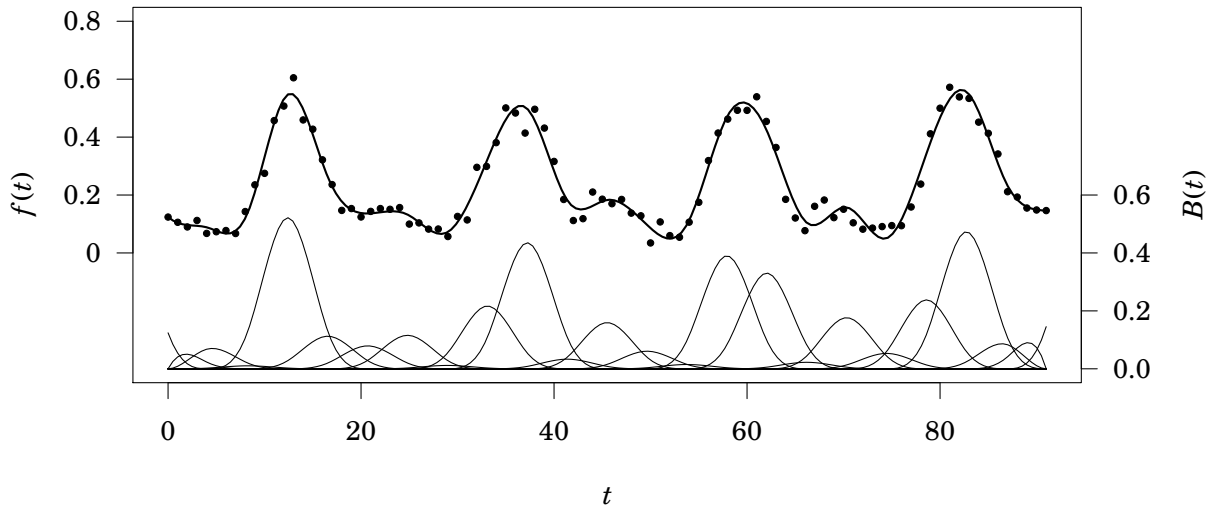


Figure 3.2: Illustration of a B-spline curve (thick solid line) and its basis functions fitted to a four-year time series of MODIS EVI values from a location in the Aksu dataset. The dots show the original values of the time series. The basis functions shown in the lower part of the figure are weighted by their appropriate coefficients. Knot locations correspond to the peak values of the basis functions.

B-splines of order $k - 1$ will have to be computed, and to calculate these two B-splines of order $k - 1$ again, three B-splines of order $k - 2$ must be computed and so forth. Whenever a number of k consecutive knots in a knot vector have the same value, at least one of the denominators in (3.10) will be zero. In this case, convention $0/0 = 0$ is adopted to define the basis functions.

The resulting curve $f(t)$ in (3.8) is a polynomial spline function of order k , its degree is $k - 1$. This means that $f(t)$ is a polynomial of degree $k - 1$ on each knot interval $x_j \leq t < x_{j+1}$. For example, a B-spline curve of order four is a piecewise cubic curve and consists of intervals of cubic polynomials. The order of a B-spline curve can be no larger than the number of basis functions. With each basis function $B_{j,k}$ is associated a B-spline coefficient c_j . B-spline basis functions are polynomial bases with compact support: they are nonzero only over a relatively small interval. Any point on a B-spline curve is created as a linear combination of only k nonzero basis functions. All curve points that fall in a given interval between two consecutive knots use the same k basis functions. Every piecewise polynomial can be represented as a weighted sum of B-splines in a similar way.

The line string formed by the set of points $\{P_1, P_2, \dots, P_m\}$, where $P_j = (g_j, c_j)$ and

$$g_j = \frac{1}{k} \sum_{r=1}^k x_{j+r}$$

is called the control polyline of the B-spline. The values g_j are referred to as *Greville abscissae* and are the averages over all the knot values for which a given B-spline basis function is nonzero. There are as many Greville abscissae as there are B-spline basis

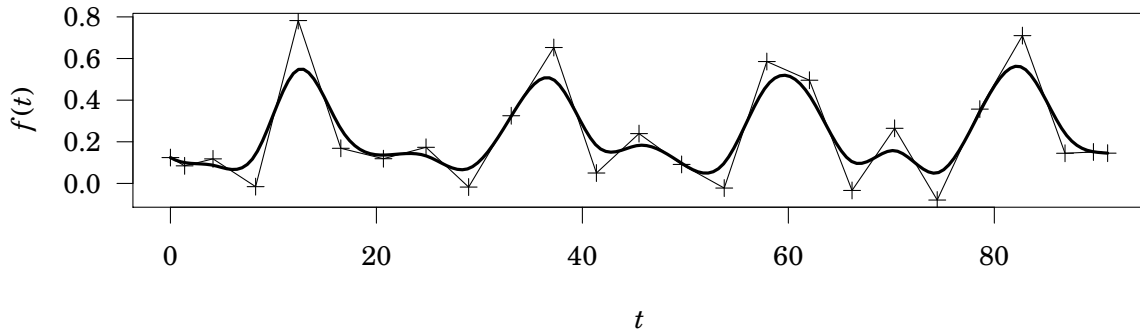


Figure 3.3: A B-spline curve (thick line) and its control polyline (thin line). The cross symbols mark the breakpoints $P_j(g_j, c_j)$ between consecutive legs of the line string.

functions. The control polyline is the line that connects the points defined by each Greville abscissa and the coefficient associated with the corresponding basis function. The spline curve generally follows the shape of its control polyline. The control polyline converges to the graph of a B-spline curve if more and more knots are used to define the curve (Dierckx, 1993; Mosier, 2009). Furthermore, the curve does not cross any straight line more often than the curve's control polygon does (Dierckx, 1993; Rogers, 2001; Mosier, 2009). This fact is useful in determining the maximum number of roots of a given B-spline by counting the oscillations of its control polygon about the horizontal line $y = 0$.

Knot vectors

To compute a basis function of order k by the recursion formula (3.10), an interval of $k + 1$ consecutive knots is needed. At the ends of the knot vector, where $k + 1$ consecutive knots do not exist, the knot vector has to be extended by a certain number of entries to be able to calculate all the m basis functions of order k over a given interval $[a, b]$. An extended knot vector that provides support on an interval $[a, b]$ must have $m + k$ knots, where $x_k = a$ and $x_{m+1} = b$ are the two exterior or end knots of the (unextended) knot sequence. The extension knots beyond the exterior knots at both ends of the knot vector may be thought of as virtual constructs that are not actually a part of the knot sequence of the spline, but are needed in order to calculate the B-spline basis functions. The extension knots of a knot vector are determining factors for the behaviour of the spline at the boundaries of the interval $[a, b]$.

Knot vectors for B-splines exist in different categories. The most fundamental difference is between *open* and *closed* or *periodic* knot vectors. Closed knot vectors are used to obtain boundary conditions suitable to produce closed, ring shaped curves with matching function values and derivatives at both ends, while open knot vectors are used to define open curves with boundary conditions that allow the function values at both ends to be different.

Open and closed knot vectors come in two varieties, they are either *uniform* or *nonuniform*, depending on whether the interior knots are positioned at equal distances or not. For a nonuniform knot vector, the positions of the interior knots may be chosen arbitrarily as long as they form a sequence that is nondecreasing. The positioning of the knots has a major influence on the shape of the resulting spline curve. This is especially true for knot vectors with multiple knots at the same position, since knot multiplicities change the smoothness conditions of the spline curve at that position (Rogers, 2001; Mosier, 2009).

Open knot vectors

Most commonly, open knot vectors are defined by repeating the values of the exterior knots $k - 1$ times at both ends (Dierckx, 1993; Rogers, 2001; Farin, 2002),

$$\begin{aligned} x_1 = x_2 = \dots = x_k = a \\ b = x_{m+k-1} = x_{m+k} = \dots = x_{m+2(k-1)} \end{aligned} \tag{3.11}$$

so that a total number of k knots including the exterior knots are coincident at either end of the knot vector. When all knots are at different positions, the derivatives of a B-spline of order k are continuous up to and including the $k - 2$ derivative. If two knots are at equal positions, then the B-spline will only be continuous up to the $k - 3$ derivative at that position (Rogers, 2001; Mosier, 2009). By placing knots at the end to make the knot multiplicity k -fold, all smoothness conditions imposed by the B-spline basis functions are dropped for the exterior knots, and the behaviour of the spline at the ends is dictated solely by the boundary conditions. This leads to a so called *clamped* B-spline curve that passes through the two end points of the control polyline P_1 and P_m . In these points, the slope of the spline curve is equal to the slope of the first and last leg of the control polyline, respectively. Figure 3.4 shows a set of eight open B-spline basis functions of order four for a knot sequence $\mathbf{x} = (x_4, \dots, x_9)$ of six knots. The corresponding open knot vector has six $(2k - 1)$ more entries corresponding to the multiple knots at both ends. Rather than actually appending the additional knots at both ends to the knot vector, a different convention is to define a corresponding multiplicity vector to a given knot vector that specifies the multiplicity of every knot in the knot vector (Farin, 2002). Following this convention, the knot vector $\mathbf{x} = (x_4, x_5, \dots, x_8, x_9)$ has multiplicity $(4, 1, \dots, 1, 4)$.

Closed knot vectors for periodic splines

Closed knot vectors are used to model closed curves, or in the case of time series, periodic signals. The knot vector of a periodic B-spline of order k defined on an interval $[a, b]$ is of

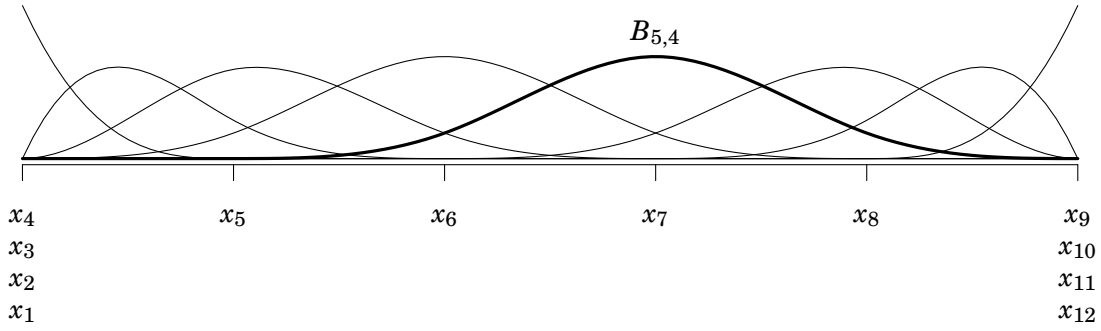


Figure 3.4: B-spline basis functions of order $k = 4$ for an open uniform knot sequence $\mathbf{x} = (x_4, x_5, x_6, x_7, x_8, x_9)$. The end knots x_4 and x_9 have 4-fold multiplicity. The thick solid line highlights the 5th B-spline basis function of order four. Each basis function is only nonzero over four (k) intervals between knots.

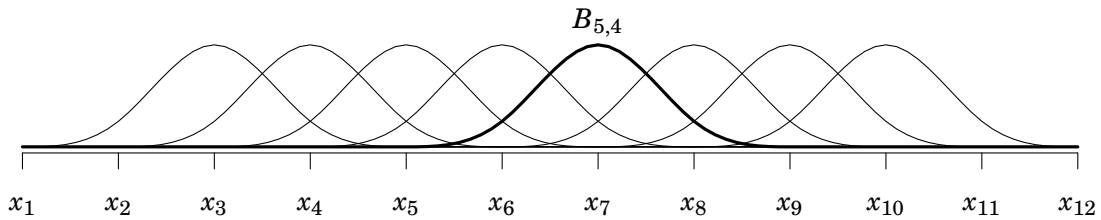


Figure 3.5: Uniform periodic B-spline basis functions of order $k = 4$ for a knot vector $\mathbf{x} = (x_1, x_2, \dots, x_{12})$. Periodic basis functions are merely translates of each other. There are no distinct end knots, as periodic B-spline bases describe closed curves without a distinct starting or ending point.

the form

$$\begin{aligned} x_{k-j} &= x_{m+k-1-j} - \tau \\ x_{j+m+k-1} &= x_{j+k} + \tau \end{aligned} \tag{3.12}$$

where $\tau = b - a$ is the period of the curve. The number of basis functions is denoted by m , and $j = 1, \dots, k - 1$. Like in the case of an open knot vector, the closed knot vector $\mathbf{x} = (x_1, \dots, x_{12})$ supports a B-spline basis of dimension eight on an interval $[a, b]$, where $a = x_4$ and $b = x_9$. Each periodic B-spline basis function is a translate of its predecessor, there are no distinct end knots and no knot multiplicities, since a periodic curve has no ends. A periodic knot vector merely creates the prerequisites for imposing periodicity conditions, but does not actually impose them. The values at both ends of a periodic spline are equal up to the $(k - 1)$ derivative. Thus, in order for a B-spline to be periodic, its coefficients must be determined so that the first and last $k - 1$ B-spline coefficients are equal:

$$c_j = c_{m-k+1+j} \quad j = 1, \dots, k - 1$$

Figure 3.6 shows an illustration how a closed elliptic curve is modelled from a set of erro-

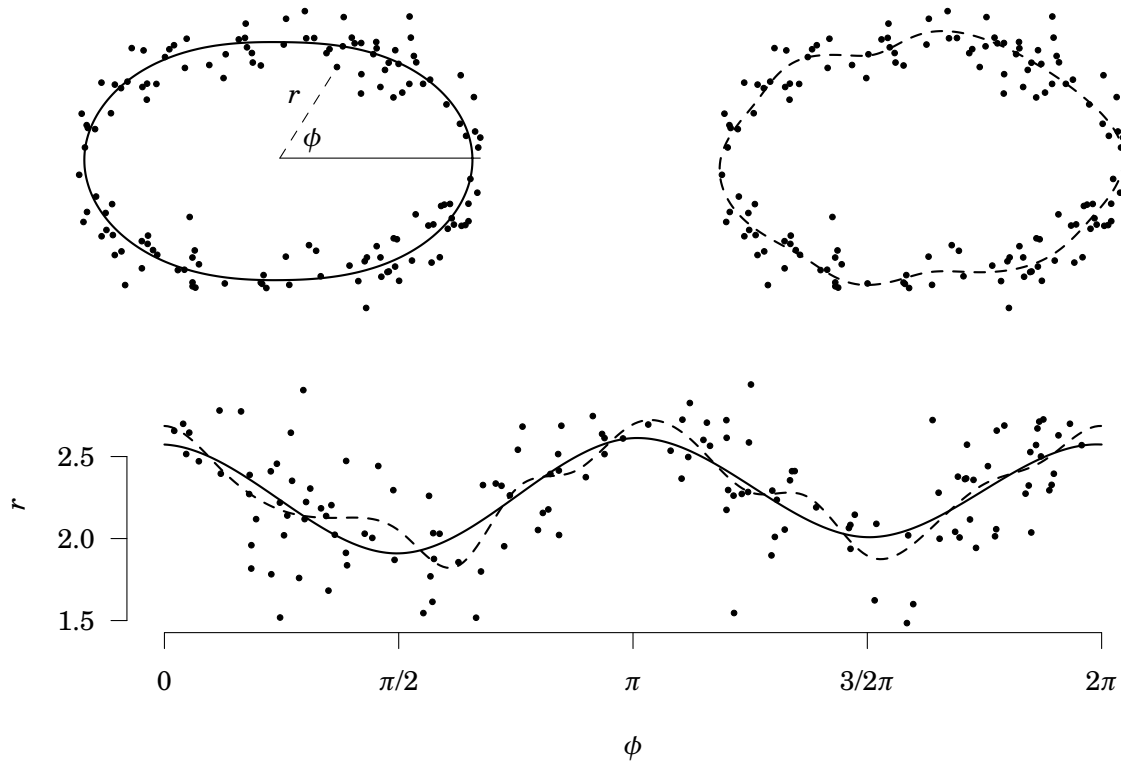


Figure 3.6: Illustration of the concept of a periodic basis and end conditions. A closed spline curve of period 2π is fitted to an elliptical cloud of points in the least squares sense. The cloud is generated by adding white noise to a perfect ellipse. The cubic spline curve in the top left panel with four knots at $\{0, \pi/2, \pi, 3/2\pi, 2\pi\}$ is appropriate to reconstruct the underlying ellipse. The cubic spline curve in the top right panel uses 16 uniformly distributed knots in the range $[0, 2\pi]$ to produce a more complex shape but slightly overfits the data. The bottom panel contains a representation of the curves and data in the polar coordinate system used in fitting the models.

neous point data using a periodic B-spline basis defined in polar coordinates. The dataset was used in unit-testing the basic B-spline implementation for the time series analysis framework. Figure 3.7 shows a periodic spline model fitted to a small dataset of NDVI values obtained from raw Landsat TM imagery over the Aksu region representative for a single year (To increase the number of data, imagery from three years, 2009, 2010 and 2011 has been combined under the assumption that there may be only negligible change within a short three year period). Periodic splines may be useful in processing datasets with incomplete coverage of only a single growing period, where they may be able to provide more realistic estimates of the length of the growing season.

3.4 Fitting B-splines to observational data

The shape of a B-spline curve is controlled by its control polyline. In computer graphics applications (e.g. Rogers, 2001; Farin, 2002) a user most typically modifies the control polygon directly to gain control over the shape of a spline curve. Fitting a curve to observational data is the same problem in reverse: given a set of possibly error prone data

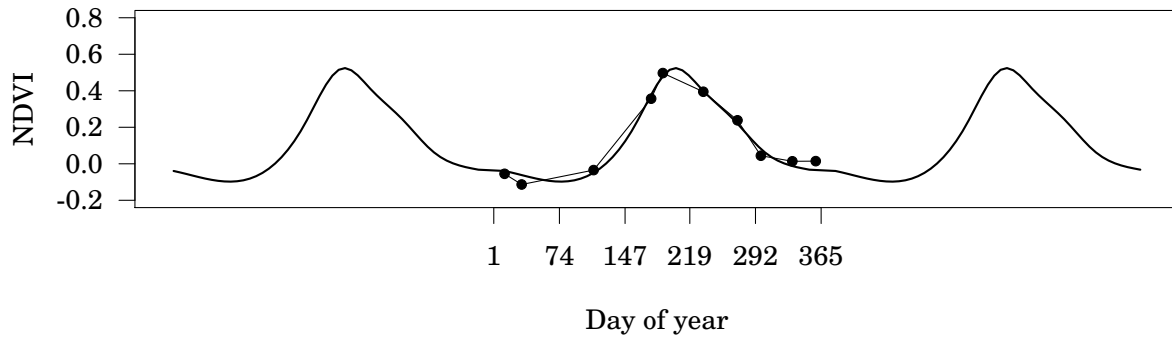


Figure 3.7: Three periods of a periodic spline model fitted to NDVI observations calculated from Landsat TM data. The model is used here to construct a continuous signal with a period of one year sampled according to the schedule of a MODIS 8-day product. The dots connected by the thin line shows the original data.

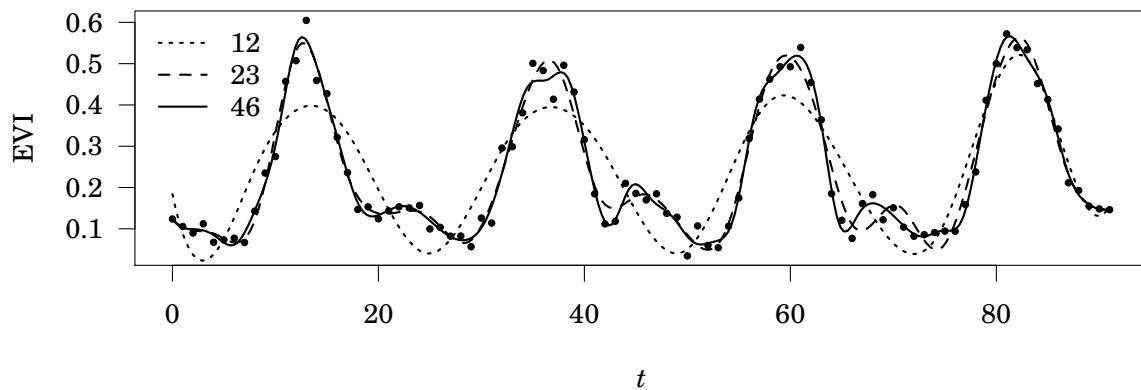


Figure 3.8: Four years of a MODIS EVI time series from the Aksu dataset, and three regression splines fitted to the data using different numbers of knots. The knot vectors used are all uniform. The number of records in the time series is $n = 92$.

with a certain shape, estimate the vertices of a control polygon that gives rise to a spline curve with a close fit to the data. The number of unknowns involved in determining a B-spline curve is thus equal to the number of control points. The B-spline representation is chosen because it is easy to manipulate and numerically stable, not because it is the solution to an optimisation problem. There are a number of different ways to put an optimisation problem to obtain the B-spline coefficients of a spline curve that fits a given data set. By putting the optimisation problem in different ways, spline curves of different properties may be obtained. Cubic spline interpolation, probably one of the most popular uses for splines, has been described in chapter 3.1 for splines in piecewise polynomial form. The equations for an interpolating cubic B-spline can be set up in a similar way using B-spline basis functions (3.10) instead of piecewise polynomials (e.g. Farin, 2002; Epperson, 2007). A remotely sensed time series typically is an error prone data set and a meaningful spline model for the data can only be fitted in a mean sense by using least squares procedures.

