

Diploma Thesis

Time and Space dependent Biophysical Land Surface Parameters from AVHRR NDVI Data for Use in Regional Climate Models

Reto Stöckli

(Department of Environmental Sciences, ETH Zürich)

Supervisor

Prof. Dr. Christoph Schär

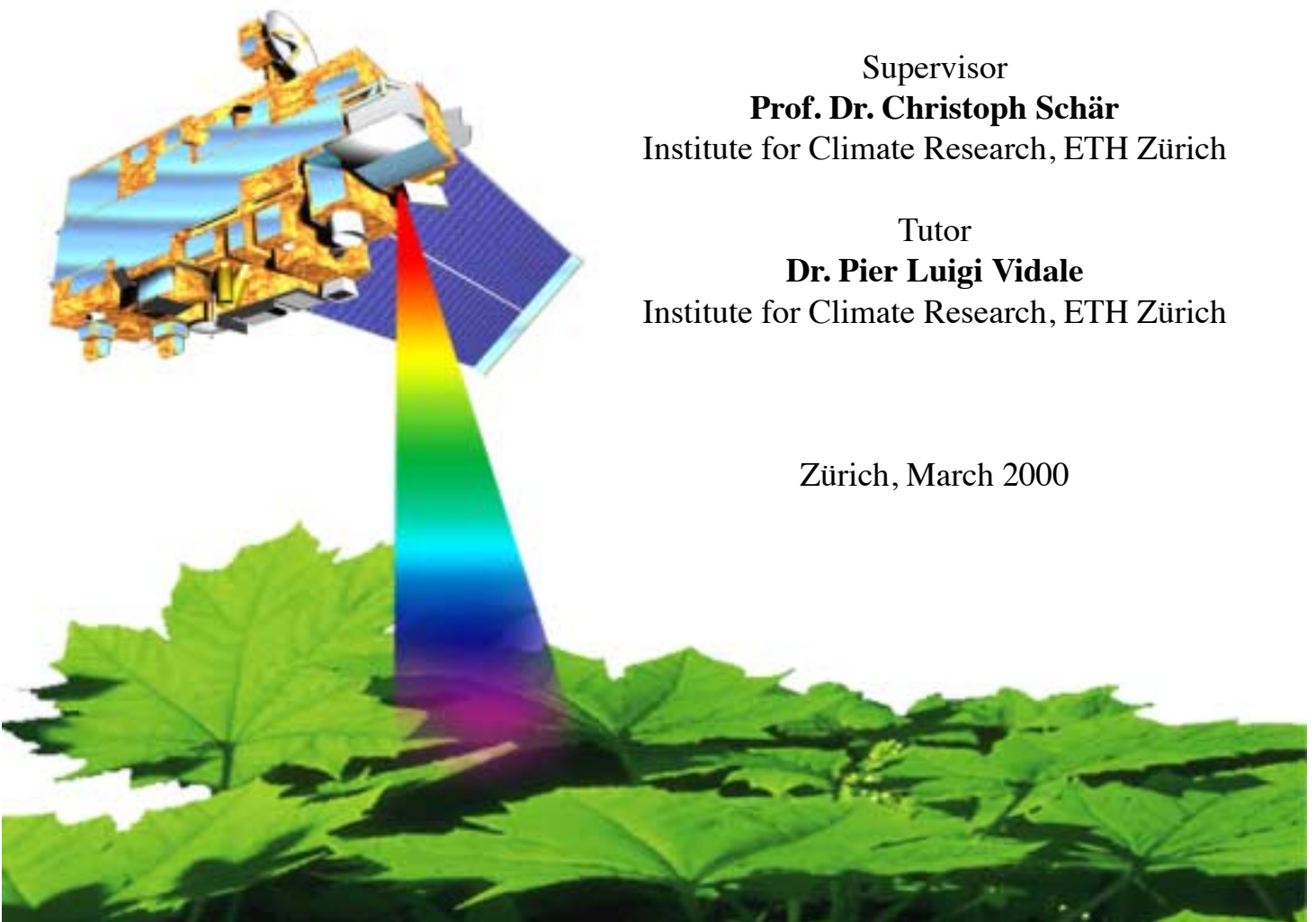
Institute for Climate Research, ETH Zürich

Tutor

Dr. Pier Luigi Vidale

Institute for Climate Research, ETH Zürich

Zürich, March 2000



Front Image: EOS TERRA spacecraft with the Moderate-resolution Imaging Spectroradiometer (MODIS) instrument operating. Spacecraft model by Reto Stöckli (1999).