Least squares approximation

Whenever there are fewer knots than data points the curve fitting problem can only be solved in a mean sense by minimising the sum of squares of the data points y_t and the values of the regression spline $f(t)$:

$$\sum [y_t - f(t)]^2 \quad (3.13)$$

More precisely, given a data set $\{(t_i, y_i)\}_{i=1}^n$, the m coefficients c_1, \dots, c_m of a regression B-spline of order k with $(m - k + 2)$ knots is found by minimising the following objective function:

$$(c_1, \dots, c_m) = \operatorname{argmin}_{c_j} \sum_{i=1}^n \left[y_i - \sum_{j=1}^m c_j B_j(t_i) \right]^2 \quad (3.14)$$

Let \mathbf{B} be an $n \times m$ matrix of B-spline basis functions of order k , where the j th basis function is in column j of \mathbf{B} , $j = 1, \dots, m$, and the entries (rows) of the column vector \mathbf{b}_j are the values for the j th basis function of order k , evaluated at locations t_i , $i = 1, \dots, n$:

$$\mathbf{B} = \begin{bmatrix} B_1(t_1) & B_2(t_1) & \cdots & B_m(t_1) \\ B_1(t_2) & B_2(t_2) & \cdots & B_m(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_1(t_n) & B_2(t_2) & \cdots & B_m(t_n) \end{bmatrix} \quad (3.15)$$

The basis function values $B_j(t_i)$ are calculated according to (3.10). In equivalence to (3.8), a B-spline representation of a discrete data set $\{(t_i, y_i)\}_{i=1}^n$ may then be written in matrix form as

$$\mathbf{y} = \mathbf{B}\mathbf{c} + \mathbf{e} \quad (3.16)$$

where \mathbf{y} is a vector of data records y_i at times t_i according to the data set, \mathbf{B} is the basis matrix and \mathbf{c} is a vector of B-spline coefficients. The n -vector $\mathbf{e} = (e_i)$ is a vector of residual errors of the spline model. To fit a spline model to a set of observational data $\{(t_i, y_i)\}$, equation 3.16 has to be solved for the coefficient vector \mathbf{c} in the least squares sense by using the normal equations (e.g. Golub and van Loan, 1996),

$$(\mathbf{B}^T \mathbf{B})\mathbf{c} = \mathbf{B}^T \mathbf{y} \quad (3.17)$$

to yield the coefficient vector \mathbf{c} as

$$\mathbf{c} = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \mathbf{y}$$

which is equivalent to minimising (3.14). How close a spline curve may be to a given set of data mostly depends on the number of basis functions m employed in the fit, and to a lesser extent on the degree of the spline's polynomial pieces.

Weighted least squares

Equation (3.17) describes an ordinary least squares fit. It has been proposed by Chen et al. (2004) that the quality information available in a number of long term archives of remotely sensed data should be used to improve the robustness of analysis methods. If such quality information is to be incorporated in a least squares procedure to make the fit less dependent on erroneous data, a weighted least squares procedure may be used, which requires only a slight modification of the normal equations (3.17) for an ordinary least squares fit. For a weighted least squares fit, a weight matrix \mathbf{W} enters (3.17):

$$(\mathbf{B}^T \mathbf{W} \mathbf{B}) \mathbf{c} = (\mathbf{B}^T \mathbf{W}) \mathbf{y} \quad (3.18)$$

and the coefficient vector becomes

$$\mathbf{c} = (\mathbf{B}^T \mathbf{W} \mathbf{B})^{-1} (\mathbf{B}^T \mathbf{W}) \mathbf{y},$$

If weighted least squares is used to adjust the relevance of the individual data points in the fit, $\mathbf{W} = \text{diag}(\mathbf{w})$ is a diagonal matrix and $\mathbf{w} = (w_i)_{i=1}^n$ is a vector of weights associated with the data points y_i in \mathbf{y} . For example, Jönsson and Eklundh (2004) propose to use weights around 1, 0.5 and 0 for data points acquired under clear sky, cloudy, and heavily clouded conditions, respectively. Weighted least squares may also be used as a part of an iterative procedure to accommodate the upper envelope of an NDVI time series by adjusting the weights between iterations according to the distance of a data point from the modelled curve (Sellers et al., 1994; Jönsson and Eklundh, 2002; Hermance et al., 2007). If, for example, a data point lies below the modelled curve in one iteration, it may be considered less important and its associated weight in the next iteration step may be lowered by a certain factor. This iterative procedure leads to a model function that is adopted to the upper envelope of the data (Jönsson and Eklundh, 2002).

3.5 Smoothing splines

In principle, if there is a record of observational data available with n samples, it is possible to fit a spline $f(t)$ to these data that consists of $n - 1$ polynomial pieces, i.e. one piece between every two data points. If such a model $y_t = f(t) + \varepsilon_t$ is fit in the least squares sense to approximate $f(t)$, a very rough curve is obtained that is likely to overfit the data. Regression splines approach this problem by reducing the number of knots in the spline model. In this way, the curve is represented by fewer polynomial pieces and a smoother, more average fit is obtained.

Smoothing splines follow a different approach: in a smoothing spline, the number of knots is equal to or at least of the order of the number of data points in a sample that is used to approximate a spline curve, but instead of merely minimising the sum of squared

errors (3.13), a different least squares criterion is employed that imposes a penalty on the roughness of the resulting spline curve. The most common of such penalised least squares criterions involves the second derivative of the spline function $f''(t)$,

$$\sum [y_t - f(t)]^2 + \lambda \int [f''(t)]^2 dt \quad (3.19)$$

where the squared second derivative is integrated over the domain of the fit and λ is a proportionality constant that determines an appropriate trade off between the closeness of fit to a given data set and the smoothness of the curve. The extra penalty term can be represented in matrix form by a penalty matrix (Hastie et al., 2009). For a B-spline, let the matrix \mathbf{D} be the $m \times m$ matrix of inner products of second derivatives of B-spline basis functions with entries

$$d_{jr} = \langle B_j''(t_i), B_r''(t_i) \rangle$$

where $\langle B_j''(t_i), B_r''(t_i) \rangle$ denotes the inner product, or dot product, of the second derivatives of the j th and r th basis functions of a B-spline basis of dimension m , evaluated at the locations of the data points $t_i, i = 1, \dots, n$

$$\langle B_j(t_i), B_r(t_i) \rangle = \sum_{i=1}^n B_j''(t_i) \cdot B_r''(t_i)$$

The matrix \mathbf{D} may be calculated as

$$\mathbf{D} = (\mathbf{B}'')^T \mathbf{B}''$$

where \mathbf{B}'' is an $n \times m$ matrix of second derivatives of basis functions similar to (3.15). The elements of \mathbf{B}'' are calculated recursively according to (3.21). The matrix \mathbf{D} is used as a penalty matrix to obtain a smooth B-spline curve by solving the following normal equations:

$$(\mathbf{B}^T \mathbf{B} + \lambda \mathbf{D}) \mathbf{c} = \mathbf{B}^T \mathbf{y} \quad (3.20)$$

A major concern with smoothing splines is the choice of the smoothing parameter λ . The degree of smoothing has a major influence on regression estimates from spline models. If λ is too low, the spline will pick up too much detail and the data will be overfitted, if it is too high, the data will be over smoothed. In both cases the spline model will not be a good estimate of the conditional mean values of the data. One possibility for estimating an optimal smoothing parameter is by cross validation (Wahba, 1990; Takezawa, 2006; Wood, 2006). By leaving out each data point in turn and fitting the model to the remaining data, a cross validation score can be determined by calculating the squared difference between the missing data point and its modelled value. Subsequently, the squared differences are averaged over the entire data set to yield an overall score. Figure 3.10 shows a cubic smoothing spline fitted to an EVI time series of one year from a location in the

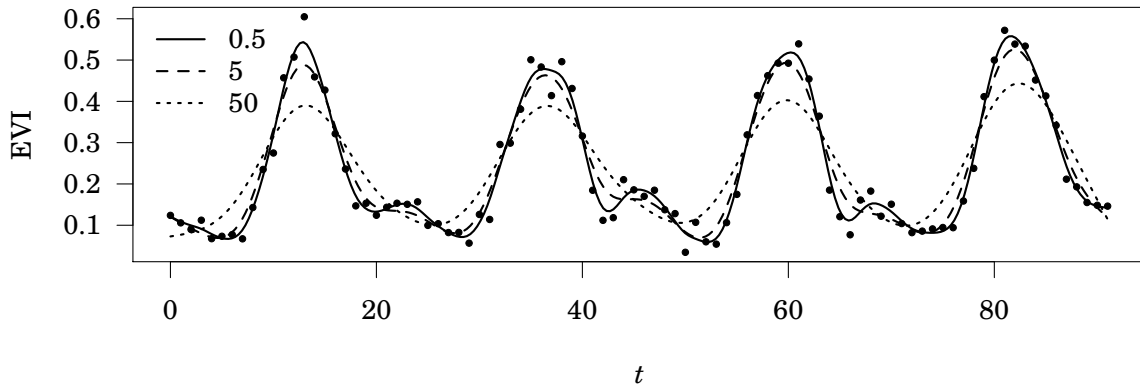


Figure 3.9: Three different smoothing splines fitted to a four year time series of the Aksu dataset with a smoothing factor λ of different orders of magnitude ($\lambda = 0.5, 5, 50$). All spline models have a number of knots equal to the number of datapoints.

Aksu dataset. The cross validation score of candidate values for λ on a logarithmic grid is evaluated to determine the best λ with the minimum score. The minimum for λ may be determined by setting up an objective function for cross validation and subsequent minimisation of this function using a method for function minimisation in one dimension, e.g. a golden section search (Press et al., 1992). However, cross validation assumes that mean and variance $\text{Var}(\epsilon_t)$ of the residual errors are constant across the entire curve. If time series modelling is understood as a means to estimate an underlying unknown function $m(\cdot)$ that gives rise to an observed signal, this may be a problem. The appropriate degree of smoothing should depend on the true smoothness of $m(\cdot)$, as expressed in the second derivative $m''(\cdot)$. In the case of time series the errors may be autocorrelated. If the properties of $m''(\cdot)$ are unknown, it is impossible to determine how much of the irregularity in the observed data is due to $m(\cdot)$, and how much is due to the irregularity in the residual errors $\text{Var}(\epsilon_t)$. If the variance in the observed data is low, this may be because $m(\cdot)$ is smooth, or it may be the result of highly autocorrelated residual errors. (Bowman and Azzalini, 1997). If the properties of $m''(\cdot)$ cannot be determined from an independent source of information, an estimate of the autocorrelation function is needed to assess the proportions of true variance in the signal and variance due to erroneous observations. Cross validation should only be used to determine λ in the case of uncorrelated residual errors.

P-splines

A different type of regularised spline models called *P-splines* was proposed by Eilers and Marx (1996). P-splines are smoothing splines that define roughness based on the shape of the spline's control polygon rather than penalising the roughness of the spline curve itself using its second derivative. Roughness is measured according to the second differences in the B-spline coefficients c_j that form the curve's control polygon. If all the

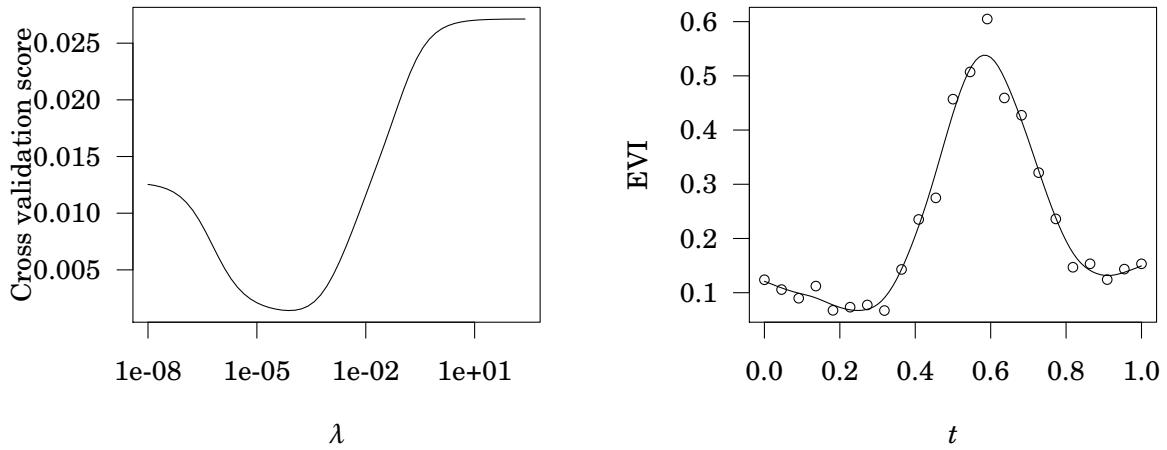


Figure 3.10: Estimation of an optimal smoothing parameter λ using cross validation. The left panel shows the cross validation score for different values of λ . Candidate values are computed on a logarithmic grid. The right panel shows the cubic smoothing spline with the optimal $\lambda = 7.5 \times 10^{-5}$ at the minimum of the curve on the left side. The positions of the spline's knots correspond to the data points. The data is from one year of an EVI time series from a location in Aksu.

second differences $\Delta^2 c_j$ are small, then the control polygon does not oscillate much and prescribes a smooth curve shape. Thus, in addition to minimising the sum of squares of the differences between measured and fitted values (3.13), Eilers and Marx (1996) aim at minimising the second differences of the coefficients of a B-spline basis with a uniform set of knots

$$\begin{aligned} c_1 - 2c_2 + c_3 &= 0 \\ &\vdots \\ c_{m-2} - 2c_{m-1} + c_m &= 0 \end{aligned}$$

which leads to the following expression for minimisation:

$$\sum_{i=1}^n [y_i - \sum_{j=1}^m c_j B_j(t_i)]^2 + \lambda \sum_{j=p+1}^m (\Delta^p c_j)^2$$

The operator Δ^p is the p th difference operator, for example, using second differences as proposed by Eilers and Marx (1996), $\Delta^2 c_j = c_j - 2c_{j+1} + c_{j+2}$. As with the second derivative penalty, the tradeoff parameter λ controls the balance between the smoothness of the resulting spline curve and the closeness of its fit to a given set of data. Setting $\lambda > 0$ leads to the fact that the absolute differences in the coefficients of adjacent B-spline basis functions are constrained. As a consequence, the curve composed of these basis functions and corresponding coefficients does not exhibit steep or rapidly changing slopes. Because of the use of differencing, the intervals between adjacent knots along the curve must all be equal, which means that the P-spline penalty can only be used with B-splines having uniform knot vectors. To obtain the coefficients for a smooth B-spline curve, the linear

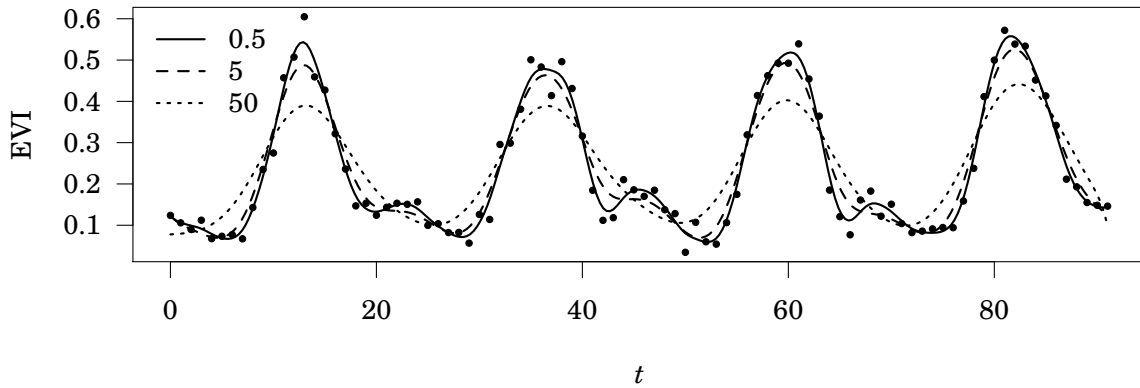


Figure 3.11: Examples of P-spline models using second difference penalties with different proportionality parameters $\lambda = 0.5, 5, 50$ fitted to the four-year Aksu timeseries.

system of equations to solve is

$$(\mathbf{B}^T \mathbf{B} + \lambda \mathbf{D}_p^T \mathbf{D}_p) \mathbf{c} = \mathbf{B}^T \mathbf{y}$$

where \mathbf{D}_p is the p th order difference matrix, obtained from differencing the identity matrix p times. The elements of a single row of \mathbf{D}_p are the p th differences of the corresponding elements of the identity matrix. For a B-spline basis of dimension m , differencing the identity matrix with m rows and columns p times yields a matrix \mathbf{D}_p of m rows and $m - p$ columns. The parameter λ is a smoothing parameter similar to that of a smoothing spline with a second derivative penalty.

P-splines conserve moments (means, variances) of data (Eilers and Marx, 1996, 2010) and are a valuable option when dealing with time series of regularly spaced observations. P-splines have the advantage that, compared to the smoothing splines with a second derivative penalty, the calculations of second differences of the coefficients are simpler and faster. Furthermore, if the second derivative of a spline curve is used in evaluating the spline model, for example to locate extreme points, it may be more appropriate to use a penalty that is not directly related to the second derivative of the curve. However, there is a close relation between the P-spline penalty and the penalty term based on the second derivative of the curve. Figure 3.11 shows three different P-splines fitted to a four-year time series of the Aksu dataset. The P-spline penalty has an effect very similar to the smoothing spline (figure 3.9), but is computationally less involved. P-splines are therefore a good alternative to smoothing splines if a uniform knot vector is used.

Special penalties

The fact that B-splines are compactly supported and have local control can be useful in defining special penalties, where only a part of the spline curve is subject to smoothing. Since the parameters of a B-spline affect only a small part of the entire spline, the linear

system of equations (3.20) can be modified to switch smoothing on and off for different parts of the spline curve. When smoothing is not desired on a given knot interval, the penalty terms in the penalty matrix \mathbf{D} for the basis functions that are nonzero on that interval are eliminated. This may be useful if the vegetation signal in a remotely sensed time series is faint and driven by erratic precipitation events, such as in observations of drylands. Greenup in dry environments often follows immediately after rainfall events (Archibald and Scholes, 2007). Smoothing is appropriate in dry periods when there is no discernible vegetation signal and the data is mostly noise. In wet periods, however, it may prevent the detection of faint, but rapidly changing phenological signals. By locally switching on and off the smoothing in dry and wet periods, a better overall model for a time series may be built.

3.6 Fitting periodic spline models

Periodic splines are useful in building models with similar properties as a Fourier series, where a time series is treated as a finite record of an infinitely repeating, periodic signal. With a suitable periodic basis, a spline can be constructed where the ends join smoothly to accommodate annual or semiannual cycles in remotely sensed time series. As already mentioned in chapter 3.3, the first and last $(k-1)$ coefficients of a B-spline of order k must be equal to ensure periodic continuity at the ends of the spline. To set up a basis matrix \mathbf{P} similar to (3.15) for a periodic spline, the last $(k-1)$ spline coefficients are eliminated by dropping the last $(k-1)$ columns of the basis matrix \mathbf{B} in (3.15) after adding their elements to the first $(k-1)$ columns of \mathbf{B} . Thus, if a periodic B-spline basis has $m = r + k - 1$ basis functions, \mathbf{P} is an $n \times r$ matrix, where n is a number of locations where the spline is to be evaluated (Dierckx, 1993; Eilers and Marx, 2010). In this way, a basis is formed where the basis functions are “wrapped around” to support a periodic curve.

3.7 Knot number and knot placement

For a regression spline model, one must select the number of knots and where to place them along the time axis. The number of knots has an important effect on the spline fit. Using a high number of knots in a regression spline yields a spline with little bias but increased variability and may result in overfitting a given data set. On the other hand, if the number of knots is too low, the resulting spline will be overly smooth, with little variability but possibly largely biased. Usually, the number of knots in a regression spline is chosen much smaller than the number of data points. In principle, the number of knots and their positions are free parameters of a spline model occurring in a nonlinear way (Wold, 1974). There exist strategies for automatic spline fitting that try to select an optimal number of knots and positions. Such strategies commonly work by systematically varying the positions of the knots until an acceptable minimum is found for the sum of squared residuals. In an iterative process, knots are added in intervals where the resid-

uals are inacceptably large (Wold, 1974; Dierckx, 1993). However, the choice of a number of knots and positions is not a complicated matter. The knots should be chosen so as to reflect the specific properties of a given data set, such as the number of available data points, the positions of extreme points. There should be a reasonable number of points per knot interval, and the number of extreme points per interval should be chosen in accordance with the degree of the spline model. For a cubic spline, there should be no more than one extreme point and one inflexion point per interval, since a cubic polynomial is not capable of approximating more variations (Wold, 1974). The common practice of placing knots at evenly spaced intervals within the range of the data makes sense if the data itself is given at uniform intervals, as is the case with most remotely sensed time series. Equally spaced knot intervals make sure that there are enough data points in each interval to obtain a reliable estimate of the underlying curve. Nonuniform knot vectors may be more appropriate if the data has obvious features. When noise and structure in the data are not equally distributed, it may be reasonable to reduce the number of knots in less structured parts of the spline curve, similar to the special penalties discussed in chapter 3.5. For example, in an NDVI time series, the number of knots in an interval outside the growing period, where the NDVI values are low and noisy, may be less than within a growing period, where more variability is needed in the fitted model to accommodate the shape of the phenological cycle. To honour the general scientific precept of Occam's razor (MacKay, 2003), there should be as few knots as possible in order to keep the number of model parameters small. Stone (1986) found that the positions of the knots in a spline model are less important than the number of knots used.

3.8 Derivatives of B-splines

The ability of a spline model to continuously represent derivatives of the modelled curve is very useful in developing algorithms to extract phenological parameters from remotely sensed time series, e.g. to identify the local minima and maxima along the curve. Many methods to determine key phenological transition dates such as the beginning of greenup and senescence rely on derivative information (e.g. Tan et al., 2011) or curvature and its rate of change (e.g. Zhang et al., 2003). For a B-spline defined as in equation (3.8), the derivative of the curve $f(t)$ may be found by computing the derivative of each of the basis functions:

$$B'_{j,k}(t) = \frac{k-1}{x_{j+k-1} - x_j} B_{j,k-1}(t) - \frac{k-1}{x_{j+k} - x_{j+1}} B_{j+1,k-1}(t) \quad (3.21)$$

It follows (Dierckx, 1993; Rogers, 2001) that the derivative of a B-spline curve is another B-spline curve defined on the original knot sequence and of order one less than the original curve

$$f'(t) = \sum_{j=1}^{m-1} d_j B_{j+1,k-1}(t) \quad (3.22)$$

where $\{d_1, \dots, d_{m-1}\}$ is a new set of basis function coefficients given by

$$d_j = \frac{k-1}{x_{j+k} - x_{j+1}}(c_{j+1} - c_j)$$

The fact that a first derivative of a B-spline is another B-spline allows that higher order derivatives may be computed recursively. Based on the derivatives, the curvature of a polynomial is given by the expression (Edwards and Gordon, 2004):

$$\kappa = \frac{f''}{(1 + f'^2)^{3/2}}$$

When f is a polynomial curve of degree greater than one, the rate of change of its curvature is given by

$$\kappa' = \frac{(1 + f'^2)f''' - 3f'f''^2}{(1 + f'^2)^{5/2}}$$

Thus, a spline curve can be efficiently analysed to estimate crucial phenological dates from time series of vegetation indices.

3.9 Finding zeroes and inverse values

The continuous derivatives of a spline curve, which are splines themselves, may be used to determine the minima and maxima of a phenological cycle by analysing the extreme points along the time series. To this end, the zeroes, or roots, of the first derivative have to be found and the second derivative must be evaluated at these points to determine whether a given extreme point is a minimum or a maximum. Since the curve and its first and second derivatives are functional representations, this happens independently of the sampling resolution of the original time series. The problem of finding the roots of polynomials is well studied (Hamming, 1973; Acton, 1990) and several elaborate algorithms are available to solve it. The method of Jenkins and Traub (1970) for real and complex polynomial root finding is common and has established as a standard (Press et al., 1992). There is a variation of the algorithm which is faster if the polynomial's coefficients are real (Jenkins, 1975). It is straightforward to apply this algorithm to a spline curve, since the zeroes of a spline function $f(t)$ may be determined by finding the real zeroes of its polynomial pieces (Dierckx, 1993).

If the spline is in its B-form (3.8), it must be converted first to its piecewise polynomial form to apply the method. It has been shown by de Boor (2001) how this conversion may be done in a numerically stable and efficient manner by evaluating the B-spline and its $(k-1)$ derivatives at the knots,

$$\mathbf{C}_{j,r} = \frac{d^{r-1}}{dt^{r-1}}f(x_j)$$

where \mathbf{C} is a $(m-3) \times k$ matrix of derivatives, where $(m-2)$ is the number of knots. The r th entry in the j th row of \mathbf{C} is the $(r-1)$ derivative of the spline curve at knot x_j ,

$r = 1, \dots, k$ and $j = 1, \dots, (m - 3)$. The method of Jenkins and Traub (1970) requires the usual polynomial coefficients rather than the derivatives, which may be obtained from the derivatives in \mathbf{C} by the relation $c_{jr} = \mathbf{C}_{j,r}/(j - 1)!$ (de Boor, 2001), where c_{jr} is the r th polynomial coefficient of the j th polynomial piece of the spline. This yields a polynomial representation such as (2.7), where the leading coefficient is the last of k coefficients for a polynomial of degree $(k - 1)$. The j th polynomial piece in the spline's piecewise polynomial representation (3.1) may thus be written as

$$f_j(t) = \sum_{r=1}^k c_{jr}(t - x_j)^{(r-1)} \quad \text{for } x_j \leq t < x_{j+1} \quad (3.23)$$

A root of the polynomial $f_j(t)$ within the open knot interval $J = [x_j, x_{j+1})$ is also a root of the spline $f(t)$, provided that $t \in J$. Root finding using the procedure of Jenkins and Traub (1970) may fail if the leading coefficient (the coefficient associated with the largest power) is close to zero. However, if this is the case, it also indicates that the curve fitting problem is not well posed and a spline model of a lower degree should be used for a given data set.

By using the Jenkins-Traub algorithm, the spline curve is evaluated at suitable locations to get good estimates of its zeroes. This can be computationally demanding, and the computational cost increases with the number of polynomial pieces in a spline. Two important properties of B-splines can be exploited to calculate zeroes more efficiently (Dierckx, 1993). First, the B-spline basis functions are always positive or zero

$$B_{j,k}(t) \geq 0 \quad \text{for all } t$$

second, the basis functions have compact support, i.e.

$$B_{j,k}(t) = 0 \quad \text{if } t \notin [x_j, x_{j+k}]$$

It follows that a B-spline $f(t)$ cannot have a zero in the knot interval $[x_j, x_{j+1}]$ if all nonzero B-spline coefficients for that interval c_r , $r = j, \dots, (j + k - 1)$ have the same sign. Thus, the effort of converting a part of a B-spline to its polynomial representation and calculating the zeroes of the resulting polynomial may be considerably reduced if only those intervals of the spline are considered that actually contain zeroes.

Another convenient property of a B-spline is its *affine invariance*. Affine invariance means that any affine transformation can be applied to a B-spline curve by applying it to the vertices of its control polyline, that is, the spline's coefficients. This helps in solving inverse problems. A common inverse problem in the determination of phenological descriptors is the determination of the point in time when a time series signal increases over, or falls below a certain threshold y (White et al., 2009). A problem equivalent to finding $t = f^{-1}(y)$ is to find the roots of a function $g(t) = f(t) + y$. The translation of a

curve by a certain amount parallel to the time axis can be seen as a simple affine transformation. Given a B-spline with a basis of dimension m with coefficients c_j , the problem of finding the times t at which the spline function $f(t)$ has a prescribed value y is equivalent to finding the roots of the B-spline with $g(t) = f(t) - y$ with coefficients $d_j = c_j - y$, $j = 1, \dots, m$. Thus, the same algorithm that is used in the determination of the zeroes of a spline can also be used to determine the times where a spline curve equals any arbitrary threshold.

4 Estimation of phenological parameters

In chapter 3, it was shown how polynomial splines, especially B-splines, can provide a framework for analysing remotely sensed time series to determine phenological descriptors. B-splines provide robust, data centred functional representations that are, unlike most other methods considered here, complete, self-contained structural models. They not only serve to estimate a phenological curve from a set of data, but also provide the means for analysing the structure of the curve and its phenological cycles. The fact that minima, maxima, inverse values and definite integrals of a B-spline can be obtained in a straightforward, analytical way renders it a very efficient data structure to define phenological descriptors and determine their values from a given time series.

There exist a number of implementations for different kinds of B-splines, e.g. in the FORTRAN language (de Boor, 1977, 2001; Dierckx, 1993), but these implementations often have the character of examples with implicit limitations on, for example, the number of knots that can be used. For a longer time series with hundreds of observations, a limit of typically 30-50 knots is too restrictive, and the restrictions are not easily removed from the code. Thus, the open and periodic regression B-spline, smoothing spline and P-spline models discussed in chapter 3 were implemented as C++ software components to allow the efficient processing of even large sets of remotely sensed data. The matrix operations required to solve the optimisation problems were implemented based on the *armadillo*¹ linear algebra library (Sanderson, 2010). All spline models are derived from a single B-spline base component that provides the facilities to compute derivatives, definite integrals, roots and inverse values of a fitted spline model. Based on the spline software components, a number of application programs were written for the exploration, fitting and exploitation of spline models in the analysis of remotely sensed time series provided as imagery data sets. Besides the spline module, the applications required two additional key components: a graphical user interface and the facility to read and write imagery datasets. The former component was realised on the basis of the *FLTK*² library. For the latter component the geospatial data abstraction library *GDAL*³ was used. An application program to explore and test spline models on remotely sensed data sets is shown in figure 4.1. The programs give the user full control over the type of spline model to use, its degree, and the number and position of the knots.

4.1 Choosing an appropriate spline

The first step in deriving phenological descriptors from a remotely sensed time series is to find an appropriate spline model as a functional representation for it. In chapter 3,

¹<http://arma.sourceforge.net>

²<http://www.fltk.org>

³<http://www.gdal.org>

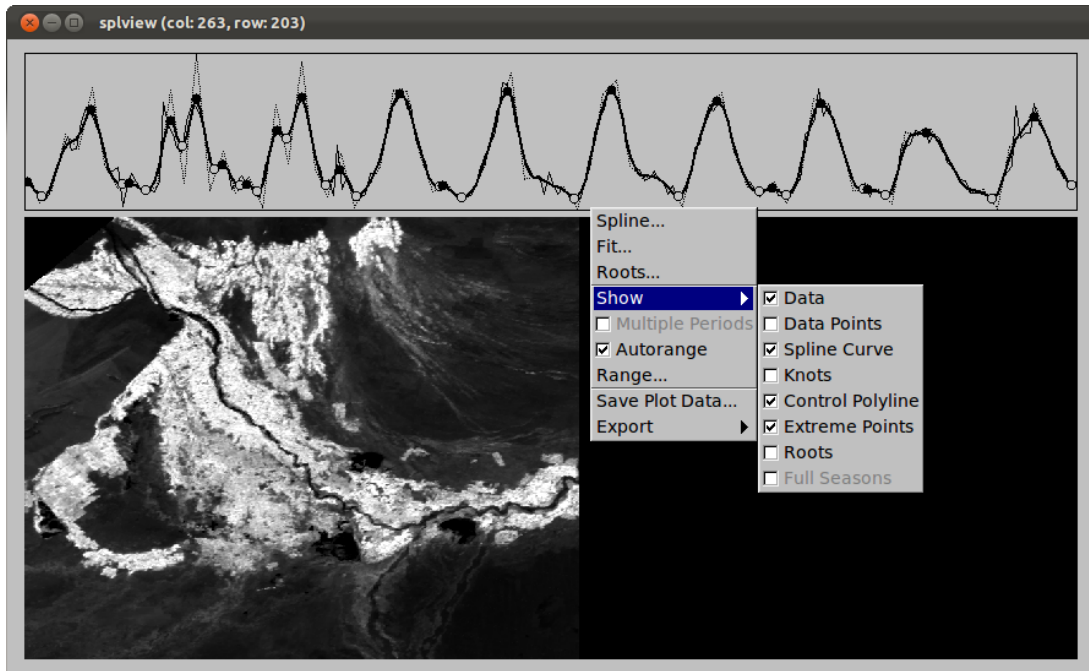


Figure 4.1: Example of an application to explore the use of different spline models with remotely sensed time series. The image viewer shows the Aksu EVI data set.

the properties of different spline models were discussed along with their specific assets and drawbacks. Along with the degree of the spline's polynomial pieces, the complexity of a regression spline is controlled chiefly by the number of knots, while the number of degrees of freedom of a smoothing spline or P-spline is adjusted through the smoothing or regularisation parameter λ . When appropriate, the techniques of regression and smoothing splines may be combined to exert even finer control on the degrees of freedom of the resulting curve. How close a fit is obtained is largely a question of the number of knots. Due to the restrictive effect of the knots on the spline's behaviour, the degree of the polynomial pieces is only a subordinate influence factor. By regularisation, the spline model is biased toward a smoother, less wiggly shape of the overall curve, as opposed to a close fit. Finding an appropriate spline model is often a balancing act between selecting an appropriate number of knots and the amount of regularisation.

Because of the difficulties in fitting higher order polynomials, the degree of a spline curve is not commonly used as a major adjusting parameter to determine the degrees of freedom of a spline fit. Least squares fitted higher order polynomials tend to oscillate and produce large discrepancies near the ends of a modelled curve. Polynomial splines were invented to solve polynomial curve fitting problems by subdividing a curve into a number of pieces that are less complex and easier to fit. But this implies that the degree of the individual polynomial pieces can be expected to be considerably lower than the degree of a single polynomial flexible enough to accommodate the whole data set. If high polynomial degrees are used in least squares fitting clamped B-spline curves, oscillation may still occur near the ends of the spline, where the continuity constraints imposed by the basis

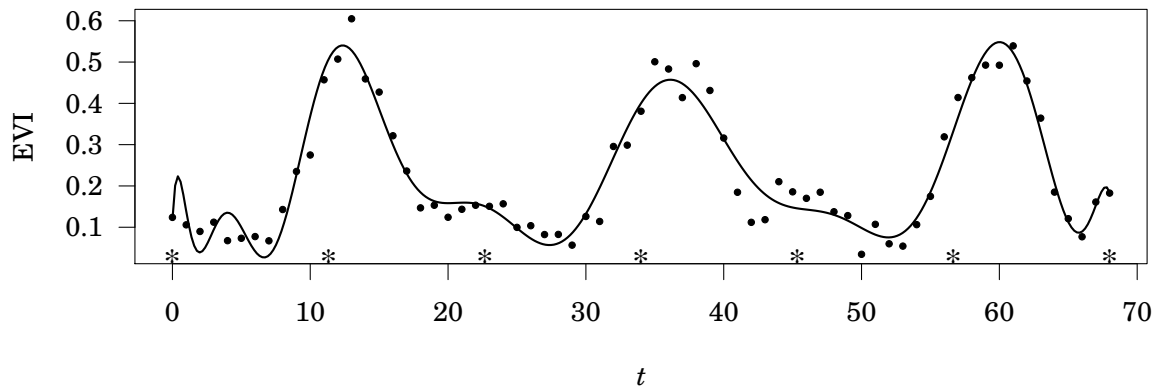


Figure 4.2: A regression spline of order 14 fitted to a three-year section of an EVI time series from a location near Aksu. Each year is subdivided into two polynomial pieces. The knots are indicated by the asterisk symbols along the horizontal axis. In its interior, the spline fits the data well. The oscillations at the left end of the curve are a result of the high polynomial degree.

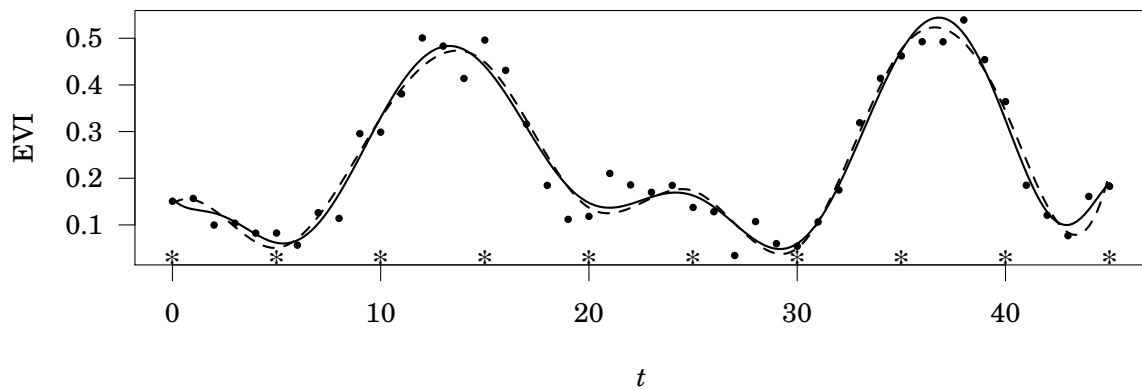


Figure 4.3: Regression splines of different order with 10 knots fitted to a two year time series from the Aksu dataset. A year is covered by 5 polynomial pieces. The grid of knots is indicated by the star symbols at the bottom, the dots represent the original data points. There is not much to choose from a spline of order 4 (dashed curve) compared to a spline of order 6 (solid curve).

functions are dropped (figure 4.2). From the ends inwards, the continuity conditions at the interior knots prevent uncontrolled oscillations of the high degree polynomial pieces. As a consequence of the restrictive conditions in the interior of a spline, the shape of a spline curve does not change much if the degree of the spline is increased (figure 4.3). The extent of the shape differences between splines of different order depends on the number of knots in the two splines. The more knots in a spline, the greater the number of interior constraints and the lesser the extent of the shape differences with increasing spline order.

Basically, two regression splines may have a similar number of degrees of freedom if one of them has only few knots, but is of a large order, while the other has many knots and a low order. An equivalent smoothing spline would have an even higher number of knots in combination with an appropriate smoothing factor. In any case, the goodness

of fit for the model depends on how well the linear optimisation problem to determine the model's coefficients may be solved for a given real-world data set. In order to fit a regression spline model of a high polynomial degree to a typical real-world time series with erroneous or missing observations, an iterative weighted least squares procedure is required to achieve an acceptable fit (Hermance et al., 2007; Bradley et al., 2007). Moreover, the evaluation of a high order B-spline curve is computationally intensive because of the recursive definition of the B-spline basis functions. Splines of lower order are less computationally demanding and numerically more stable. If only the shape of the modelled curve itself is of interest, no more than a cubic spline may be needed to model remotely sensed time series. A cubic is the lowest degree that can support an inflection, which allows to create curves with shapes such as asymmetries or shoulders commonly found in time series of vegetation indices (Chen et al., 2004; Jönsson and Eklundh, 2002). Splines of high degree (> 5) only make sense if an appropriate number of smooth derivatives is required for any purpose. The first derivative of a cubic spline, for example, is a second degree piecewise polynomial, its second derivative is a piecewise linear curve, and its third derivative is a piecewise constant step function. The cubic spline itself and its first derivative are continuous and smooth, while the second derivative is still continuous, but not smooth. The third derivative is not even continuous. A vivid example for when cubic splines are no longer appropriate is in computer aided manufacturing, where the second derivative of a spline stands for an acceleration profile: if a robot arm is to handle a grinding tool to grind off some material along a path prescribed by a spline curve, that curve's second derivative should be a smooth continuous curve. Otherwise, the robot would not accelerate and direct the tool steadily following the curve (Mosier, 2009). This is why third degree splines are seldomly seen in computer aided manufacturing applications. Because of the demand for smooth second derivatives, robot engineers typically use splines of fifth degree (Mosier, 2009). In applying splines to time series analysis problems in remote sensing, one of the most important things is to find extreme points. In the process of assessing extrema, the second derivative is only evaluated to determine whether a given extreme point is a minimum or a maximum, and need not be smooth. The piecewise linear second derivative of a cubic spline serves fine for the purpose of distinguishing minima and maxima. Some methods for determining phenological parameters rely on third derivative information (Zhang et al., 2003; Tan et al., 2011). In such a case, it may be appropriate to use a quintic spline, whose third derivative is a smooth, quadratic spline curve. Splines of higher degree than five are not generally needed in the analysis of remotely sensed time series. Thus, instead of fitting a high order spline with only a few knots to a remote sensing time series, it is more efficient to use lower order splines and increase the number of polynomial pieces appropriately. In addition, the number of polynomial pieces in a spline is a parameter that is more intuitively adjusted by an analyst than the rather incomprehensible weights in an iterative least squares fit of a high order spline model.

4.2 Deriving phenological descriptors from splines

Even if an optimal spline model of a time series has been established and is ready to be exploited in order to determine phenological descriptors, the challenge remains to identify and quantify vegetation cover by relating the model parameters to principal phenological attributes. In order to derive a set of phenological descriptors from a time series, the foremost thing is to identify a consistent set of minima and maxima that enclose the individual vegetation periods or cycles. The top panel of figure 4.4 shows all the extreme points identified from a detailed regression B-spline model of a 10-year time series from a location near Aksu. The extrema were determined by using the Jenkins-Traub algorithm to find the roots of the spline's first derivative, as discussed in chapters 3.8 and 3.9. An advantage of this method is that the extreme points may occur in any arbitrary pattern along the time axis, since the spline model contains no implicit assumptions of regularity, unlike, for example, the low order Fourier model used by Jönsson and Eklundh (2002). In addition, the accuracy of the resulting extreme points does not depend directly on the dynamics and sampling resolution of original data, since the root-finding algorithm evaluates the spline model at appropriate positions to localise zeroes in the first derivative within a prespecified tolerance. However, the more detailed a given spline model, the more extrema it will generally show, and not all of them will be relevant in the determination of successive vegetation periods. The challenge is now to distinguish the extreme points that bracket a phenological cycle from those that emerge from minor features within a cycle.

The method proposed here to determine consistent extrema that bracket the vegetation periods starts with the largest maximum in a time series and places a grid along the time axis, centered around the location of the global maximum (figure 4.5). The grid size is determined by the expected period length of a vegetation cycle. In the case of the Aksu EVI data, which exhibit annual growth cycles, the expected period length is one year or 23 composites. The grid is used to determine a set of consecutive maxima that mark the middle of a number of growing periods. Subsequently, each maximum is visited in turn by the algorithm to determine the lowest minima to the left and to the right of a given maximum, called the *early* and *late minima*, respectively. If no reasonable local minimum can be identified as the early or late minimum, e.g. at the ends of a time series record, the algorithm bails out and leaves the early and/or late minimum associated with the given maximum value undefined. Any incomplete growth cycles with one or two undefined minima, as well as cycles where the time difference between the early and late minimum is no longer than half the expected period length, are eliminated. In this way, each growing cycle is bracketed by two local minima enclosing the largest local maximum between them. The algorithm works entirely without value-based thresholds, just by assessing the time pattern underlying the occurrence of the extreme points. Optionally, a value-based threshold may be applied to eliminate candidate growing cycles where the

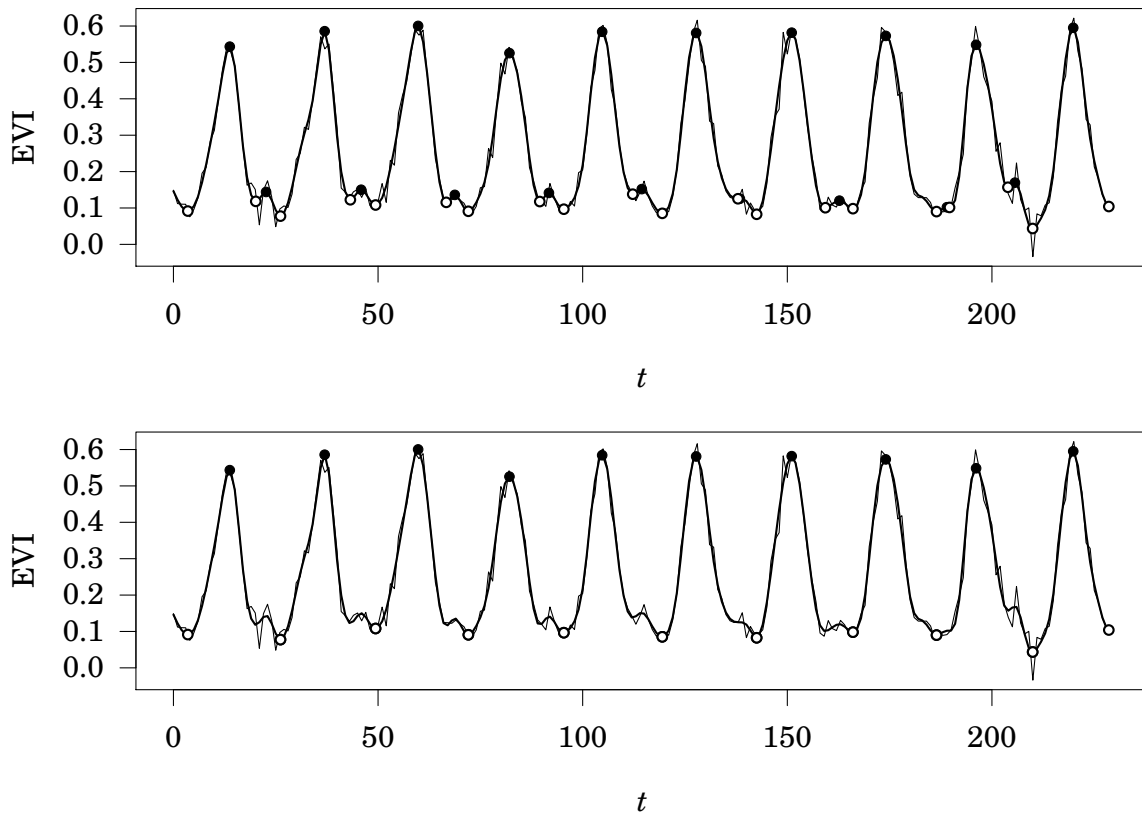


Figure 4.4: Top panel: Extreme values computed from a spline model of a MODIS EVI time series of ten years (230 observations). The model uses a rich B-spline basis with 80 knots. Bottom panel: Consistent set of minima and maxima used to define the 10 seasons.

difference in values between the local maximum and the larger of the two minimums is below a fixed threshold. The threshold setting is useful in the processing of time series of dryland environments, where there is not always a noticeable vegetation response every year during longer dry periods. The result of the overall procedure is a segmentation of a given time series into a number of growth cycles, excluding annual segments where no clear vegetation signal is observed. The consistent set of minima and maxima bracketing the growth cycles of the Aksu EVI example is shown in the bottom part of figure 4.4. The method of season identification is appropriate for time series with one growing period per year. If a semiannual cycle is present, the annual bracketed segments must be scanned for secondary maxima and minima in a similar way. Unlike other modelling techniques, e.g. piecewise logistic or Gaussian functions, where a model can only be fitted *after* assessing the structure of a time series (e.g. Zhang et al., 2003; Jönsson and Eklundh, 2004), a spline model is by itself able to represent curve structure. If, for any purpose, a different scheme for growth cycle identification is more appropriate, it can be built on top of a spline representation of a time series without substantially changing the implementation of the spline model as a whole.