Contents

Acknowledgements	3
Abstract	4
1. Introduction	5
2. Satellite Remote Sensing of Vegetation	10
2.1 Spectral Properties of Land Surface Vegetation	10
2.2 Vegetation Monitoring with the NOAA AVHRR Sensor	12
2.3 The Pathfinder NDVI Dataset	13
2.4 Analysis of Pathfinder NDVI Time-Series	14
2.5 Fourier Adjustment of NDVI Time-Series	17
2.5.1 Introduction	17
2.5.2 Fourier series	17
2.5.3 Fourier adjustment algorithm	18
2.5.4 First Fourier adjustment	18
2.5.5 Apply a second weighted Fourier adjustment	19
2.5.6 Weighting the Fourier fit	19
2.6 Effectiveness of the Fourier Adjustment Technique	20
3. The European Fourier-Adjusted and Interpolated NDVI	23
3.1 Data Retrieval, Reprojection and Scaling	24
3.2 Programming and Implementation	25
3.3 Substitution of Missing Data	26
3.4 An Enhanced Fourier Adjustment Procedure	26
3.4.1 Adjustments to the weighting scheme	27
3.4.2 Detection of positive outliers	30
3.4.3 A third weighted Fourier fit	31
3.4.4 Repeated main period	32
3.5 Interpolation of Missing Data	33
3.6 Examples of Adjusted NDVI Time-Series	34
3.7 Remaining Anomalies	35
3.7.1 Positive outliers at the end of the vegetation period	35
3.7.2 Overestimation in boreal forest habitats	36
3.7.3 Agriculture	37
3.7.4 Growth-season start and end	37

4. Results: The 1983-1993 EFAI-NDVI	38
4.1 Latitudinal Variation	38
4.2 Seasonal Variation	39
4.3 Inter-Annual Variation	42
4.4 Spatial Variation	45
4.5 Comparison to the FASIR NDVI	47
4.6 Error Estimation	48
5. Land Surface Parameterization.....	50
5.1 Introduction	50
5.2 Land Cover Classification	52
5.3 Soil Type Classification	54
5.4 Estimation of Net Primary Production through NDVI	56
5.5 Calculation of the Surface Parameters from NDVI	56
5.5.1 SR - FPAR	56
5.5.2 NDVI - FPAR	57
5.5.3 SR-NDVI intermediate model	57
5.5.4 Leaf area index and canopy greenness	58
5.5.5 Roughness length	59
5.6 The MAPPER Program	60
6. Summary and Discussion	61
References	63
Abbreviations	67
Appendix	69

Acknowledgements

First of all, I would like to thank Prof. Christoph Schär, who made this diploma thesis possible. I would like to express my gratitude to Dr. Pier Luigi Vidale, who served as my advisor for this project. He not only provided professional support, guidance and cooperation, but also allowed independence and had trust in my research.

I also would like to thank Jim Collatz from the Biospheric Sciences Branch (Code 923), NASA/Goddard Space Flight Center, for providing the directions to a pre-release of the ISLSCP-II datasets. Fritz Hasler from the Goddard Laboratory for Atmospheres (Code 912) at NASA/GSFC receives my thanks for the possibility to use the Fortran-90 licence and the computing power on “Mitch” - it greatly improved the efficiency of this project. My thanks also go to David Fanning from Fanning Software Consulting for the very helpful online IDL support and for his IDL Programming Techniques book (Fanning 1999).

I always appreciated the numerous discussions, coffee-breaks and the friendly environment in the Institute for Climate Research. I would finally like to give my warmest thanks to my parents and to my sister for their love and support. A very special thank goes to Eva for inspiring my life so much (your precious advice “non docere impransus” was of great use throughout this project!).

The Land Pathfinder NDVI data used in this study were produced through funding from the Earth Observing System Pathfinder Program of NASA’s Mission to Planet Earth in cooperation with National Oceanic and Atmospheric Administration. The data were provided by the Earth Observing System Data and Information System (EOSDIS), Distributed Active Archive Center at Goddard Space Flight Center which archives, manages, and distributes this data set.