Once the time series is structured into a number of annual growing periods, a set of phenological descriptors can be calculated for each year. After revision of the various phenological parameters presented in chapter 1.3, a set of 20 descriptors was selected

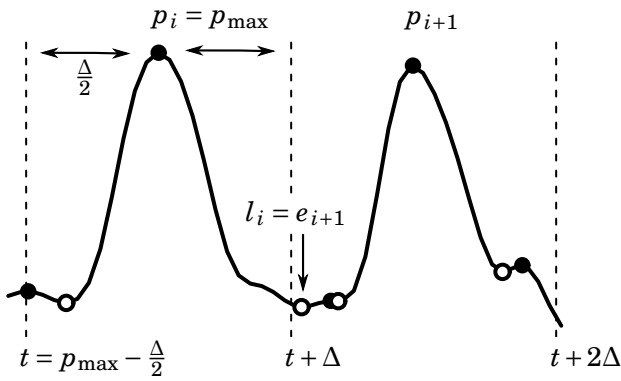


Figure 4.5: Algorithm to identify a consistent set of local minima and maxima that bracket a phenological cycle. The grid with size Δ centered on the global maximum p_{\max} is used to determine a set of annual maxima. The lowest minimum between two successive maxima p_i and p_{i+1} is at the same time the late minimum of the i th year l_i and the early minimum of the $(i+1)$ year e_{i+1} .

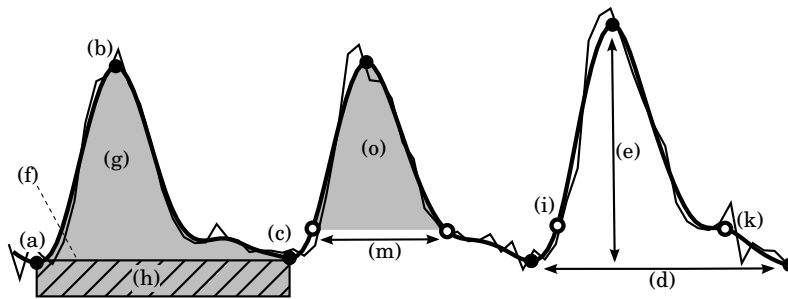


Figure 4.6: Three years of a MODIS EVI time series (thin line) from a selected location in the Aksu oasis illustrating various phenological parameters. The thick line shows the fitted smooth B-spline. The solid dots show location (day of year) and modelled value of the early minimum, peak and late minimum of a season (parameters (a), (b) and (c), respectively). Parameter (d) gives the duration in days between successive minima. The open dots at (i) and (k) mark the beginning and end of greenness, with (m), the duration of greenness, defined as the time span in days between (i) and (k). The parameters (e) and (f) give amplitude and base level (latent) EVI value, with amplitude defined as the difference between peak and latent EVI. Latent EVI is given by the average of the early and late minimum values. There are four integral parameters, the integral between two successive minima (g), the latent integral (h) and the greenness integral (o). The total integral is defined by the sum of (g) and (h). There are two additional parameters not shown above: the rate of greenup, which is defined as the slope of a line connecting the point of the onset of greenness and the annual peak value; and the rate of senescence, which is the (positive) slope of the line connecting the annual peak and the point of end of greenness.

and routines for their derivation were implemented in C++ on top of the B-spline base component. The individual parameters are illustrated in figure 4.6. The first six parameters are given by the structuring into annual growing periods: they are defined as the days of year where the early minimum, local maximum and late minimum of a growing period occurs, and the corresponding vegetation index values at these days. The *duration* of a growing period in days is determined by the time difference between the early and late minima. With each growing period is associated a *latent value* determined by the average values at the two minima. *Amplitude* is defined as the difference between

the maximum value and the latent value of a growing period. In addition, three integral quantities are determined called the *total integral*, *latent integral* and the *min.-min. integral*. The total integral is given by the total area under the spline curve between two successive minima of a growth cycle and may be taken to represent total biomass. The total integral is split into a latent and a min.-min. integral. The latent integral is described by a “box” whose width and height correspond to the duration and latent value of a growing period, respectively (figure 4.6). The total integral may be taken to represent the biomass permanently sustained at a given site (Tucker et al., 1981). The min.-min. integral is obtained by subtraction of the total and latent integrals. The beginning and end of the period of active vegetation development, or *greenness*, within an annual growth cycle were determined using a method of adaptive thresholds. Beginning (end) of greenness were defined as the day of year where the time series reaches a value equivalent to the early (late) minimum plus a threshold expressed as a percentage of the difference between the annual peak value and the early (late) minimum value. For the Aksu EVI data, this threshold was set at 20 percent. Parameters related to greenness are the time series values at the beginning and end of greenness, the duration of greenness, defined as the number of days between the beginning and end of greenness, and the greenness integral, which is the area of the spline curve over a horizontal line that intersects the curve at the average of the values at the start and end of greenness (marked (o) in figure 4.6).

4.3 Comparison to other methods and algorithms

To examine and verify the results of the spline based strategy for deriving phenological descriptors, two other available methods for calculating phenological parameters were selected to compare the results. The first method, implemented in the software `ts_phen_ind`⁴, follows the method of Reed et al. (1994) to determine the onset and offset of the growing period, which is based on digital filtering. To determine the start and end of the vegetation growing season, backwards and forwards lagged moving average (MA) curves are calculated from a time series (figure 4.3), where the time lag is chosen as one standard deviation (1SD) from the barycentre of a season. The beginning and end of a growing period is then determined by the points where the backwards and forwards lagged moving averages intersect the original time series (Ivits et al., 2008). The method of piecewise Gaussian function fitting implemented in the TIMESAT software (Jönsson and Eklundh, 2004) is chosen as the second method. Jönsson and Eklundh (2004) use a variable threshold method to determine the onset and offset of the vegetation growing season, where the beginning of a season is defined as the point on the fitted piecewise Gaussian curve where the value has increased by a certain percentage of the distance

⁴The software was written and kindly provided by Wolfgang Mehl, European Commission Joint Research Centre (JRC), Ispra, Italy.

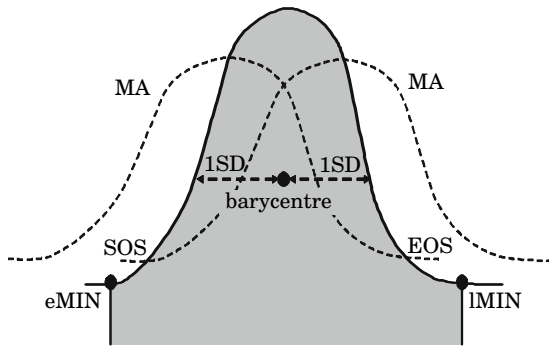


Figure 4.7: Schematic explanation of phenological parameters calculated from a time series curve (solid line) using moving averaging (MA). Start (SOS) and end (EOS) of season are determined by the points of intersection of the curve and the time delayed MA's. The early and late minima are denoted by eMIN and lMIN, respectively (after Ivits et al., 2008).

between the early minimum level and the seasonal maximum. The end of the season is defined in a similar way by using the same percentage of the distance between the late minimum level and the maximum (Jönsson and Eklundh, 2004).

The digital filtering, piecewise Gaussian and spline based models for the derivation of phenological parameters were parameterised independently of each other by visual inspection of a number of curve fits. In the case of the digital filtering based method, the model was parameterised by the author of the `ts_phen_ind` software, Wolfgang Mehl at JRC in Ispra. The Gaussian curves were fitted in order to track the upper envelope of the data (Jönsson and Eklundh, 2002). For the spline based approach a regression spline with 80 polynomial pieces and a uniform open knot vector was fitted to the original data set. The models were then used to calculate phenological parameters for the Aksu EVI data set. Unlike the moving average based method, both the piecewise Gaussian method and the spline based approach use variable thresholds (see chapter 1.3) to determine the on- and offset of greenness. In order to compare the individual approaches, the percentage thresholds for both the piecewise Gaussian and the spline based method were set at 20 percent.

4.4 Comparison sampling design

Due to the characteristics of the imaging process, remotely sensed imagery typically exhibits dependencies among neighbouring pixels, a phenomenon known as spatial autocorrelation (Craig, 1979). When using analysis methods on image data with a statistical basis, this spatial autocorrelation must be taken into consideration, since statistical analysis typically requires a set of data that comprise independent samples. It was shown that spatial autocorrelation influences the results and accuracy assessment of land cover classifications obtained from remotely sensed data (Labovitz and Masuoka, 1984; Congalton, 1988; Dobbertin and Biging, 1996; Hammond and Verbyla, 1996). If the distance between a number of sampled pixels is lower than a specific autocorrelation distance inherent in a data set, e.g. when sampling homogeneous blocks of pixels (Hammond and Verbyla, 1996), a sample is more likely to contain similar features due to the autocorrelations among the sample elements. As a consequence, the variability in the data is underestimated based on the sample. In assessments of land cover classification accu-

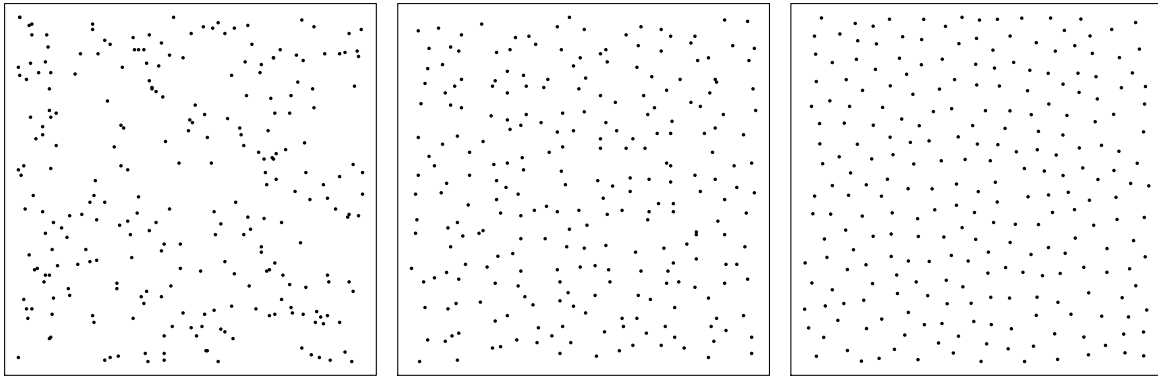


Figure 4.8: Three types of spatial sampling. The horizontal and vertical coordinates of the sample points in the left panel have been chosen randomly within the width and height of the area. The central panel shows a jittered grid: the area is divided into a grid, and one or more points are randomly selected from every grid cell. Poisson disk sampling is shown in the right panel.

racy, this may lead to overly optimistic accuracy reports (Hammond and Verbyla, 1996). In the same way, the assessment and comparison of the phenological features may be affected by an optimistic bias if spatial autocorrelation is not taken into account. A good sampling design should thus fulfil the following three criteria:

- * To get a fair sample where no parts of an area are over- or underrepresented, the sample points should be uniformly distributed over the entire sampling area, no clumping or gaps should occur.
- * Some minimum distance should be maintained between the sample points to avoid contamination of the sample by autocorrelation.
- * Within the limits set by the former two criteria, a sample should be random.

A common strategy is to use *uniform random sampling*, where the horizontal and vertical coordinates of the points in a sample are chosen randomly from a given coordinate range, with the result that the number of sample points per unit area follows a Poisson distribution (O’Sullivan and Unwin, 2010), as in the leftmost panel of figure 4.8. Uniform random sampling leads to an irregular spatial pattern, and there may be considerable gaps where no sampling points are placed. Then again, there are clumpy spots where two or more sample points are at close distance, which means that a uniform random sample is susceptible to autocorrelation bias. Besides, if the gaps become too extensive, a sample may no longer be representative for a given area as a whole. Stoffels et al. (2011) used a regular sampling grid to obtain training samples for the parameterisation of a maximum likelihood classifier. In Gaussian maximum likelihood classification, per-class variance-covariance matrices are used as parameters, so Stoffels et al. (2011) set the size of the sampling grid well in excess of the autocorrelation distance to avoid the underestimation of the per-class variances and covariances based on the training samples. Sampling on a regularly spaced grid, where the distance between sample points in the horizontal and vertical directions is set constant to avoid autocorrelated samples, is diametrically

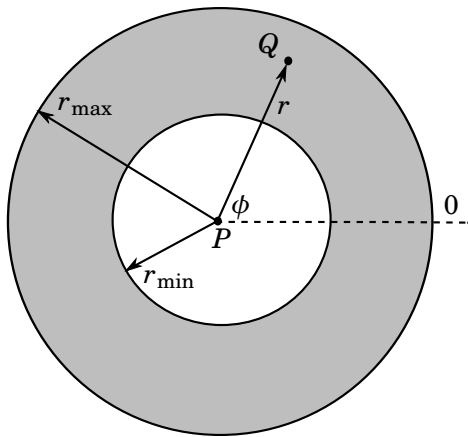


Figure 4.9: Generation of a new sample point in Poisson disk sampling. The new point Q is selected from the annulus (the disk, shown in grey) surrounding an already existing point P . A point on the disk is addressed by polar coordinates (ϕ, r) , $r \in [r_{\min}, r_{\max}]$. The process is repeated by iterating over a set of existing points until a prescribed sampling density is reached (after Tulleken, 2008).

opposite to uniform random sampling and contradicts the above requirement of randomness. The former technique is too regular, the latter is too messy. An alternative method that, contrary to a regular grid, retains a somewhat random sampling pattern while at the same time mitigating autocorrelation bias is the *jittered grid* method, also known as *jittered sampling* (Suffern, 2007). In jittered sampling, the total area to be sampled is divided into a number of grid cells by subdividing the horizontal and vertical axes into regular intervals. Subsequently, a random sample is placed within each cell of the grid. The sampling density of a jittered grid is controlled by the number of grid cells and the number of uniform random samples to place in each cell. The jittered grid sample shown in the middle panel of figure 4.8 was created using 16 divisions along the horizontal and vertical axes to partition the entire sampling area into 256 grid cells, placing one random sample per cell. While jittered sampling leads to a more uniform random arrangement of the sampling points, it does not completely avoid autocorrelated sample elements, since the positions of the samples within the cells are completely arbitrary and there is no hard limit on the minimum distance between individual sample points. A method that fulfils all three of the above criteria for a good sampling design is the *Poisson disk sampling* method (Tulleken, 2008). It is used in computer graphics to generate textures or randomly place objects whenever the objects have to be tightly packed together; but no closer to each other than a specified minimum distance. As shown in the rightmost panel of figure 4.8, Poisson disk sampling leads to a well spaced, random set of points.

The basic idea of Poisson disk sampling is to generate new points around existing points, so as not to disturb the minimum distance requirement. To this end, a new random point is selected from an annulus surrounding an already existing point using a polar coordinate system with its origin at the existing point (figure 4.4). The extent of the annulus is defined by the minimum distance r_{\min} and a maximum distance r_{\max} . The algorithm for Poisson disk sampling proposed by Tulleken (2008) starts by creating a grid similar to a jittered grid with a cell size of $(r_{\min}/2)$, where r_{\min} is the minimum distance between the sample points. In the next step, a first point is randomly chosen

and a number of k new points are generated from its annulus. For every newly generated point, the grid is used to check for other points that are too close to the new point, in which case the point is rejected. The algorithm now iterates over the set of existing points by picking points from the set at random, trying to add at most k points to its annulus until there are no more points that can be added without violating the minimum distance requirement. The parameter k is used in addition to the minimum point distance r_{\min} to control the sampling density. The larger k , the more sample points are generated within limits of r_{\min} .

In order to compare the digital filtering, piecewise Gaussian and spline based approaches for deriving phenological descriptors, Poisson disk sampling was implemented in Python following (Tulleken, 2008), and used to obtain well spaced paired samples of picture elements from the phenological parameter images. The minimum allowed distance between sample points r_{\min} was set to 10 pixels to avoid autocorrelation bias (Craig, 1979). The outer diameter of the annulus r_{\max} was set at twice the minimum distance (Tulleken, 2008). For the parameter k , a value of 5 was used.

5 Results and discussion

To verify the spline based approach for the derivation of phenological parameters, the resulting phenological parameters from the uniform regression spline model discussed in chapter 4.2 were compared to the results obtained from the two alternative methods (chapter 4.3). Figure 5.1 shows three examples of phenological descriptors derived from the regression spline model for the year 2005 based on the Aksu data set. The descriptors shown are the start of the growing season (parameter (i) in figure 4.6), the season length (parameter (m) in figure 4.6) and the integral of EVI over the growing season (parameter (o) in figure 4.6). To check the plausibility of the phenological descriptors derived from the regression spline model, a subset of modelled phenological parameters was compared to the results from the other two models (moving average filtering and Gaussian curve fitting). In regard of the temporal variability present in the Aksu EVI data, comparisons were done for three individual years within the period covered by the data: 2002, 2005 and 2009. The first and last year of the entire period were ignored to avoid disturbances due to edge effects in the determination of the phenological parameters. The subset of descriptors shown in figure 5.1 comprises three most important phenological descriptors that are derived equivalently by all three phenological assessment methods. The timing of the growing season and its length are important features e.g. in global change research (White et al., 2009), while the seasonal integral can be used as a proxy to assess the amount of biomass produced in a given year (Tucker et al., 1981). To include only picture elements from vegetated areas for which the determination of phenological descriptors is meaningful, pixels with an amplitude in a given year determined by the regression spline model (parameter (e) in figure 4.6) below 0.2 EVI units were excluded from the analysis. The constant EVI threshold of 0.2 was chosen by inspecting the parameter images of the seasonal amplitude. The agricultural areas in the Aksu area could well be delineated from the surrounding desert areas by using the aforementioned threshold. Under the restriction of the amplitude threshold, the Poisson disk sampling yielded a set of 512 sample points for the year 2002, for the 2005 data, 547 samples were obtained, and for 2009, sampling at the prespecified point density yielded a sample size of 577. Since the sampling density in Poisson disk sampling is constant, this indicates an approximate increase in the vegetated area of 13 percent between 2002 and 2009. The three equivalent phenological parameters derived from the three time series analysis methods were compared on a point to point basis by paired random samples from the imagery containing the phenological descriptors for 2002, 2005 and 2009. Table 5.1 shows summary statistics for the individual samples and parameters. In the mean, the start of season dates were predicted within a window of six days by all three methods. The greatest discrepancy of six days occurred between the spline and the moving average based model in 2009. The most consistent estimates for the start of the growing seasons occurred in 2005, where the difference in means for the starting times was only two days.

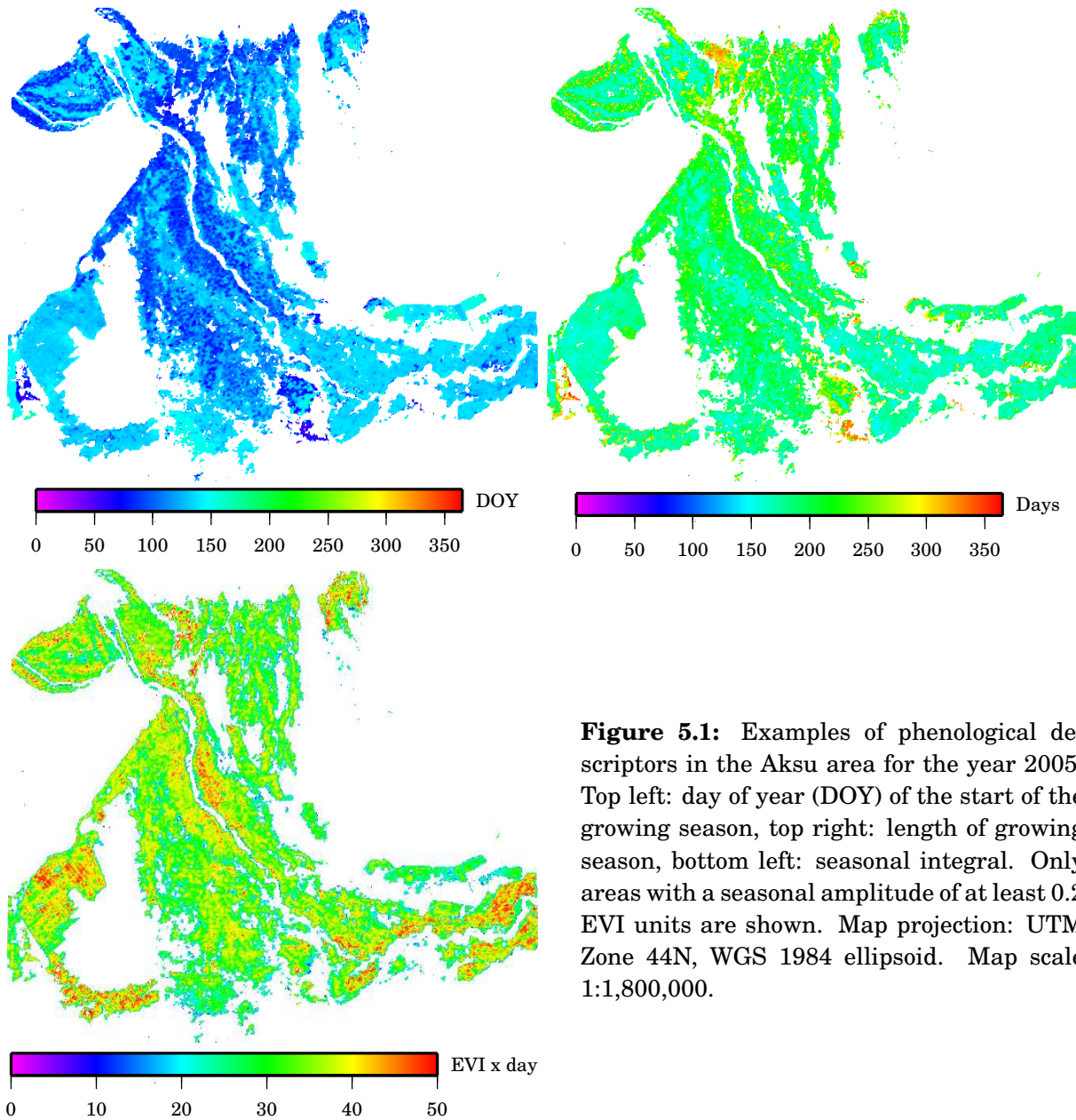


Figure 5.1: Examples of phenological descriptors in the Aksu area for the year 2005. Top left: day of year (DOY) of the start of the growing season, top right: length of growing season, bottom left: seasonal integral. Only areas with a seasonal amplitude of at least 0.2 EVI units are shown. Map projection: UTM Zone 44N, WGS 1984 ellipsoid. Map scale 1:1,800,000.

Despite the large sample sizes, a Kruskal-Wallis rank sum test (Spiegel and Stephens, 2008) performed on the estimates of the start of the growing season in 2005 was not significant (p-value 0.3729). The null hypothesis of the test states that the estimates for the start of the growing season are equal in their central tendency. The differences between the methods in 2002 and 2009 were significant according to the Kruskal-Wallis statistics. For the other two parameters, growing season length and the seasonal integral, Kruskal-Wallis statistics indicated significant differences in all years. Differences in the mean growing season lengths for the three models ranged from about one week in 2002 to as much as two weeks in 2009, but the differences were always below, or at least of the order of the standard deviations for the estimates of a given method in a given year. Of all three methods, the moving average based method showed the least intra-annual variability of season length estimates. Intra-annual variability of season

		2002			2005			2009		
		MA	Gauss.	Spline	MA	Gauss.	Spline	MA	Gauss.	Spline
Start	\bar{y}	123	121	119	121	121	119	130	125	124
	$\hat{\sigma}$	14.66	25.42	24.97	14.65	20.70	20.41	15.67	21.08	17.22
	\hat{c}_v	0.119	0.210	0.209	0.121	0.171	0.171	0.121	0.169	0.139
Length	\bar{y}	183	179	190	189	182	195	187	194	202
	$\hat{\sigma}$	16.73	35.70	38.41	15.36	31.79	38.98	14.87	30.07	31.31
	\hat{c}_v	0.091	0.199	0.202	0.081	0.175	0.199	0.080	0.155	0.155
Integral	\bar{y}	34.82	40.58	31.91	36.42	43.54	32.80	35.54	44.45	33.01
	$\hat{\sigma}$	7.45	11.83	7.78	7.92	12.01	8.00	6.59	9.98	7.75
	\hat{c}_v	0.214	0.292	0.244	0.218	0.276	0.244	0.185	0.225	0.235

Table 5.1: Summary statistics of the samples obtained from the three phenological data sets derived from digital filtering (MA), Gaussian curve fitting (Gauss.) and spline curve fitting (Spline). Samples are taken for three years at the beginning, middle and end of the time series. The summary statistics for the start of the growing season (Start), the length of the growing season (Length) and the seasonal integral (Integral) are the sample mean (\bar{y}), the standard deviation ($\hat{\sigma}$) and the coefficient of variation (\hat{c}_v). Season start and length are given in days, integrals are given in units of (EVI \times day).

lengths for the piecewise Gaussian and spline approaches was of the same order. For the seasonal integrals, the spline based approach provided the lowest estimates on average, consistently over all three years that were considered. The mean seasonal integral values of the moving average model were always in between the results of the spline model and the piecewise Gaussian model, which constantly yielded the highest average seasonal integrals.

5.1 Start of season

Figure 5.2 shows three scatterplots of the start of season estimates for the moving average (MA) based model for the years 2002, 2005 and 2009. The estimates of both models do not scatter much around the 1:1-line, however, when the growing season of a sample begins relatively early, the MA model has a tendency to provide slightly later estimates compared to the spline based model. As can be seen in figure 5.3, the start of season estimates for the piecewise Gaussian curve fitting method in comparison to the spline model line up more consistently along the 1:1-line. There is a slight scatter around the line, but no indication of a general bias. The fact that the spline model agrees better with the piecewise Gaussian method than with the moving average based approach is in accordance with the structure of the three approaches. In the piecewise Gaussian and spline models, the start of season is determined in a similar way by a variable threshold method. In both methods, the start of the season is determined as the point where the EVI increases over a threshold that amounts to 20 percent of the annual amplitude. Since the threshold parameter was set equivalently in both models, it can be expected that the

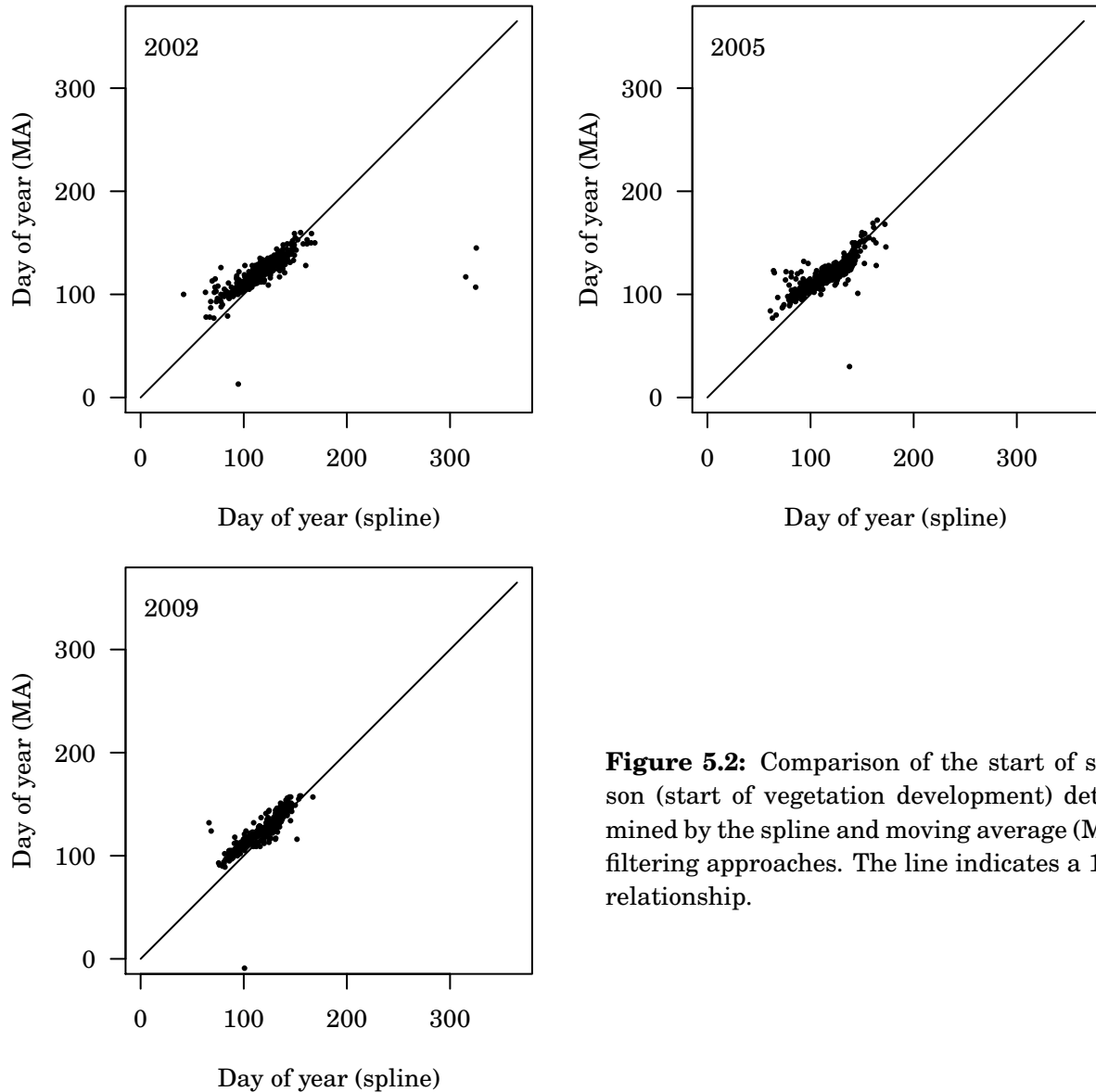


Figure 5.2: Comparison of the start of season (start of vegetation development) determined by the spline and moving average (MA) filtering approaches. The line indicates a 1:1-relationship.

estimates of these two models are more in agreement as compared with the moving average method, which uses a completely different approach to determine the start of the season. However, the scatterplots of figures 5.2 and 5.3 indicate that the start of season estimates of all three modelling approaches are consistent.

5.2 Length of season

A less consistent picture is shown by the comparison of the estimates for the length of the growing season between the moving average, piecewise Gaussian and spline based approaches to phenological analysis. The point clouds in the scatterplots of season lengths derived from moving averaging (MA) vs. season lengths obtained from the spline model are considerably tilted against the 1:1-line (figure 5.4). As already noticed from the summary statistics, the moving average estimates of season length are less variable than the spline estimates. For some sample elements, the spline estimates indicate a longer sea-

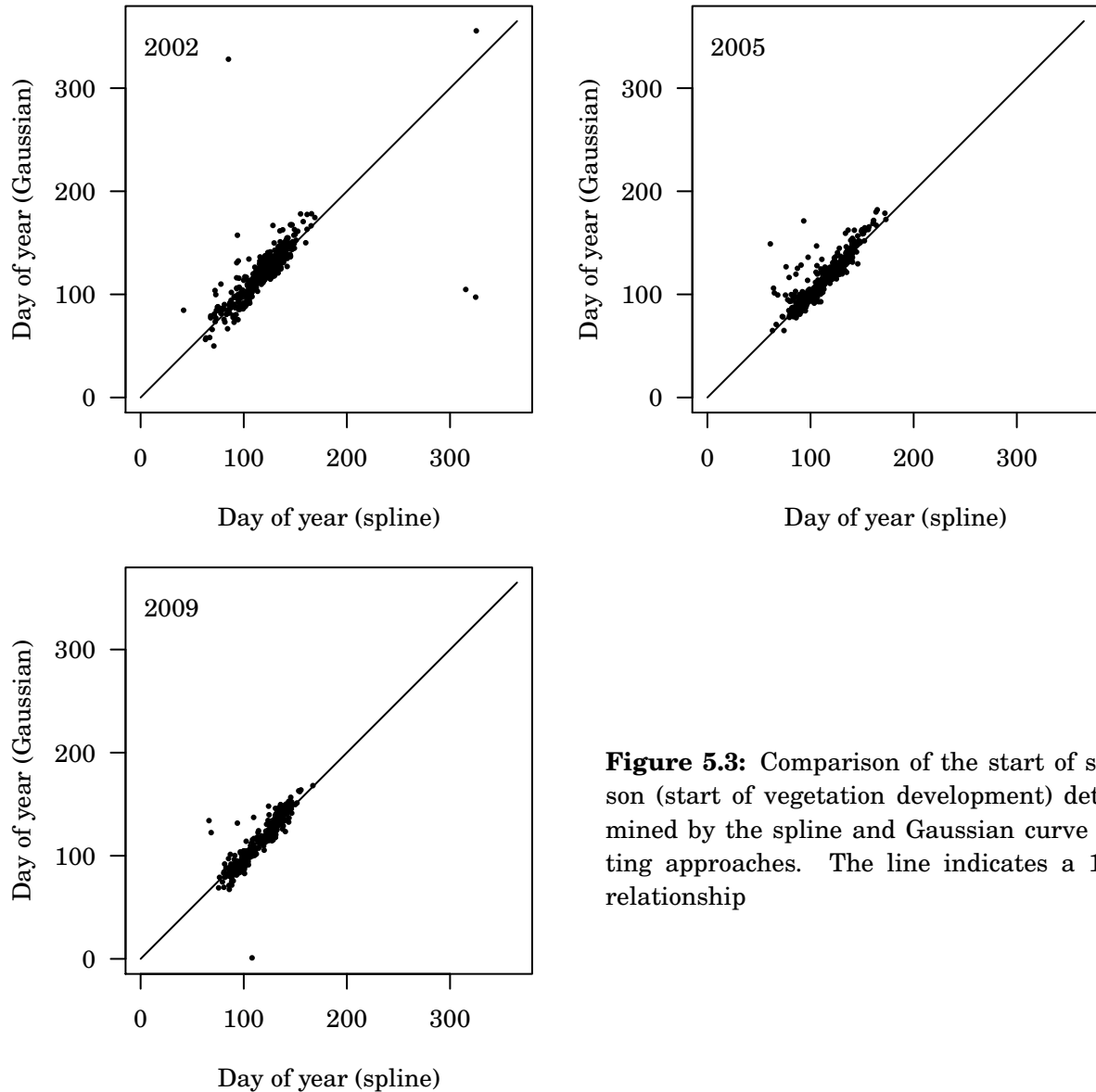


Figure 5.3: Comparison of the start of season (start of vegetation development) determined by the spline and Gaussian curve fitting approaches. The line indicates a 1:1-relationship

son than the MA estimates, but there is no evident relationship. Also, the scattering in the point cloud is noticeably less in 2009 than in 2002 and 2005. The season lengths from the piecewise Gaussian model, again, are more in alignment with the 1:1-line (figure 5.5). The point cloud is not tilted as with the MA estimates, but there is a certain amount of scattering that affects both approaches. Again, the tighter packing of points along the 1:1-line for the piecewise Gaussian model as opposed to the moving average model indicates that the two curve fitting methods are more akin to each other as opposed to the moving average method, concerning the way of timing the key phenological transitions. While the curve fitting methods use adaptive thresholds, the MA method uses backwards and forwards lagged moving averages. However, the length of season estimates of the three methods differ considerably more between the methods than the start of season estimates discussed in chapter 5.1. The season length is defined as the number of days between the start and end of the season. Since the start of season estimates are well in

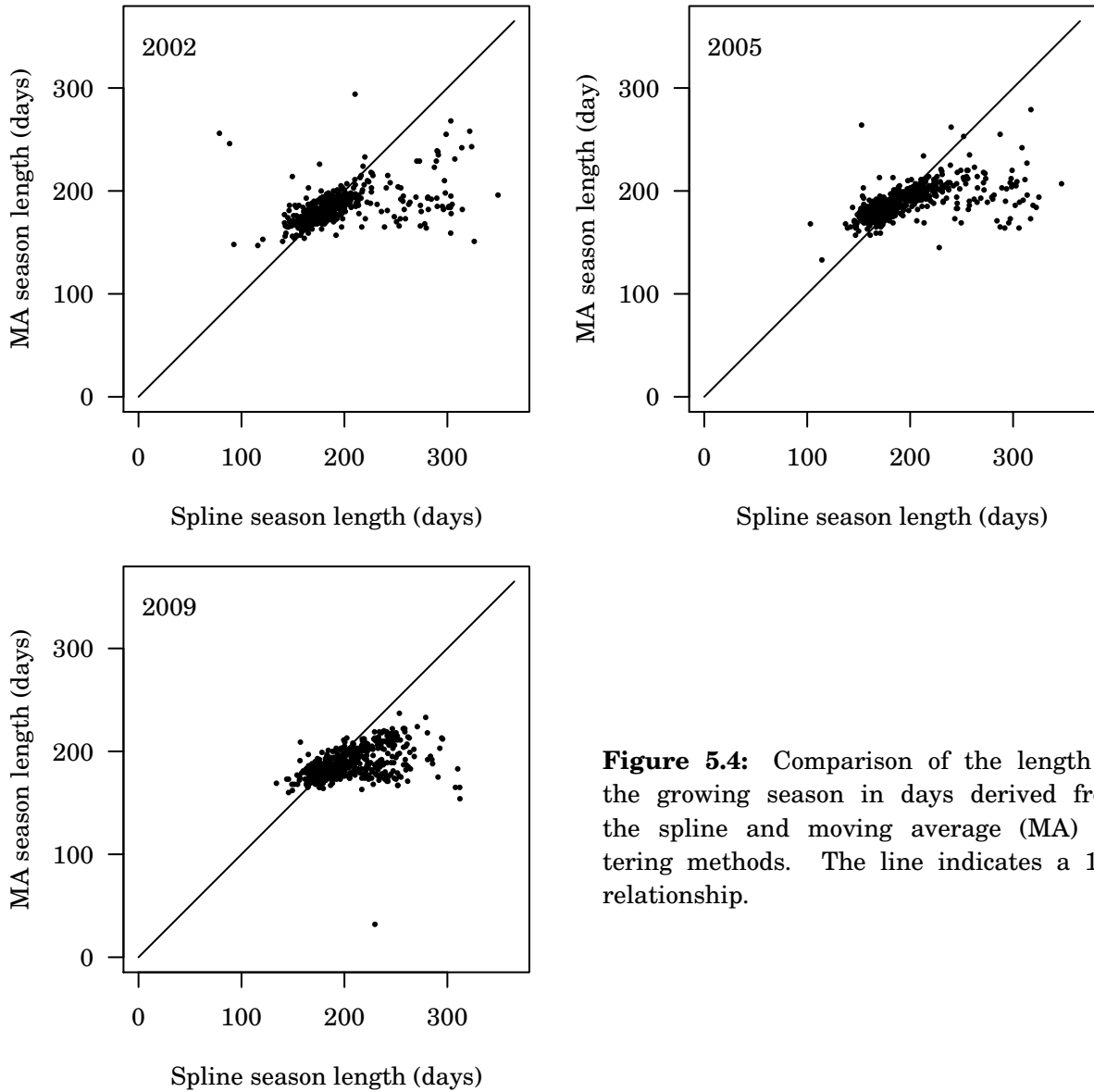


Figure 5.4: Comparison of the length of the growing season in days derived from the spline and moving average (MA) filtering methods. The line indicates a 1:1-relationship.

agreement between the three tested methods, the discrepancies in the length of season estimates can only be attributed to differences in the determination of the end of the season. Figure 5.6 shows a section of two years of the original EVI time series for a selected location in the Aksu area in combination with the start and end of growing season estimates for all three methods for the two years. The profiles typically show a distinct, sharp increase in EVI values at the beginning of the growing period. Accordingly, there is not much difference in the timing of the start of season between the three methods. All three approaches are able to detect the edge in the profile equally well. At the end of the growing period, the profiles show more variability in shape. There is a rapid drop in the EVI values followed by a slowly decaying shoulder that may be more or less pronounced in individual years. The end of season estimates for the three methods are more in agreement in years when the shoulder is less pronounced, as for example in the first of the two years in figure 5.6. In the second year, where the shoulder is more pronounced,

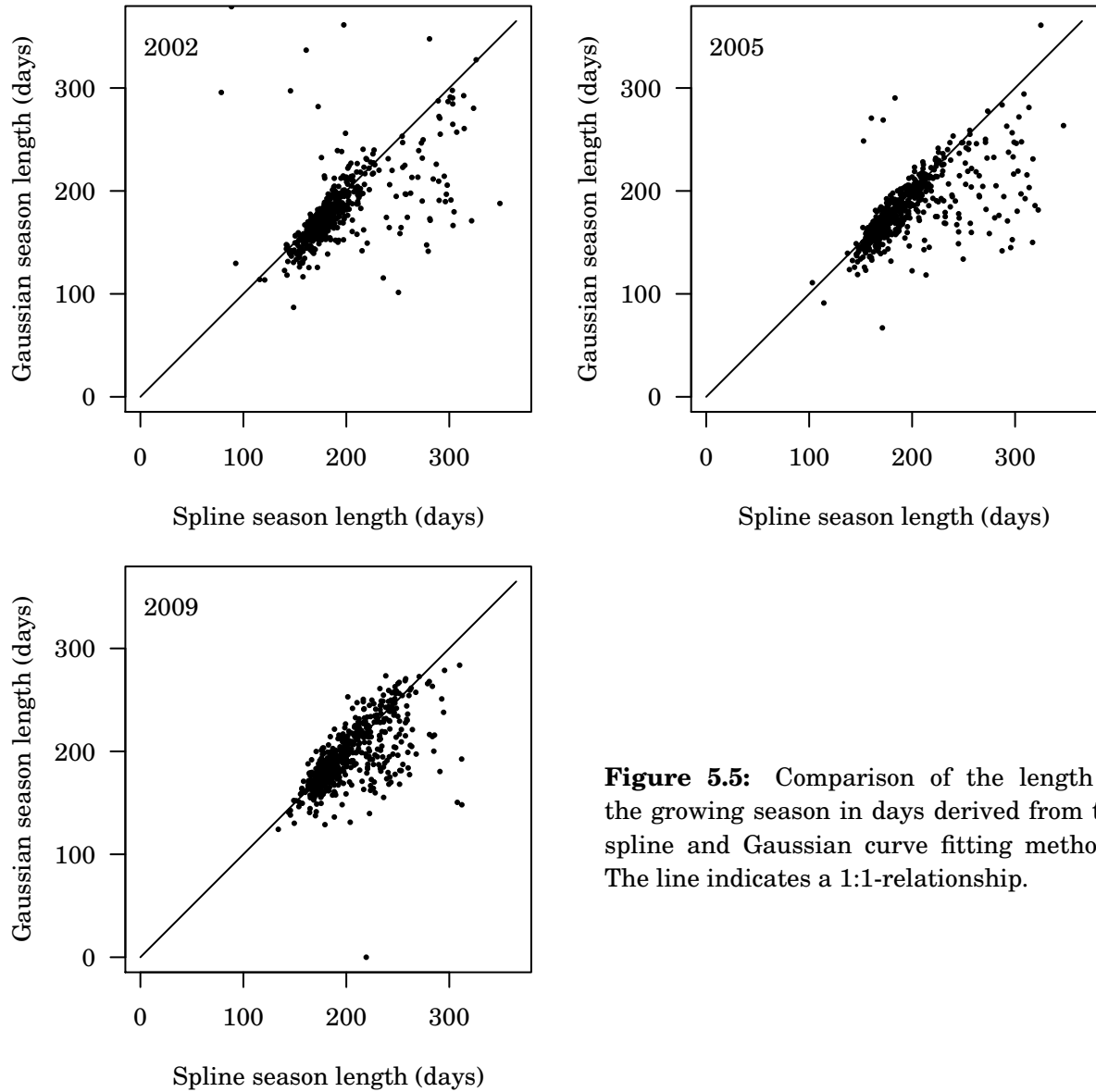


Figure 5.5: Comparison of the length of the growing season in days derived from the spline and Gaussian curve fitting methods. The line indicates a 1:1-relationship.

the results of the end of season estimates differ. Under the assumption that the key phenological transition at the end of the season is indicated by the sharp drop in EVI, the spline and Gaussian methods provide better estimates of the end of the season, and thus, of the season length, compared to the moving average method. However, the scattering in the season length estimates between the methods indicates that the determination of the end of season in the case of subtle profiles as in figure 5.6 is difficult for all three models.