Abstract

Remotely-sensed normalized difference vegetation index (NDVI) from the NOAA Land Pathfinder dataset is used to estimate biophysical land surface parameters for the years 1983 to 1993. 10-day NDVI composites with a spatial resolution of 0.1° by 0.1° are analyzed for the European domain and error sources in the dataset are discussed. A Fourier adjustment procedure developed by Sellers et al. (1994, 1996b) and Los (1998) allows to separate reliable NDVI values from interferences, so that the vegetation signal can be extracted from an error-contaminated time-series. This procedure is reviewed and tested for its effectiveness on idealized time-series. The Fourier adjustment technique is then applied to the Land Pathfinder NDVI dataset and special cases are presented, where the original algorithm does not hold. An enhanced Fourier adjustment procedure is developed by taking these special cases in consideration. An 11-year European Fourier-Adjusted and Interpolated NDVI dataset (EFAI-NDVI) is processed and the following associated biophysical land surface parameters are estimated with simple relationships: Fraction of absorbed Photosynthetically Active Radiation (FPAR), Leaf Area Index (LAI), vegetation cover fraction, greenness fraction and roughness length.

Results show, that the enhanced Fourier adjustment algorithm performs well for most areas within the European domain. The correction scheme is most effective where suspect data are supposed to occur due to atmospheric interferences. Certain inconsistencies need further evaluation: Vegetation types with a large seasonality occurring mainly at high latitudes lead to overestimated NDVI values when sudden changes in vegetation activity occur. Positive outliers during time of low vegetation activity could only be partly removed, what needs further examination. The spatial variability of the EFAI-NDVI is detected even at small scales in the examined dataset, which encourages the use of high resolution vegetation datasets in Soil-Vegetation-Atmosphere Transfer Schemes (SVATS). The correct spatial and temporal accuracy of NDVI is important for the estimation of the associated land surface properties. Future research will involve the use of these properties in sensitivity studies with coupled biosphere-atmosphere models. A better representation and understanding of land surface processes in regional climate modelling is expected as a future prospect.

1. Introduction

During the past decades, the understanding of human-induced and natural variations in the global climate has become very important. The climate system consists of several components interacting with each other: the atmosphere, the hydrosphere, the cryosphere, the lithosphere and the biosphere. Especially the biosphere is important in its ability to sustain life, in regulating the land surface water and carbon balance and in determining the radiative energy fluxes. Biosphere properties are strongly dependent on the temporal dynamics and spatial distribution of land surface vegetation. Vegetation therefore indirectly drives physical and chemical processes in the atmosphere by altering the transfers of sensible heat, water and momentum between the ground surface and the planetary boundary layer. The vegetation physiology itself is very sensitive to climate forcings and therefore provides several feedback mechanisms between the land surface and the atmosphere (Bounoua et al. 1999).

Biosphere properties need to be considered in climate simulations to account for the interactions mentioned above. The formulation of these land surface parameterizations (LSPs) has evolved in the last two decades from simple prescriptions of surface parameters (BATS of Dickinson 1984, SiB of Sellers et al. 1986) to very sophisticated physical and biochemical models (Sellers et al. 1996a), that include advances in plant physiology research made during the 1980s and 1990s (Farquhar et al. 1980, Collatz et al. 1998). These advanced biosphere models are coupled with atmospheric and ocean general circulation models (AGCMs) and regional circulation models (RCMs), such that climatic responses can be simulated with altered global or regional change scenarios (Sellers et al. 1997, Heck 1999, Pielke et al. 1999). Los et al. (2000) shows that changes in the global vegetation activity strongly affect the land surface hydrology and also produce a significant change in the land surface energy balance.

Table 1.1: Sensitivity of land surface parameters to increased FPAR (Los et al. 2000).

Driving parameter	30°S-30°N	40°N-70°N	Global
Δ FPAR (-)	0.148	0.183	0.157
Affected climate parameters	30°S-30°N	40°N-70°N	Global
Δ Assimilation ($\text{mol/m}^2/\text{yr}$)	33.7	26	28.5
Δ Precipitation (mm/yr)	11.1	33.7	19.1
Δ Evapotranspiration (mm/yr)	50.9	48.8	46.1
Δ Soil water (mm/yr)	-59.7	-27.7	-41.8
Δ Soil latent heat flux (W/m^2)	-8.11	-4.28	-6.1
Δ Soil sensible heat flux (W/m^2)	-9.03	-6.41	-7.02
Δ Canopy latent heat flux (W/m^2)	12.2	8.16	9.77
Δ Canopy sensible heat flux (W/m^2)	4.81	2.51	3.52

Increased FPAR¹ leads to an increased carbon assimilation, which enhances the transpiration and indirectly results in an increased precipitation rate, as seen in Table 1.1. Soil fluxes decrease, when FPAR increases. A more differentiated view is provided by Heck (1999) with a sensitivity experiment of the European climate with altered continental-scale vegetation. Northern Europe soil-atmosphere fluxes respond with the same sign to increased vegetation as shown in Table 1.1. In the Mediterranean region increased vegetation leads to moister and cooler spring but are followed by a warmer and drier summer with very low soil moisture available. This study shows that regional climate modeling provides a better understanding of local climate response to continental-scale vegetation changes.

The current third generation SVATS (SiB2 by Sellers et al. 1996a, SiB 2.5 by Vidale, personal communication) formulate biosphere processes from time varying fields of vegetation, static land cover types and soil properties. Vegetation data is not collected very frequently and with a large area coverage by ground measurements as this is done for other climate data like rainfall and temperature. This fact poses a serious limitation to the estimation of biospheric processes in climate simulations. Several methods have been developed to obtain vegetation data from satellite observations. These methods are based on the unique spectral behavior of vegetation: leaf material absorbs radiation in the visible and strongly reflects radiation in the near-infrared part of the solar spectrum. The normalized difference vegetation index reflects this response: $NDVI = (\rho_N - \rho_V) / (\rho_N + \rho_V)$, where ρ_V is the visible land surface reflectance and ρ_N is the near-infrared land surface reflectance (Tucker and Sellers 1986). During the past two decades polar orbiting satellites from the National Oceanic and Atmospheric Administration (NOAA) have been collecting data in these spectral bands using the Advanced Very High Resolution Radiometer (AVHRR) instrument. These measurements provide high temporal and spatial coverage to access vegetation behavior of regional and global scales for climate studies.

Although potential use of satellite remote sensing for the estimation of land surface properties looks very promising, spatial and temporal consistency of these datasets needs to be carefully reviewed. Serious compromises are found due to sensor degradation, clouds, viewing geometry, registration errors and atmospheric effects (Holben 1986, Goward et al. 1991, Gutman and Ignatov 1995). Many approaches have been undertaken to generate corrected global and regional vegetation datasets for use in land surface parameterizations, such as the NOAA Global Vegetation Index (GVI) dataset (Goward et al. 1993 and 1994), the NOAA Land Pathfinder dataset (James and Kalluri 1994) and the Global Inventory Monitoring and Modelling System (GIMMS) dataset (Los et al. 1994). The estimation of land surface properties from the 1° by 1° global FASIR corrected GIMMS NDVI dataset has been successfully performed by Sellers et al. (1996a) and validations were carried out at field experiments (FIFE: Sellers et al. 1988, OTTER: Angelici et al. 1991, BOREAS: Hall et al. 1993).

The intention of the present thesis was to use existing vegetation datasets for the estimation of land surface properties at regional scale. They will improve the land surface parameterization scheme SiB2.5, which is in preparation for being coupled to the new ETH Regional Climate Model (CHRM by Lüthi et al. 1997). At the current state, no multi-year regional vegetation dataset for the European domain is available for climatic use. The GIMMS FASIR-NDVI (Los et al. 2000) will be extended for 0.25° grid resolution and monthly time steps earliest until summer 2002.

1. FPAR: Fraction of Photosynthetically Active Radiation absorbed by the green vegetation, part of the solar spectrum, which is used for the photosynthetic carbon assimilation process. FPAR is used to estimate the net primary production of land surface vegetation (See section 5.4).