5.3 Seasonal integrals

As can be already seen from the summary statistics in table 5.1, the spline estimates of the seasonal integrals were lower than the estimates of the moving average and piecewise Gaussian approaches. Figure 5.7 shows scatterplots of the seasonal integral derived from the moving average method versus the seasonal integrals computed from the spline

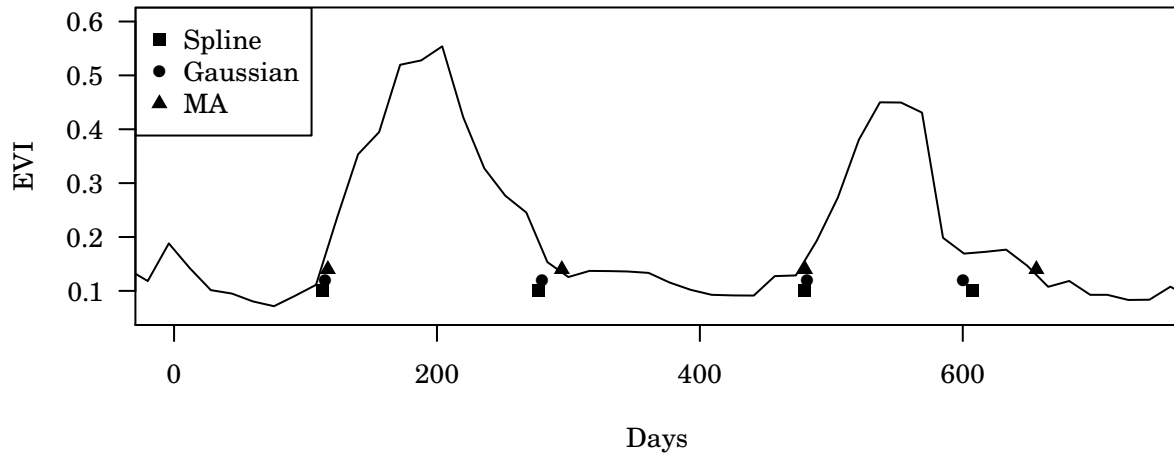


Figure 5.6: Comparison between the start and end of season estimates of the three methods for a section of the time series of two years from a selected location. The curve shows the original time series, the symbols denote the timings of the start and end of season estimates as determined by the three methods. The slowly decaying shoulder at the end of the season is responsible for the scattering in the length of season estimates. The horizontal axis is labelled in terms of days after the first day of year in 2008.

model. The difference in the means of the estimates apparent in table 5.1 turns out to be due to a consistent bias. The point clouds for all three years are tightly packed and stretched along the 1:1-line, but are shifted slightly upwards parallel to the line. This indicates a relationship between the moving average estimates and the spline estimates of the seasonal integrals that can be described by an additive constant. The seasonal integrals computed from the moving average method are, on average, slightly higher by a certain amount than the integrals calculated from the spline model. Thus, regarding seasonal integrals, the results from the moving average and the spline model may be considered as equivalent up to a constant. Since there is such a clear relationship, it is not likely that the differences in the moving average and spline based integrals can be explained by the rather irregular differences in season length between the two approaches. The fact that the variability present in the end of season estimates obviously does not spread to the estimates of the seasonal integrals indicates that a typical EVI curve of the investigated Aksu data set is rather flat and low at the end of the growing season, so that an increased interval of integration due to a longer season does not influence the values of the integrals much. The seasonal integrals of the spline model compared to those computed from the piecewise Gaussian curve fit are shown in figure 5.8. As is the case with the seasonal integrals obtained from moving averaging, the piecewise Gaussian integrals are almost always higher than the integrals computed from the spline model, but the discrepancy is greater compared to the moving average method. Furthermore, the discrepancies are not approximately constant, but tend to increase with increasing integral values.

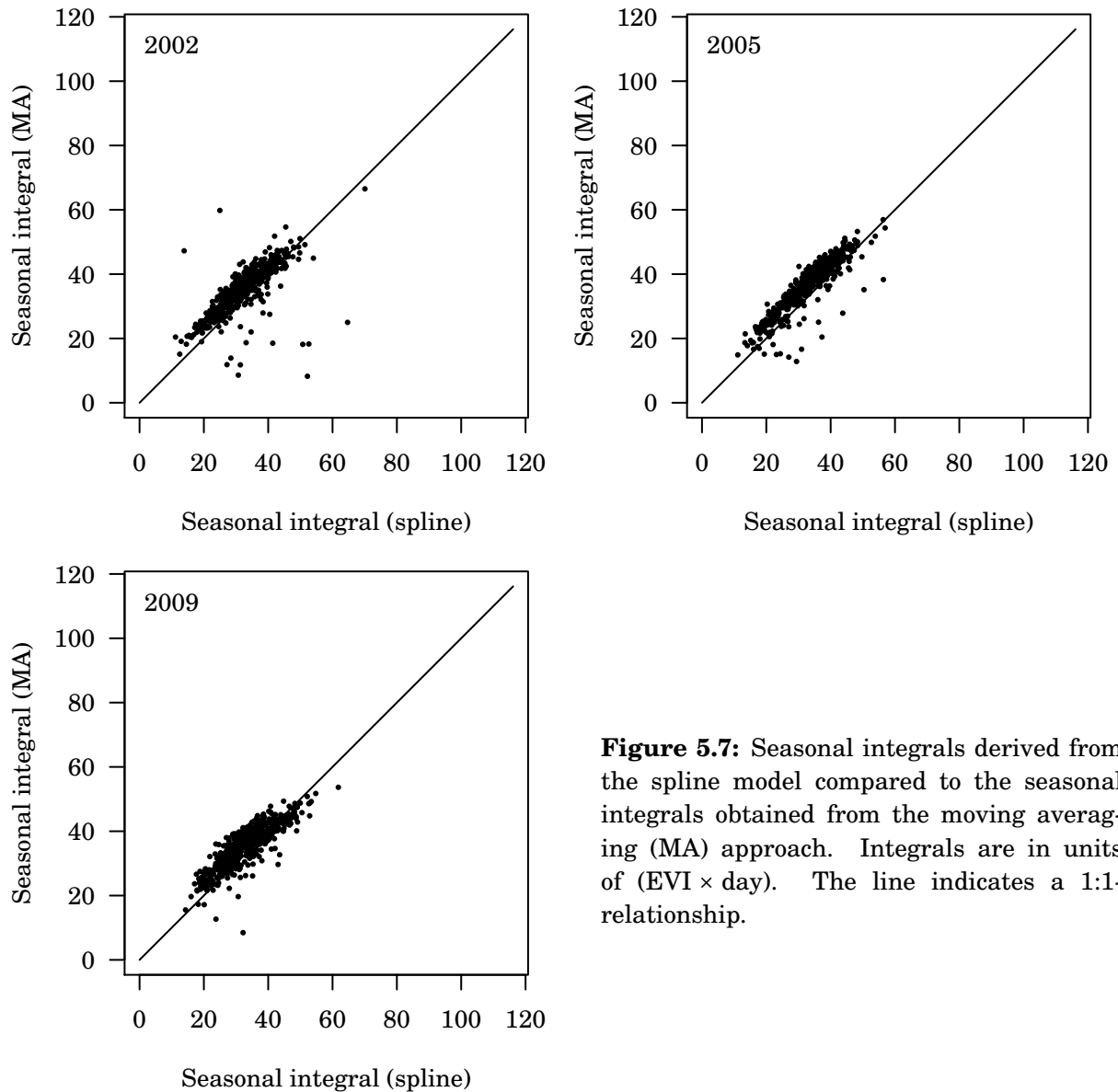


Figure 5.7: Seasonal integrals derived from the spline model compared to the seasonal integrals obtained from the moving averaging (MA) approach. Integrals are in units of $(EVI \times \text{day})$. The line indicates a 1:1-relationship.

5.4 Trends

So far, it was established that, with the exception of the season length, the selected phenological parameters obtained from the three models agree well for instantaneous data sets of individual years. Since phenological parameters are typically used to assess vegetation dynamics, the phenological markers should also agree in their trends before the models may be regarded as fully consistent. In chapter 1.4 it was mentioned that the dynamics of land use in the Aksu area can be studied in a spatial context by analysing the trends in phenological parameters over the course of time. Figure 5.9 shows the regression coefficients of linear trends over 8 years from 2002 to 2009. The trends were computed based on the phenological parameters of the spline model using the TimeStats software (Udelhoven, 2011). The regression coefficients give the annual rates of change for the individual phenological parameters. For the start and length of the growing season, the regression coefficients are in units of $(\text{days}/\text{year})$. The regression coefficients for

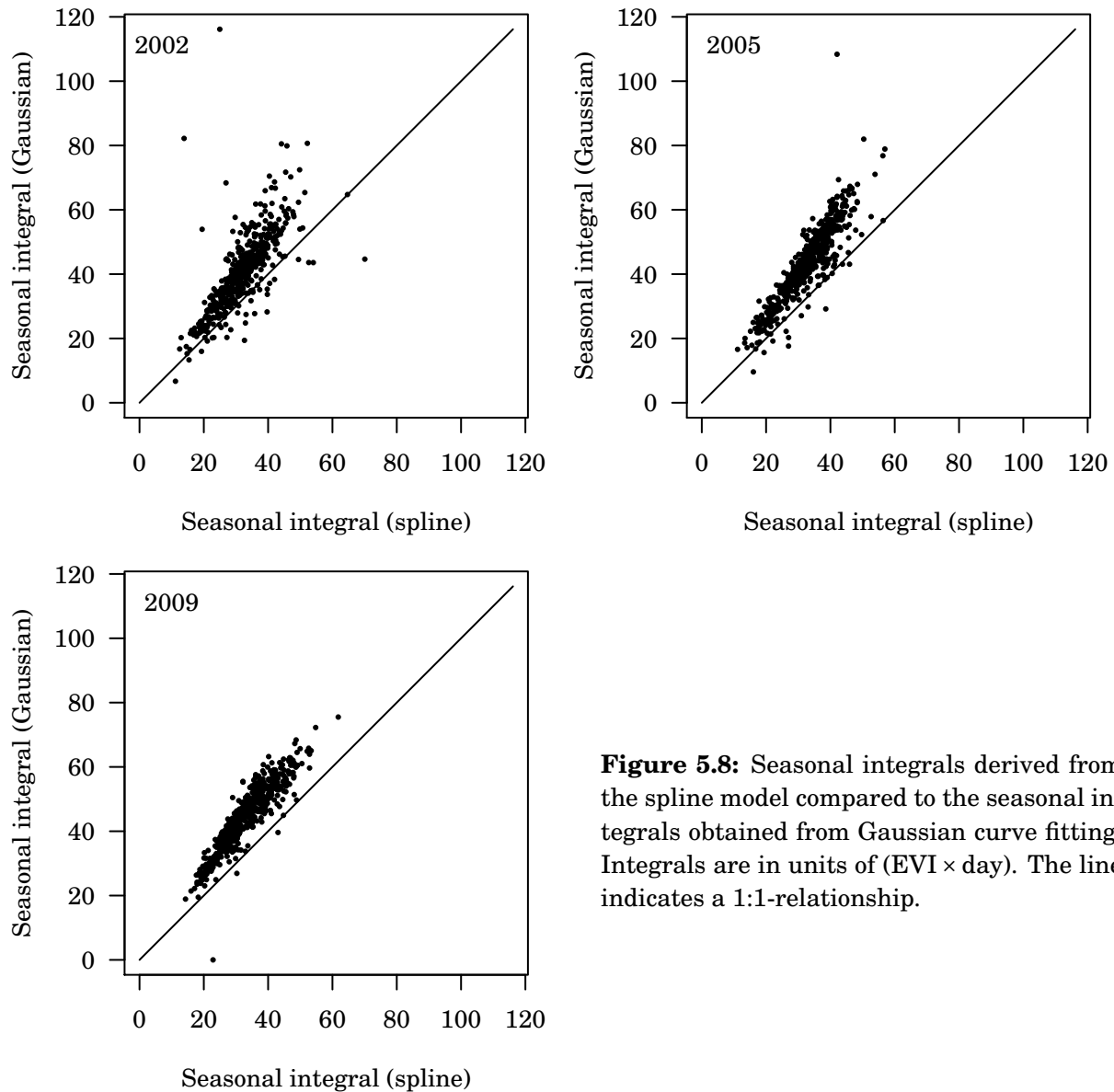


Figure 5.8: Seasonal integrals derived from the spline model compared to the seasonal integrals obtained from Gaussian curve fitting. Integrals are in units of $(EVI \times \text{day})$. The line indicates a 1:1-relationship.

the seasonal integral are in units of $[(EVI \times \text{day})/\text{year}] \times 10^3$. As can be seen in the bottom left panel of figure 5.9, the areas with the newly created agricultural fields mentioned in chapter 1.4 (marked “B”) show relatively large positive regression coefficients for the seasonal integral, indicating that the vegetation amount over time is increasing. In contrast, the regression coefficients for the apparently stable areas marked “A” in figures 1.3 and 5.9 are moderately negative. In an actual trend analysis, statistical significance testing would have to be done to establish the validity of the trends indicated by the regression coefficients. It is, however, not within the scope of this thesis to analyse and interpret the trends in the phenological parameters in their environmental context. The regression coefficients are merely used to assess whether the results of the three modelling approaches agree in their trends.

Figure 5.10 shows a series of scatterplots to compare the regression coefficients obtained from TimeStats (Udelhoven, 2011) between the moving average (MA), piecewise Gaus-

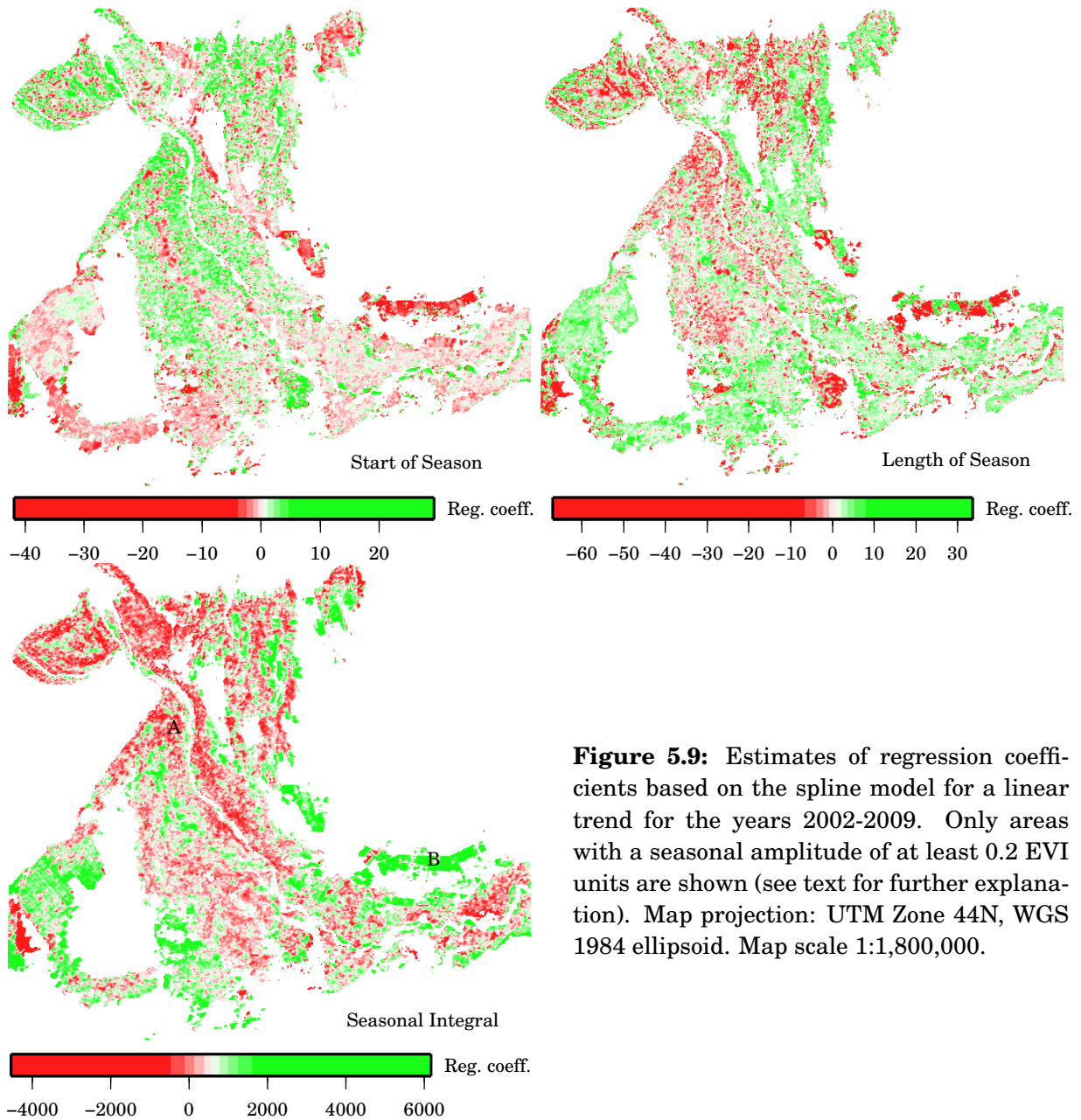


Figure 5.9: Estimates of regression coefficients based on the spline model for a linear trend for the years 2002-2009. Only areas with a seasonal amplitude of at least 0.2 EVI units are shown (see text for further explanation). Map projection: UTM Zone 44N, WGS 1984 ellipsoid. Map scale 1:1,800,000.

sian, and spline based models. The sample points were taken from the random Poisson disk samples for the year 2009 discussed earlier in this chapter. The scatterplots show pairwise comparisons of the spline based model with the two alternative methods. The regression coefficients for the start of season trend are shown in the top row of figure 5.10, the middle row shows the two scattergrams for the trends in season length. Trends for the seasonal integral are shown in the bottom row.

The scatterplots show that, with the exception of the season length, the linear trends of the selected phenological parameters are consistent between the three methods, too. The regression coefficient estimates generally agree well for the start of season. The point cloud formed by the moving average vs. the spline based start of season regression coefficient estimates is slightly tilted against the 1:1-line, suggesting that the moving average based approach provides marginally lower positive or negative trend estimates

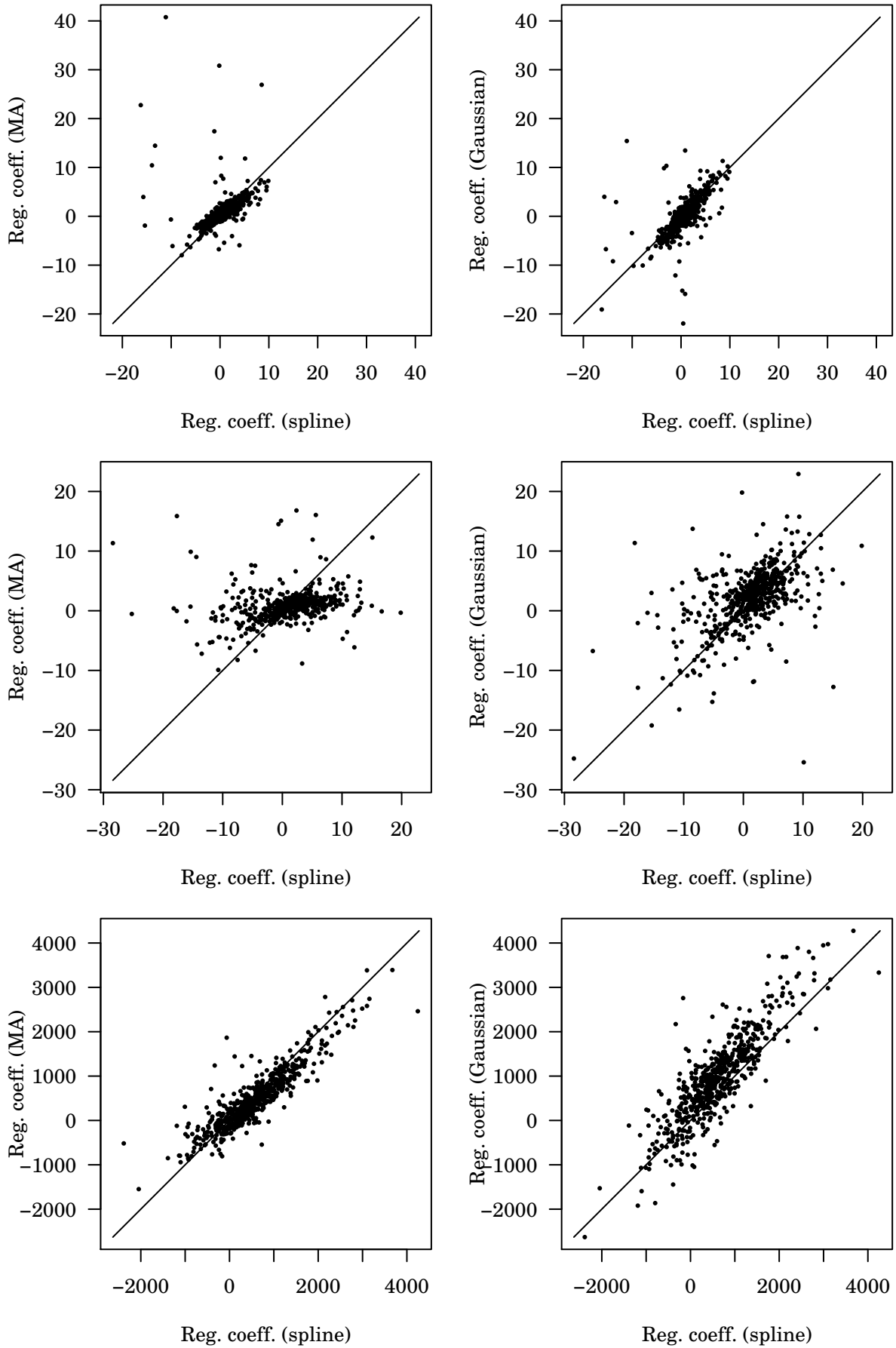


Figure 5.10: Comparison between regression coefficients of the 8-year trends. Top row: start of season, middle row: season length, bottom row: seasonal integral.

in comparison to the spline based method. The directions of the trend estimates of the seasonal integral are also in good agreement, but the scattering between the piecewise Gaussian and spline estimates is larger compared to the relationship formed by the spline estimates with those of the moving average based method. In contrast, the directions of the trend in season length do not agree well between the methods, which is a result of the difficulties encountered by the models in timing the end of the growing season.