The present work therefore focuses on the generation of a regional European vegetation dataset with associated land surface properties for use in coupled biosphere-atmosphere circulation models like SiB2.5-CHRM. Higher spatial resolutions are needed for a proper estimation of the regional land surface variability in this context. The NOAA AVHRR Pathfinder NDVI dataset, in its maximum 8km¹ spatial resolution, was selected for this purpose. This dataset is continuously available from 1981 to the present and therefore allows the research on inter-annual vegetation variabilities and the estimation of multi-year trends. The temporal scale is set from 1983 until 1993 by selecting 10-day intervals, such that short time features like leaf-out in spring is detected more efficiently than with monthly composites used by Los et al. (2000). This dataset is perfectly scaled for regional climate modelling, but it is heavily contaminated with data dropouts and outliers induced to its temporal and spatial characteristics².

Sellers et al. (1996b) and Los (1998) have shown that it is possible to extract the proper vegetation signal from error-contaminated time-series with a Fourier curve fitting, when the following assumptions are made:

A-1 Vegetation activity varies continuously over time and shows a seasonal variability.

A-2 Errors resulting from clouds and atmospheric contamination tend to decrease the NDVI signal

Considering these, cloud interferences, data dropouts and atmospheric effects can be removed by trusting the high values within the time-series and rejecting (low) outliers. In this study the Fourier adjustment technique by Los et al. (1994), Sellers et al. (1996b) and Los (1998) is revised and carefully adapted to the highly error-contaminated 0.1° by 0.1° Land Pathfinder NDVI dataset. Several extracted cases show that the original Fourier adjustment technique does not hold well for the present dataset. Positive outliers are observed, which is in disagreement with the second assumption, A-2. By taking into account these cases and by modifying the weighting scheme in the Fourier adjustment, a consistent multi-year NDVI dataset for the European domain is the main result of the present work.

To examine the quality of the created vegetation dataset, it is compared to the lower resolution ISLSCP II initiative (International Satellite Land Surface Climatology Project II: Los et al. 2000) data collection. General trends in the NDVI are reproduced well, although dataset properties and processing steps were quite different for the two datasets. Differences in the NDVI amplitude and slight phase shifts are observed between the two products, what needs further examination. The vegetation product shows inter-annual variability in the seasonal amplitude and in the length of the growing season, what looks promising for its use in sensitivity experiments in regional climate scenarios. A positive NDVI trend of 1.4% per year is detected in the examined time period from 1983 to 1993, what can relate to an observed increase of the net primary production in northern hemisphere (Myneni et al. 1997). However, such decadal trends from satellite remote sensing are subject to measurement inconsistencies with the same magnitude caused by successive replacement of the satellites (Hurrell and Trenberth 1997). In contrast, vegetation-type dependent seasonality shows high amplitude differences as observed in the presented vegetation dataset and matches well with vegetation type characteristics described by Justice et al. (1985), Moulin et al. (1997) and Fries et al. (1998).

1. 8km spatial resolution at nadir corresponds to approximately an 0.1° lat/lon grid cell at 45° latitude.

2. Having datasets spatially and temporally subsampled reduces errors induced by measurement difficulties (Section 2.2) and also reduces high frequency components within the dataset.

By combining the time- and space-dependent vegetation (NDVI) product with land cover and soil type maps, land surface properties can be assigned or calculated with simple relationships and intermediate models¹ (see Figure 1.1). In the present work, a remotely-sensed land cover type map by DeFries et al. (1998) is used to estimate vegetation-type dependent properties and the FAO soil type map (FAO, 1978) is applied to have soil properties assigned.

Time-dependent fields like the fraction of absorbed photosynthetically active radiation ($FPAR$), the total leaf area index (L_T), vegetation cover fraction (f_V), the greenness fraction (N) and roughness length (z_0) are calculated using a methodology by Los (1998) and Los et al. (2000) and are presented without further examination of the products in the appendix C of this work.