In summary, the spline based approach to determine phenological descriptors is consistent with the two alternative methods for the test data set. Consistency exists in both the descriptors themselves and the directions of their linear trends. The variations in the length of season estimates present in all three models are a consequence of the shape of the phenological profiles recorded in the test data set.

6 Conclusions and outlook

In this thesis, a general framework for the analysis of remotely sensed time series based on polynomial spline models was developed, with a focus on the use of these models to derive phenological descriptors from remotely sensed data. Spline models provide the advantage of a data driven, locally controlled fit that does not anticipate a certain shape of a phenological profile. Furthermore, splines have favourable mathematical properties that can be exploited to efficiently analyse a modelled curve, e.g. to obtain its minimum and maximum values. Most important, the derivatives and integrals of spline curves are themselves splines and can be analytically derived from the coefficients of a fitted spline model. Thus, splines are self contained analytical models for remotely sensed time series that can be used with raw time series data and do not generally require preprocessing of the data, e.g. for noise reduction.

The utility of the implemented spline models in deriving phenological descriptors from remotely sensed time series was tested and verified by applying a particular spline model to MODIS EVI time series at a spatial resolution of 250 m from the Aksu agricultural area in northwestern China. From a visual inspection of the curve fits of different regression and smoothing spline models to a number of time series samples from the data set, it was concluded that a simple regression spline provided a good model for the data. To assess the plausibility of the spline method to derive phenological parameters, the results of the spline model were tested against the results of two other well established models to derive phenological descriptors. The comparisons showed that all three models were consistent in the prediction of the phenological metrics considered. Thus, the spline based method to derive phenological parameters performed equally well compared to the other two methods. An advantage of the spline model is that its parameters, such as the number of polynomial pieces that comprise a curve, are intuitive and can be easily manipulated by an analyst in a trial-and-error technique to find a well fitting spline model for a given data set.

A validation of the results derived from the phenological analysis of remotely sensed time series imagery by ground truthing is generally difficult. Time series imagery products such as the MODIS collections and the SPOT VEGETATION archive typically have a coarse spatial resolution between 250 m and 1 km. It is thus prohibitive to compare ground point measurements directly to pixel values of remotely sensed data products because of the scale difference (Liang et al., 2002; Liang, 2004). Such comparisons could only be made if an observed surface was large and homogeneous. In a spatially heterogeneous environment, such as the agricultural plots in the Aksu oasis, a direct comparison between point measurements and 250 m MODIS pixels is not appropriate. The same is true for the collection of reference information, e.g. if a set of phenological descriptors is to be used for land cover classification. The more inhomogeneous an observed area in

relation to the ground resolution of a remote sensing instrument, the less likely it will be that a scene observed by that instrument will contain pure pixels that are occupied by a single land cover type. Instead, the remotely sensed time series as well as any phenological descriptors derived from it will be the product of a mixture of different land cover types.

The problem of synergy effects (Cracknell, 1998) due to mixed pixels may be addressed by semiquantitative classification methods for remotely sensed time series such as the approach developed by Evans and Geerken (2006), which not only identifies the dominant land cover type for a given picture element, but also gives an estimate of its relative fractional coverage. This allows, for example, to collect reference pixels that are not exclusively covered by a single land cover type (Evans and Geerken, 2006). The approach of Evans and Geerken (2006) is based on analytical derivatives of a Fourier series of NDVI time series observations. Splines share similar properties with a Fourier series regarding the form of the derivatives. A derivative of a spline is another spline in the same way that a derivative of a sinusoid is another sinusoid (Smith, 2003). Thus, spline models may be used to develop similar classification models. In trying to discriminate subtle differences between vegetation types in an arid environment, Evans and Geerken (2006) encountered problems such as spurious noise in higher harmonics that stem from the global fit of a Fourier series. As opposed to a Fourier series, splines can provide compact support, and it may well turn out that spline based semiquantitative classifiers are more robust than their Fourier based counterparts.

Other researchers proposed to try to bridge the scale gap between ground observations and coarse level satellite observations by scaling up the ground observations using a cascade of imagery products at different spatial resolutions (Liang, 2004; Fisher et al., 2006). Fisher et al. (2006) used logistic functions to scale up field measurements of phenological parameters to higher resolution satellite records (Landsat). These scaled up data could be used in turn to calibrate or validate coarse resolution imagery products. A framework for time series analysis should therefore be able to cope with not only vegetation index products obtained on a regular schedule by globally observing platforms, but also with shorter, possibly discontinuous time series of vegetation parameters produced from observations at a regional scale that provide higher spatial resolution but less temporal coverage. The flexible properties of spline models may allow in the future to compute phenological parameters from short discontinuous time series like Landsat observations in a way that is consistent with other spline models for hypertemporal observations at a coarser spatial scale.

An entirely different strategy to enhance spatial resolution is to use image fusion techniques in the spatial domain to augment the spatial resolution of coarse level imagery by merging it with imagery products of higher spatial resolution. Gao et al. (2006) developed a spatially and temporally adaptive reflectance fusion model (STARFM) to combine

spectral reflectances of Landsat and MODIS sensors. The model is based on a weighting function to fuse the MODIS and Landsat data by using information from spectrally similar neighbouring pixels. The STARFM algorithm may be valuable for applications that require high resolution in both time and space. Since splines are general, data driven models, they may also be applied to these data sets.

It was not possible within the frame of this work to examine all aspects of spline models relevant to remote sensing phenology in detail. However, the application programs described in appendix A, developed to fit spline models to time series of remotely sensed data and derive phenological parameters, are useful and perform equally well compared to other methods for the Aksu test data set. The programs are written in such a way as to provide convenient user interfaces to calculate phenological parameters from time series imagery. Thus, these software products already provide valuable tools in the analysis of Earth observation data that complement existing implementations. Whether splines are superior to other methods in the determination of phenological descriptors across different ecosystems and data archives may turn out in the future, when the operational software tools provided through this work will be used on more and more data sets with different characteristics. In addition, the implementation of the spline framework for time series analysis that forms the basis for the application programs in appendix A may just as well serve as a basis for other analysis methods of Earth observation time series such as semiquantitative classifiers.

A Description of Computer Programs

Based on the C++ implementations of the various spline models discussed in chapter 3, a number of application programs have been written in the course of this work to fit and evaluate spline models. There are four different programs: The **splfit** application is intended for the exploration of remotely sensed time series imagery using spline models. If an acceptable model is found, the **splfit** program can be used to perform the actual fitting of the model for an entire image data set. A spline model fitted by **splfit** can be evaluated using the **splcal** and **phencal** programs. The **splcal** program can be used to evaluate the spline model or its derivatives, e.g. to obtain a resampled data set based on the spline model's predictions, while the **phencal** program calculates various phenological parameters from the B-spline representation of a time series.

A.1 splview

The program **splview** provides a graphical user interface to explore different spline models for the analysis of remotely sensed time series. The interface consists of a grayscale image viewer combined with a time series plot. Time series vectors from different locations of an image containing a time series may be analysed using various spline models. Roots, extreme points and seasonal information (beginning/end of a season) may be displayed. Most of the options are available via context menus (right mouse button). The following display operations are available: 'SHIFT' key and left mouse button (region selection); 'STRG' key and left mouse button (pan); mouse wheel (zoom). The middle mouse button resets the zoom window to display the full extent of the loaded image.

Command line synopsis

splview [-a *name*] [-A *name*] [-b *band*] [-c *length*] [-k *number*|-K *name*] [-l *offset/gain*] [-m] [-p] [-t *threshold*] *image-file*

Command line options

-a name -A name The **-a** option may be used to specify the name of a file containing the abscissae values for the observations in the input image. The number of entries must match the number of bands in the image. Default abscissae values are defined as (0,1,...,n-1) where n is the number of bands, so this option must be used in case of irregularly spaced observations. The option **-A** specifies the name of a file that stores the abscissae values where the spline is to be evaluated. The files must be text files with numbers separated by white space.

-b band The index (1-based) of the band to open in the image viewer (defaults to the first band)

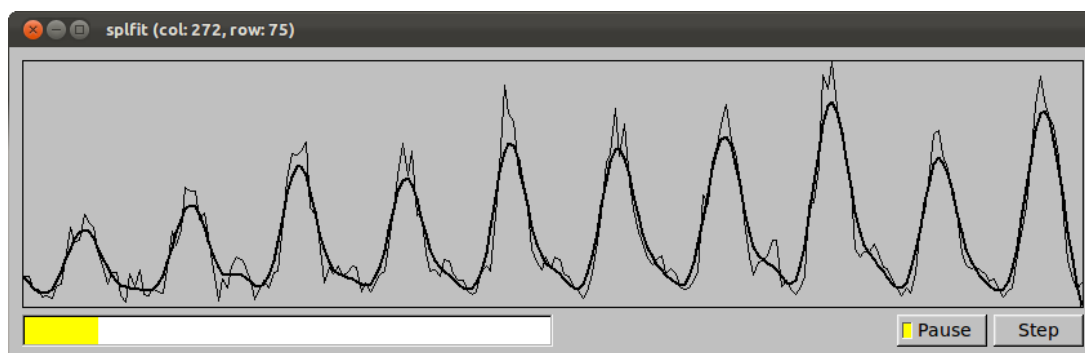


Figure A.1: Graphical user interface of the `splfit` program. Processing may be stopped at any time to examine the model fit to the current time series record in detail.

-c length The cycle length, i.e. the length of a compositing period in days. For example, the cycle length for standard MODIS vegetation index products is 16. Must be given if display of seasonal information is desired.

-k number -K name Options to control the number of spans/knots of the spline. Per default, the number of knots is equal to the number of data points and the knot positions are distributed uniformly across the data range. Option **-k** is used to specify the number of spans. The number of knots is then $number + 1$, and the knots are positioned uniformly along the horizontal axis. Option **-K** gives the name of a text file containing the positions of the knots separated by white space. It can be used to define nonuniform knot vectors. Only one of **-k** or **-K** should be given, capital **-K** overrides lowercase **-k** if both options are present. The configuration of the knot vector cannot be changed using the graphical interface if **-K** is used.

-l offset/gain A set of linear coefficients to transform the input data values

-m Flag indicating that compositing is month bounded (e.g. SPOT VEGETATION) rather than bounded by year (e.g. MODIS)

-p Indicates that a periodic spline model is to be used.

-t threshold A threshold value applied when computing seasonal information. A seasonal pattern is detected only if the difference between minimum and maximum value in a given year is in excess of the given threshold. A seasonal pattern is always matched by default (i.e. the default threshold is zero).

A.2 `splfit`

The program `splfit` is used to fit spline models of a given form to remotely sensed time series in an image. The output of `splfit` is another image containing the coefficients of the spline in B-form. The output is in the format of a standard ENVI¹ BIL file. Special

¹ENVI is a trademark of EXELIS Visual Information Solutions (<http://www.exelisvis.com>)

header tags are used to indicate the spline model used in the fit. By default **splfit** displays a plot window with a context menu (right mouse button) that allows to analyse the observational data and the resulting spline curves while the image is processed. The type of spline model that is fitted depends on the presence or absence of specific command line options related to the spline (see below).

Command line synopsis

splfit [-a *name*] -d *degree* [-f *factor*] [-i *number*] [-k *number*|-K *name*] [-l *offset/gain*] [-n *value*] [-o *value*] [-p] [-r *xmin/xmax/ymin/ymax*] [-s *value*] [-u *number*] [-w *time*] [-q|-Q]
input-image output-image

Command line options

The only mandatory spline related command line parameter is the degree (-d). If a degree is given as the only parameter, a B-spline is fitted in the least squares sense. If, in addition, -s is specified, a smooth B-spline model is fitted by imposing a penalty term on the integral of the second derivative of the fitted spline model. The value of -s is a proportionality constant that adjusts the smoothness of the spline, greater values of -s lead to smoother curves. If both -s and -o are specified, a P-spline is used as the model function, where the value of the -s argument is a proportionality constant as explained before and the value of -o is the order of the penalty.

-a name The name of a text file containing the abscissae values for the observations in the input image. Values must be separated by white space. The number of entries must match the number of bands in the image. Default abscissae values are defined as (0,1,...,n-1) where n is the number of bands, so this option must be used in case of irregularly spaced observations.

-l offset/gain A set of linear coefficients to transform the input data values

-n value Input nodata value. Defaults to zero.

-d degree The polynomial degree of the spline curve.

-k number -K name Options to control the number of spans/knots of the spline. Per default, the number of knots is equal to the number of data points and the knot positions are distributed uniformly across the data range. Option -k is used to specify the number of spans. The number of knots is then *number* + 1, and the knots are positioned uniformly along the horizontal axis. Option -K gives the name of a text file containing the positions of the knots separated by white space. It can be used to define nonuniform knot vectors. Only one of -k or -K should be given, capital -K overrides lowercase -k if both options are present.

- s** *value* Smoothness of the spline: *value* is a proportionality constant that defines the amount of smoothing applied when fitting a spline model. The larger the real number given by *value*, the smoother the fitted curve (i.e. the larger the penalty imposed on the integral of the curve's second derivative).
- o** *value* Order of the penalty for a P-spline model (usually 2.0).
- p** Indicates that a periodic spline model is to be used.
- u** *number* The number of iterations in fitting the upper envelope of the observational data.
- f** *factor* The multiplication factor when fitting the upper envelope (see **-u**). Between two successive iterations, the weights of data points above the curve are multiplied by *factor*.
- i** *number* Plot interval: when an image is processed, a new plot is displayed every *number* samples. Defaults to 1 (every sample in the image is plotted).
- r** *xmin/xmax/ymin/ymax* A tuple of four real values indicating the ranges of the horizontal and vertical axis of the plot.
- w** *time* Wait time: Time delay in (seconds*10) between the display of two successive plots. Defaults to zero.
- q** **-Q** Quiet: **-q** will suppress plotting of time series and show only a progress bar, **-Q** (scripting mode) will suppress all graphical output.
- y** Yes, plot the spline's control polygon and knots when displaying the spline curve. When using many knots, this may result in a confusingly packed display.

A.3 splcal

The **splcal** application program takes as input an image produced by its sibling, the **splfit** program. It outputs an image file containing time series as modeled by the spline defined by the B-coefficients from the output of **splfit**.

Command line synopsis

splfit [-**A** *name*] [-**D** *derivative*] [-**e**] [-**L** *offset/gain*] [-**N** *value*] [-**p** *number*] [-**t** *name*] *input-file* *output-file*

Command line options

- A name** The name of a text file containing a set of abscissae values as real numbers separated by blank space (blanks, tab or newline characters). If the argument is present and the file exists, the spline model specified in the input file is evaluated at these abscissae values. If the argument is omitted or the file cannot be opened, the spline model is evaluated at the default locations indicated in the header of the input file.
- D derivative** The derivative to evaluate. Defaults to 0, i.e. the spline curve itself.
- e** When outputting periodic data, include the points at both ends of the signal. Default is to generate a strictly periodic signal by omitting the last point (which is equal to the first).
- L offset/gain** A set of linear coefficients to transform the modeled data before writing the results to the output file. Default is to directly write the modeled data without transforming them.
- N value** Output nodata value. Defaults to zero.
- p number** The number of periods to output for a periodic spline. Defaults to one. This option is ignored if the spline model defined in the input is not a periodic one.
- t name** Output data type name. One of 'Byte', 'Integer' or 'Real' to produce output data as unsigned 8-bit, signed 16-bit or single precision (32-bit) float. Default is single precision float. If any other type is used, it is likely that the output data must be transformed to match the range of the selected output type (see **-L**).

A.4 phencal

The **phencal** software computes phenological parameters from a spline model obtained using **splfit**. The output is the format of a standard ENVI BIL file with a set of metafiles to access the individual phenological parameters. The various phenological parameters are explained under *output* below.

Command line synopsis

[-a] [-A] -b mm/dd/yyyy -c length [-g fraction] [-m] [-t threshold] input-file output-file

Command line options

- a -A Augment:** flag indicating that incomplete seasons at the beginning or end of the time series record should be completed (augmented) using the first or last recorded value. The default is to ignore incomplete seasons and clip the time series. The

option **-a** causes season augmenting at the beginning of the time series record, a capital **-A** indicates that augmenting should be done at the end of the time series. If both options are present, incomplete seasons are augmented at both the beginning and the end of the record.

- b** *mm/dd/yyyy* The date of the first observation (begin) of the time series.
- c** *length* The cycle length, i.e. the length of a compositing period in days. For example, the cycle length for standard MODIS vegetation index products is 16.
- g** *fraction* Greenness parameter: determines how onset and cease of vegetation growth in a season are determined. At present, *fraction* is a fraction of the seasonal amplitude. The onset (end) of greenness is defined as the point where the measurements in the time series reach a value equivalent to the early minimum (late minimum) plus *fraction* times the amplitude (i.e. the difference between the seasonal peak value and the latent value). If **-g** is omitted, only a reduced set of phenological parameters is computed (see below).
- m** Flag indicating that compositing is month bounded (e.g. SPOT VEGETATION) rather than bounded by year (e.g. MODIS)
- t** *threshold* A threshold value applied when computing seasonal information. A seasonal pattern is detected only if the difference between minimum and maximum value in a given year is in excess of the given threshold. A seasonal pattern is always matched by default (i.e. the default threshold is zero).

Output

This section describes the various phenological parameters output by **phencal**. In addition to the default parameters, an extended parameter set is available by specifying a value for the greenness (**-g**) option. A seasonal cycle is defined by looking at the sequence of minima and maxima present in a modelled time series. A cycle is bounded by the lowest minima to the left (early minimum) and right (late minimum) of an annual peak. In general, the late minimum of a given year coincides with the early minimum of the following year.

DOY_Early_Min The day of year of the early minimum.

DOY_Peak The day of year of the annual peak value.

DOY_Late_Min The day of year of the late minimum.

Early_Min_Val The modelled value at the time of the early minimum.

Peak_Val The modelled annual peak value.

Late_Min_Val The modelled value at the time of the late minimum.

Min_Min_Duration The number of days between the early and late minima of a given year.

Amplitude The difference between the modelled peak value and the latent value (see *Latent_Val* below).

Latent_Val The average of the modelled values at the early and late minima.

Min_Min_Integral The time integrated modelled value from the early to the late minimum of a given year.

Latent_Integral The time integrated latent value from the early to the late minimum of a given year (a "box" of width *Duration* and height *Latent_Val*).

Total_Integral The sum of *Min_Min_Integral* and *Latent_Integral*.

The following parameters are controlled by the **-g** option (see above):

DOY_Start_Green The day of year of the onset of vegetation growth (greenness).

DOY_End_Green The day of year when vegetation ceases to develop, i.e is completely senescent.

Green_Duration The number of days between the start and end of greenness, a proxy for the duration of active vegetation development.

Green_Integral The time integrated modelled value over the duration of greenness (see above).

Greenup_Rate The slope of a line connecting the point of onset of greenness and the annual peak value. Describes the approximate rate of vegetation growth.

Senescence_Rate The (positive) slope of a line connecting the annual peak value and the point of end of greenness. Describes the approximate rate of advancement of vegetation senescence.

References

- Aboufadel, E. and S. Schlicker (1999). *Discovering Wavelets*. New York: John Wiley & Sons.
- Acton, F. S. (1990). *Numerical Methods That Work*. Washington D.C.: The Mathematical Association of America.
- Andres, L., W. A. Salas, and D. Skole (1994). Fourier analysis of multi-temporal AVHRR data applied to a land cover classification. *International Journal of Remote Sensing* 15(5), 1115–1121.
- Archibald, S. and R. J. Scholes (2007). Leaf green-up in a semi-arid African savanna – separating tree and grass response to environmental cues. *Journal of Vegetation Science* 18, 583–594.
- Azzali, S. and M. Menenti (2000). Mapping vegetation-soil-climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. *International Journal of Remote Sensing* 21(5), 973–996.
- Badhwar, G. B. (1984). Automatic corn–soybean classification using Landsat MSS data. II. Early season crop proportion estimation. *Remote Sensing of Environment* 14, 31–37.
- Balridge, A. M., S. J. Hook, C. I. Grove, and G. Rivera (2009). The ASTER spectral library version 2.0. *Remote Sensing of Environment* 113, 711–715.
- Bloomfield, P. (2000). *Fourier Analysis of Time Series* (2nd ed.). Wiley Series in Probability and Statistics. New York: John Wiley & Sons.
- Bowman, A. W. and A. Azzalini (1997). *Applied Smoothing Techniques for Data Analysis*, Volume 18 of *Oxford Statistical Science Series*. Oxford: Clarendon Press.
- Bradley, B. A., R. W. Jacob, J. F. Hermance, and J. F. Mustard (2007). A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment* 106, 137 – 145.
- Chatfield, C. (2004). *The Analysis of Time Series: An Introduction* (6th ed.). Texts in Statistical Science. Boca Raton: Chapman & Hall/CRC.
- Chen, J., P. Jönsson, T. M., Z. Gu, B. Matsushita, and L. Eklundh (2004). A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment* 91, 332–344.
- Cleveland, W. S. (1994). *The Elements of Graphing Data* (Revised ed.). Summit, N. J.: Hobart Press.
- Cleveland, R. B., W. S. Cleveland, J. E. McRae, and I. Terpenning (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics* 6(1), 3–73.
- Congalton, R. G. (1988). Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. *Photogrammetric Engineering & Remote Sensing* 54(5), 587–592.
- Cooley, J. W. and J. W. Tukey (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation* 19, 297–301.
- Cracknell, A. P. (1998). Synergy in remote sensing-what's in a pixel? *In* 19(11), 2025–2047.
- Craig, R. G. (1979). Autocorrelation in Landsat data. *In Proceedings of the International Symposium on Remote Sensing of Environment*, Volume 13/3, Ann Arbor.