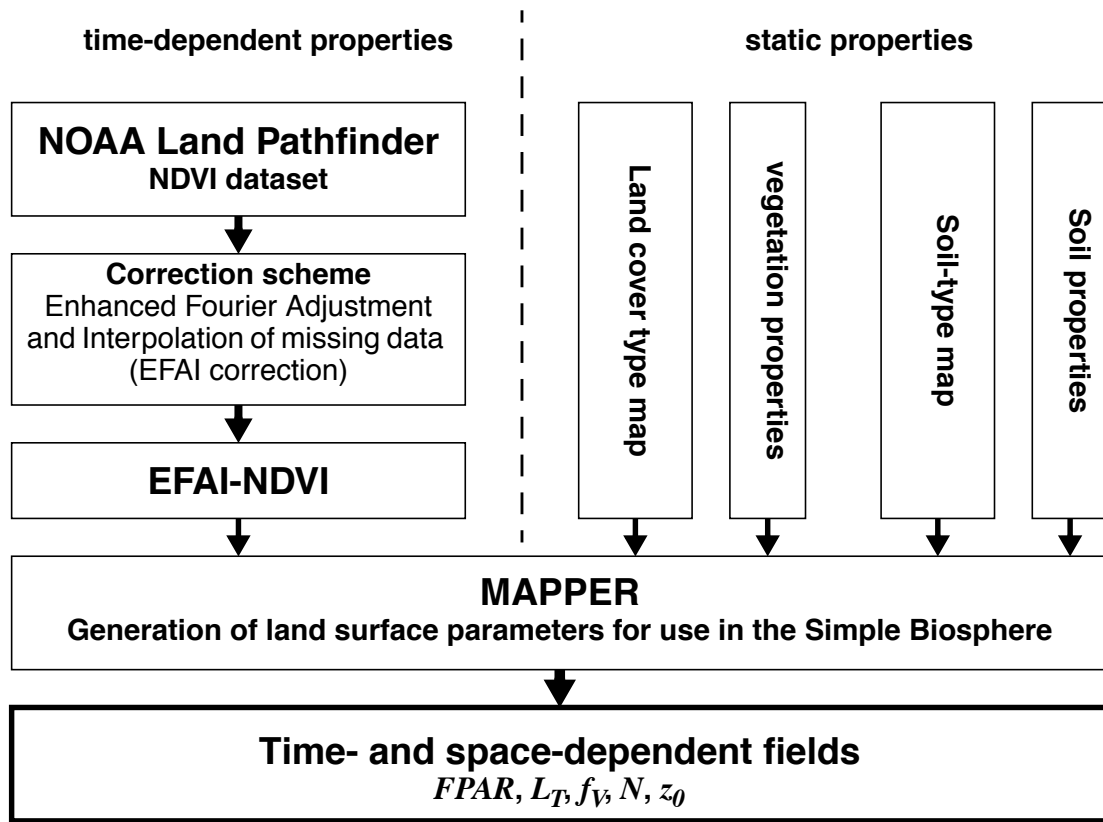


Figure 1.1: The calculation of land surface properties combines time-dependent NDVI-fields with static land cover and soil type maps through simple relationships (presented in section 5.5). The accuracy of the NDVI data is important, since all land surface properties are first order dependent on NDVI. The present work focuses on the generation of the EFAI-NDVI from the NOAA Pathfinder NDVI dataset. Resulting land surface properties, calculated with the SiB-preprocessor “MAPPER”, are presented in the appendix C.

1. There are alternatives to the NDVI-based land surface parameters described in this work: A multitude of specialized vegetation classes, which all have their own phenology curves. Example is the USGS vegetation classification for the USA with over 270 classes. A practical problem rises from the fact, that no consistent parameters are available for most of these classes.

The remaining chapters are arranged as follows: chapter 2 outlines the theory behind remote sensing of vegetation from space and discusses the associated difficulties and available solutions. The Fourier adjustment technique is presented. In chapter 3 the Fourier adjustment technique is applied to generate an 11-year NDVI dataset. Various special cases are evaluated and modifications to the original algorithm are discussed and applied. In chapter 4 the resulting product is compared to the few available sources and its usability is discussed. Future updates in the generation of vegetation datasets are outlined. In Chapter 5 the calculation of land surface parameters, resulting from the generated EFAI-NDVI, is outlined. Chapter 5 then focuses on collecting and reclassifying the land cover and soil type maps from public available sources to fit the used calculation schemes. The program MAPPER is then used to produce European fields of time- and space-varying land surface parameters. Chapter 6 provides final summarized thoughts and a future outlook from the present work. The appendix then shows tables and maps used and presents the land surface fields calculated in chapters 3 and 5.