- Crist, E. P. and R. J. Kauth (1986). The tasseled cap de-mystified. *Photogrammetric Engineering & Remote Sensing* 52(1), 81–86.
- Curran, P. (1980). Multispectral remote sensing of vegetation amount. *Progress in Physical Geography* 4, 315–341.
- de Boor, C. (1972). On calculating with B-splines. *Journal of Approximation Theory* 6, 50–62.
- de Boor, C. (1977). Package for calculating with B-splines. *SIAM Journal on Numerical Analysis* 14(3), pp. 441–472.
- de Boor, C. (2001). *A Practical Guide to Splines* (Revised ed.), Volume 27 of *Applied Mathematical Sciences*. New York: Springer.
- de Jong, R., S. de Bruin, A. de Wit, M. E. Schaepman, and D. L. Dent (2011). Analysis of monotonic greening and browning trends from global NDVI time-series. *Remote Sensing of Environment* 115, 692–702.
- DeFries, R., M. Hansen, and J. Townshend (1995). Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. *Remote Sensing of Environment* 54, 209–222.
- Dierckx, P. (1993). *Curve and Surface Fitting with Splines*. Monographs on Numerical Analysis. Oxford: Clarendon Press.
- Dobbertin, M. and G. S. Biging (1996). A simulation study of the effect of scene autocorrelation, training sample size and sampling method on classification accuracy. *Canadian Journal of Remote Sensing* 22(4), 360–367.
- Eastman, R. J. and M. Fulk (1993). Long sequence time series evaluation using standardized principal components. *Photogrammetric Engineering & Remote Sensing* 59(8), 1307–1312.
- Edwards, S. and R. A. Gordon (2004). Extreme curvature of polynomials. *The American Mathematical Monthly* 111(10), 890–899.
- Eilers, P. H. C. and B. D. Marx (1996). Flexible smoothing with B-splines and penalties. *Statistical Science* 11(2), 89–121.
- Eilers, P. H. C. and B. D. Marx (2010). Splines, knots and penalties. *Wiley Interdisciplinary Reviews: Computational Statistics* 2(6), 637–653.
- Epperson, J. F. (2007). *An Introduction to Numerical Methods and Analysis* (Revised ed.). Hoboken: John Wiley & Sons.
- Evans, J. P. and R. Geerken (2006). Classifying rangeland vegetation type and coverage using a Fourier component based similarity measure. *Remote Sensing of Environment* 105, 1–8.
- Faraway, J. J. (2006). *Extending the Linear Model with R*. Texts in Statistical Science. Boca Raton: Chapman & Hall/CRC.
- Farin, G. (2002). *Curves and Surfaces for GAGD* (5th ed.). San Francisco: Morgan Kaufmann.
- Fensholt, R., I. Sandholt, and M. Schultz Rasmussen (2004). Evaluation of MODIS LAI, fAPAR and the relation between fAPAR and NDVI in a semi-arid environment using in situ measurements. *Remote Sensing of Environment* 91, 490–507.
- Fischer, A. (1994a). A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters. *Remote Sensing of Envi-*

- ronment* 48, 220–230.
- Fischer, A. (1994b). A simple model for the temporal variations of NDVI at regional scale over agricultural countries. validation with ground radiometric measurements. *International Journal of Remote Sensing* 15(7), 1421–1446.
- Fisher, J. I., J. F. Mustard, and M. A. Vadeboncoeur (2006). Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sensing of Environment* 100, 265–279.
- Fisher, R. A. (1929). Tests of significance in harmonic analysis. *Proceedings of the Royal Society, Series A* 125, 54–59.
- Fligge, M., S. K. Solanki, and J. Beer (1999). Determination of solar cycle length variations using the continuous wavelet transform. *Astronomy and Astrophysics* 346, 313–321.
- Foley, J. D., A. van Dam, S. K. Feiner, and J. F. Hughes (1997). *Computer Graphics in C. Principles and Practice* (2nd ed.). Boston: Addison–Wesley.
- Fugal, D. L. (2009). *Conceptual Wavelets in Digital Signal Processing*. San Diego: S & ST Space and Signals Technical Publishing.
- Galford, G. L., J. F. Mustard, J. Melillo, A. Gendrin, C. C. Cerri, and C. E. P. Cerri (2008). Wavelet analysis of MODIS time series to detect expansion and intensification of row-crop agriculture in Brazil. *Remote Sensing of Environment* 112, 576–587.
- Gao, F., J. Masek, M. Schwaller, and F. Hall (2006). On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing* 44(8), 2207–2218.
- Geerken, R., B. Zaitchick, and J. P. Evans (2005). Classifying rangeland vegetation type and coverage from NDVI time series using Fourier filtered cycle similarity. *International Journal of Remote Sensing* 26(24), 5535–5554.
- Gershenfeld, N. (1999). *The Nature of Mathematical Modeling*. Cambridge: Cambridge University Press.
- Golub, G. H. and C. F. van Loan (1996). *Matrix Computations* (3rd ed.). Baltimore: Johns Hopkins University Press.
- Hamming, R. W. (1973). *Numerical Methods for Scientists and Engineers* (2nd ed.). New York: Dover Publications.
- Hamming, R. W. (1989). *Digital Filters* (3rd ed.). Mineola, N.Y.: Dover Publications.
- Hammond, T. O. and D. L. Verbyla (1996). Optimistic bias in classification accuracy assessment. *International Journal of Remote Sensing* 17(6), 1261–1266.
- Hao, X., Y. Chen, C. Xu, and W. Li (2008). Impacts of climate change and human activities on the surface runoff in the Tarim river basin over the last fifty years. *Water Resources Management* 22, 1159–1171.
- Hastie, T., R. Tibshirani, and J. Friedman (2009). *The Elements of Statistical Learning* (2nd ed.). Springer Series in Statistics. Springer.
- Hermance, J. F. (2007). Stabilizing high-order, non-classical harmonic analysis of NDVI data for average annual models by damping model roughness. *International Journal of Remote Sensing* 28(12), 2801–2819.
- Hermance, J. F., R. W. Jacob, B. A. Bradley, and J. F. Mustard (2007). Extracting phenological signals from multiyear AVHRR NDVI time series: Framework for applying

- high-order annual splines with roughness damping. *IEEE Transactions on Geoscience and Remote Sensing* 45(10), 3264–3276.
- Hill, J., M. Stellmes, T. Udelhoven, A. Röder, and S. Sommer (2008). Mediterranean desertification and land degradation mapping related land use change syndromes based on satellite observations. *Global and Planetary Change* 64, 146–157.
- Hird, J. N. and G. J. McDermid (2009). Noise reduction of NDVI time series: An empirical comparison of selected techniques. *Remote Sensing of Environment* 113, 248–258.
- Holben, B. N. (1986). Characteristics of maximum value composite images for temporal AVHRR NDVI data. *International Journal of Remote Sensing* 7, 1435–1445.
- Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83, 195–213.
- Huete, A. and R. D. Jackson (1987). Suitability of spectral indices for evaluating vegetation characteristics on arid rangelands. *Remote Sensing of Environment* 23, 213–232.
- Huete, A. R. (1988). A soil-adjusted vegetation index (savi). *Remote Sensing of Environment* 25, 295–309.
- Ivits, E., M. Cherlet, W. Mehl, and S. Sommer (2008). Estimating the ecological status and change of riparian zones in Andalusia assessed by multi-temporal AVHRR datasets. *Ecological Indicators* 9(3), 422–431.
- Jackson, R. D. and A. R. Huete (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine* 11, 185–200.
- Jakubauskas, M. E., D. R. Legates, and J. H. Kastens (2001). Harmonic analysis of time-series AVHRR NDVI data. *Photogrammetric Engineering & Remote Sensing* 67(4), 461–470.
- Jakubauskas, M. E., D. L. Peterson, J. H. Kastens, and D. R. Legates (2002). Time series remote sensing of landscape-vegetation interactions in the southern Great Plains. *Photogrammetric Engineering & Remote Sensing* 68(10), 1021–1030.
- Jasinski, M. F. (1990). Sensitivity of the normalized difference vegetation index to sub-pixel canopy cover, soil albedo and pixel scale. *Remote Sensing of Environment* 32, 169–187.
- Jenkins, M. A. (1975). Zeros of a real polynomial. *ACM Transactions on Mathematical Software* 1(2), 178–189.
- Jenkins, M. A. and J. F. Traub (1970). A three-stage algorithm for real polynomials using quadratic iteration. *SIAM Journal on Numerical Analysis* 7(4), 545–566.
- Jensen, J. R. (1983). Biophysical remote sensing. *Annals of the Association of American Geographers* 73(1), 111–132.
- Jiang, L., T. Jufen, Z. Zhijie, L. Tianhong, and L. Jianhua (2005). Water resources, land exploration and population dynamics in arid areas – the case of the Tarim river basin in Xinjiang of China. *Population and Environment* 26(6), 471–503.
- Jones, H. G. and R. A. Vaughan (2010). *Remote Sensing of Vegetation*. Oxford: Oxford University Press.
- Jönsson, P. and L. Eklundh (2002). Seasonality extraction by function fitting of time series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing* 40(8), 1824–1832.

- Jönsson, P. and L. Eklundh (2004). Timesat – a program for analyzing time-series of satellite sensor data. *Computers & Geosciences* 30, 833–845.
- Justice, C. O., E. Vermote, J. R. G. Townshend, R. Defries, D. P. Roy, D. K. Hall, V. V. Salomonson, J. L. Privette, G. Riggs, A. Strahler, W. Lucht, R. B. Myneni, Y. Knyazikhin, S. W. Running, R. R. Nemani, W. Zhengming, A. R. Huete, W. van Leeuwen, R. E. Wolfe, L. Giglio, J. P. Muller, P. Lewis, and M. J. Barnsley (1998). The moderate resolution imaging spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing* 36(4), 1228–1249.
- Keele, L. (2008). *Semiparametric Regression for the Social Sciences*. Chichester: Wiley.
- Labovitz, M. L. and E. J. Masuoka (1984). The influence of autocorrelation in signature extraction—an example from a geobotanical investigation of Cotter Basin, Montana. *International Journal of Remote Sensing* 5(2), 315–332.
- Lambin, E. F. (1996). Change detection at multiple temporal scales: Seasonal and annual variations in landscape variables. *Photogrammetric Engineering & Remote Sensing* 62(8), 931–938.
- Lengyel, E. (2004). *Mathematics for 3D Game Programming & Computer Graphics* (2nd ed.). Boston: Charles River Media.
- Li, Z. and M. Kafatos (2000). Interannual variability of vegetation in the United States and its relation to El Nino/southern oscillation. *Remote Sensing of Environment* 71, 239–247.
- Liang, S. (2004). *Quantitative Remote Sensing of Land Surfaces*. Hoboken: John Wiley & Sons.
- Liang, S., H. Fang, M. Chen, C. J. Shuey, C. Walthall, C. Daughtry, J. Morisette, C. Schaaf, and A. Strahler (2002). Validating MODIS land surface reflectance and albedo products: methods and preliminary results. *Remote Sensing of Environment* 83, 149–162.
- Lobo, A. and P. Maisongrande (2006). Searching for trends of change through exploratory data analysis of time series of remotely-sensed images of SW Europe. In J. A. Sobrino (Ed.), *Second Recent Advances in Quantitative Remote Sensing*, Proceedings of the Second International Symposium on Recent Advances in Quantitative Remote Sensing, pp. 534–537.
- Lu, H., M. R. Raupach, T. R. McVicar, and D. J. Barrett (2003). Decomposition of vegetation cover into woody and herbaceous components using AVHRR NDVI time series. *Remote Sensing of Environment* 86, 1–18.
- MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge: Cambridge University Press.
- Menenti, M., S. Azzali, W. Verhoef, and R. Swol (1993). Mapping agroecological zones and time lag in vegetation growth by means of Fourier analysis of time series of NDVI images. *Advances in Space Research* 13(5), 233–237.
- Moody, A. and D. M. Johnson (2001). Land-surface phenologies from AVHRR using the discrete Fourier transform. *Remote Sensing of Environment* 75, 305–323.
- Mosier, R. G. (2009). Spline functions. In R. L. Norton (Ed.), *Cam Design and Manufacturing Handbook* (2nd ed.), Chapter 5, pp. 69–124. New York: Industrial Press.
- Moulin, S., L. Kerogat, N. Viovy, and G. Dedieu (1997). Global-scale assessment of vegetation phenology using noaa/avhrr satellite measurements. *Journal of Climate* 10, 1154–1170.

- Myneni, R. B., C. D. Keeling, C. J. Tucker, G. Asrar, and R. R. Nemani (1997). Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* 386, 698–702.
- Neuenschwander, A. L. and K. A. Crews (2008). Disturbance, management, and landscape dynamics: Harmonic regression of vegetation indices in the lower Okavango delta, Botswana. *Photogrammetric Engineering & Remote Sensing* 74(6), 753–764.
- Nocedal, J. and S. J. Wright (2006). *Numerical Optimization* (2nd ed.). Springer Series in Operations Research. New York: Springer.
- Olsson, L. and L. Eklundh (1994). Fourier series for analysis of temporal sequences of satellite sensor imagery. *International Journal of Remote Sensing* 15(18), 3735–3741.
- O’Sullivan, D. and D. J. Unwin (2010). *Geographic Information Analysis* (2nd ed.). Hoboken: John Wiley & Sons.
- Percival, D. B. and A. T. Walden (2000). *Wavelet Methods for Time Series Analysis*. New York: Cambridge University Press.
- Peters, A. E., M. D. Eve, E. H. Holt, and W. E. Whitford (1997). Analysis of desert plant community growth patterns with high temporal resolution satellite spectra. *Journal of Applied Ecology* 34, 418–432.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in C*. New York: Cambridge University Press.
- Price, J. C. (1993). Estimating leaf area index from satellite data. *IEEE Transactions on Geoscience and Remote Sensing* 31(3), 727–734.
- Reed, B. C., J. F. Brown, D. VanderZee, T. R. Loveland, J. W. Merchant, and D. O. Ohlen (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science* 5, 703–714.
- Reed, B. C., M. White, and J. F. Brown (2003). Remote sensing phenology. In M. D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science*, pp. 365–381. Dordrecht: Kluwer.
- Richardson, A. J. and C. L. Wiegand (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering & Remote Sensing* 43(12), 1541–1552.
- Roderick, M. L., I. R. Noble, and S. W. Cridland (1999). Estimating woody and herbaceous vegetation cover from time series satellite observations. *Global Ecology and Biogeography* 8, 501–508.
- Roerink, G. J., M. Menenti, and W. Verhoef (2000). Reconstructing cloudfree NDVI composites using fourier analysis of time series. *International Journal of Remote Sensing* 21(9), 1911–1917.
- Rogers, D. F. (2001). *An Introduction to NURBS*. San Francisco: Morgan Kaufmann.
- Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics* 11(4), 735–757.
- Ruppert, D., M. P. Wand, and R. J. Carroll (2003). *Semiparametric Regression*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge University Press.
- Sakamoto, T., M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, and H. Ohno (2005). A crop phenology detection method using time-series modis data. *Remote Sensing of Environment* 96, 366–374.

- Sanderson, C. (2010). Armadillo: An open source C++ linear algebra library for fast prototyping and computationally intensive experiments. Technical report, NICTA.
- Savitzky, A. and M. J. E. Golay (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 36(8), 1627–1639.
- Sellers, P. J., C. J. Tucker, G. J. Collatz, S. O. Los, and C. O. Justice (1994). A global 1° by 1° NDVI data set for climate studies. part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *International Journal of Remote Sensing* 15(17), 3519–3545.
- Shumway, R. H. and D. S. Stoffer (2011). *Time Series Analysis and Its Applications* (3rd ed.). Springer Texts in Statistics. New York: Springer.
- Smith, S. W. (2003). *Digital Signal Processing*. Boston: Newnes.
- Spiegel, M. R. and L. J. Stephens (2008). *Statistics* (4th ed.). New York: McGraw-Hill.
- Stoffels, J., S. Mader, J. Hill, W. Werner, and G. Ontrup (2011). Satellite-based stand-wise forest cover type mapping using a spatially adaptive classification approach. *European Journal of Forest Research* 131(4), 1071–1089.
- Stone, C. J. (1986). Comment: Generalized additive models. *Statistical Science* 2, 312–314.
- Strang, G. (1994). Wavelets. *American Scientist* 82, 250–255.
- Suffern, K. (2007). *Ray Tracing From The Ground Up*. Wellesley: A K Peters.
- Takezawa, K. (2006). *Introduction to Nonparametric Regression*. Wiley Series in Probability and Statistics. Hoboken: John Wiley & Sons.
- Tan, B., J. T. Morisette, R. E. Wolfe, F. Gao, G. A. Ederer, J. Nightingale, and J. E. Pedelty (2011). An enhanced TIMESAT algorithm for estimating vegetation phenology metrics from MODIS data. *IEEE Journal on Selected Topics in Applied Earth Observations and Remote Sensing* 4(2), 361–371.
- Tucker, C. J., B. N. Holben, J. H. Elgin, Jr., and J. E. McMurterey III (1981). Remote sensing of total dry-matter accumulation in winter wheat. *Remote Sensing of Environment* 11, 171–189.
- Tucker, C. J., J. E. Pinzon, M. E. Brown, D. Slayback, E. W. Pak, R. Mahoney, E. Vermote, and N. El Saleous (2005). An extended AVHRR 8-km ndvi data set compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing* 26(20), 4485–5598.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, Mass.: Addison–Wesley.
- Tulleken, H. (2008). Poisson disk sampling. *Dev.Mag* 21, 21–25.
- Udelhoven, T. (2011). TimeStats: A software tool for the retrieval of temporal patterns from global satellite archives. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4(2), 310–317.
- Unser, M. (1999). Splines. A perfect fit for signal and image processing. *IEEE Signal Processing Magazine* 16(6), 22–38.
- Unser, M., A. Aldroubi, and M. Eden (1993a). B-spline signal processing: Part I—theory. *IEEE Transactions on Signal Processing* 41(2), 821–832.
- Unser, M., A. Aldroubi, and M. Eden (1993b). B-spline signal processing: Part II—efficient design and applications. *IEEE Transactions on Signal Processing* 41(2), 834–848.

- van Dijk, A. (1987). Smoothing vegetation index profiles: An alternative method for reducing radiometric disturbance in NOAA/AVHRR data. *Photogrammetric Engineering & Remote Sensing* 53(8), 1059–1067.
- Velleman, P. F. (1980). Definition and comparison of robust nonlinear data smoothing algorithms. *Journal of the American Statistical Association* 75(371), 609–615.
- Verbesselt, J., R. Hyndman, G. Newnham, and D. Culvenor (2010). Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment* 114, 106–115.
- Vinciková, H., M. Hais, J. Brom, J. Procházka, and E. Pecharová (2010). Use of remote sensing methods in studying agricultural landscapes – a review. *Journal of Landscape Studies* 3, 53–63.
- Viovy, N., O. Arino, and A. S. Belward (1992). The best index slope extraction (BISE): A method for reducing noise in NDVI time series. *International Journal of Remote Sensing* 13(8), 1585–1590.
- Wagenseil, H. and C. Samimi (2006). Assessing spatio-temporal variations in plant phenology using Fourier analysis on NDVI time series: results from a dry savannah environment in Namibia. *International Journal of Remote Sensing* 27(16), 3455–3471.
- Wahba, G. (1990). *Spline Models for Observational Data*, Volume 59 of *CMBS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia: Society for Industrial and Applied Mathematics.
- Wang, G., Y. Shen, J. Zhang, S. Wang, and W. Mao (2010). The effects of human activities on oasis climate and hydrologic environment in the Aksu river basin, Xinjiang, China. *Environmental Earth Sciences* 59, 1759–1769.
- Warner, R. M. (1998). *Spectral Analysis of Time-Series Data*. New York: Guilford Press.
- White, M. A., K. M. de Beurs, K. Didan, D. W. Inouye, A. D. Richardson, O. P. Jensen, J. O’Keefe, G. Zhang, R. R. Nemani, W. J. D. van Leeuwen, J. F. Brown, A. de Wit, M. Schaepman, X. Lin, M. Dettinger, A. S. Bayley, J. Kimball, M. D. Schwartz, D. D. Baldocchi, J. T. Lee, and W. K. Lauenroth (2009). Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. *Global Change Biology* 15(10), 2335–2359.
- White, M. A., P. E. Thornton, and S. W. Running (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles* 11(2), 217–234.
- Wold, S. (1974). Spline functions in data analysis. *Technometrics* 16(1), 1–11.
- Wood, S. N. (2006). *Generalized Additive Models*. Boca Raton: Chapman & Hall/CRC.
- Xu, H., M. Ye, and Y. Song (2005). The dynamic variation of water resources and its tendency in the Tarim river basin. *Journal of Geographical Sciences* 15(4), 467–474.
- Yandell, B. S. (1993). Smoothing splines – a tutorial. *The Statistician* 42, 317–319.
- Yang, W., D. Huang, B. Tan, J. C. Stroeve, N. V. Shabanov, Y. Knyazikhin, R. R. Nemani, and R. B. Myneni (2006). Analysis of leaf area index and fraction of PAR absorbed by vegetation products from the terra MODIS sensor: 2000-2005. *IEEE Transactions on Geoscience and Remote Sensing* 44(7), 1829–1842.
- Zhang, X., M. A. Friedl, C. B. Schaaf, A. H. Strahler, J. C. F. Hodges, F. Gao, and B. C. Reed (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Envi-*

ronment 84, 471–475.

Zhang, X., J. C. F. Hodges, C. B. Schaaf, M. A. Friedl, A. H. Strahler, and F. Gao (2001). Global vegetation phenology from AVHRR and MODIS data. In *Proceedings of the International Geoscience and Remote Sensing Symposium*, Volume 5, pp. 2262–2264.

Lebenslauf

Persönliche Daten

Sebastian Mader
Maximinstraße 11
54292 Trier

Geboren am 22.04.1976 in Dachau
Staatsangehörigkeit: deutsch

Schulbildung

- 09/1983–10/1986 Grundschule Vierkirchen (Bayern)
- 11/1986–06/1987 Grundschule Schillerschule, Bietigheim-Bissingen (Baden-Württemberg)
- 08/1987–06/1996 Gymnasien I und II im Ellental, Bietigheim-Bissingen
Abschluss mit dem Abitur

Studium

- 10/1997–09/2006 Diplomstudiengang Angewandte Physische Geographie an der Universität Trier
Diplomarbeit:
Die Nutzung hyperspektraler Aufnahmesysteme zur erweiterten Parametrisierung eines Modellansatzes zur Landschaftsbewertung in mediterranen Trockengebieten

Sonstige Tätigkeiten

- 05/2006–09/2006 Wissenschaftliche Hilfskraft mit Abschluss im Fachbereich Geographie/Geowissenschaften der Universität Trier, Fach Fernerkundung

Berufliche Tätigkeiten

- seit 10/2006 Wissenschaftlicher Mitarbeiter im Fachbereich Geographie/Geowissenschaften der Universität Trier, Fach Fernerkundung