2. Satellite Remote Sensing of Vegetation

An overview of vegetation monitoring with remotely-sensed data from the NOAA Advanced Very High Resolution Radiometer (AVHRR) is provided in this chapter. Spectral absorption characteristics of plants allow to effectively estimate their spatial and temporal distribution from satellite sensors with empirical relationships. This technique provides global coverage for short time intervals at reasonable resolution, but data is subject to several external interferences and systematic measurement errors.

These inconsistencies in remotely-sensed vegetation datasets are discussed for the NOAA Land Pathfinder dataset used in this work. The magnitude of errors within the dataset can be significantly reduced with an adjustment procedure that bases on Fourier series with least squares fitting. Two assumptions are made to distinguish unreliable data from the vegetation signal, what serves to reconstruct time-series in remotely-sensed vegetation data. This adjustment procedure is very effective at simulated time-series and shows better performance when the temporal resolution of the time-series is increased.

2.1 Spectral Properties of Land Surface Vegetation

There is currently much interest in the quantitative characterization of temporal and spatial vegetation patterns with remotely-sensed data for the study of earth science and global change. During the past two decades, various techniques and algorithms have been evaluated to create accurate datasets with global as well as regional scale. They all base on the discontinuity of the spectral absorption curve for green vegetation around $0.7\ \mu\text{m}$ (Gutman 1991). Light is absorbed strongly within the green leaves by the plant pigments present. Most of this absorption is due to chlorophyll a and chlorophyll b as well as the carotenoids in different spectral bands (Figure 2.1): Most of the photosynthetically activity is found between $0.4 - 0.5\ \mu\text{m}$ (blue) and $0.62 - 0.7\ \mu\text{m}$ (red). The range from $0.5 - 0.62\ \mu\text{m}$ (green, yellow) has reduced chlorophyll absorption and results in the green appearance of vegetation to our eyes. The spectral band $0.74 - 1.1\ \mu\text{m}$ (infrared) is highly reflected by the plant leaves, especially for dense canopies (Figure 2.2). Higher wavelengths above $1.1\ \mu\text{m}$ show an increasing absorption curve by the plants liquid water as described by Tucker and Sellers (1986).

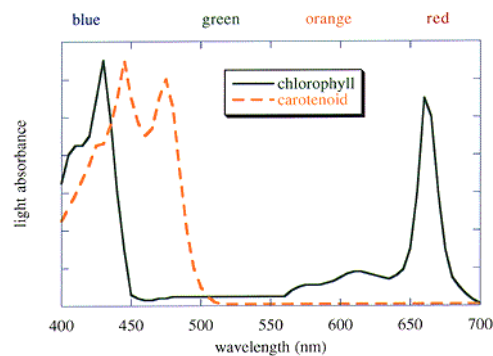


Figure 2.1: Absorption spectrum of isolated chlorophyll and carotenoid species. The colors associated with the various wavelengths are indicated above the graph. (Vermaas 1998)

The spectral characteristics of green leaves can be used to determine the amount of photosynthetic active biomass by combining different wavelengths to vegetation indices. The information contained in a single spectral band is usually insufficient to characterize vegetation status, since atmospheric constituents and soil reflectances can have similar spectral responses at certain wavelengths (Qi et al. 1994).

Two remotely-sensed vegetation indices are in common use, defined as: (Sellers et al. 1996b)

$$SR = \frac{\rho_N}{\rho_V} \quad (2.1)$$

$$NDVI = \frac{\rho_N - \rho_V}{\rho_N + \rho_V} \quad (2.2)$$

with

SR = Simple Ratio

NDVI = Normalized Difference Vegetation Index

ρ_V, ρ_N = reflectances or radiances, for visible (VIS) and near-infrared (NIR) wavelength

High photosynthetic activity will result as a low reflectance in the visible band ρ_V and high response in the near-infrared wavelengths ρ_N . Although leaf spectra vary (Figure 2.2), the large differences between red and infrared reflectance occur for all photosynthetically active leaves and this allows to distinguish leaves from other remotely-sensed objects, such as water, soils and clouds. Both, SR and NDVI, show a high response to green vegetation and have higher values at increasing photosynthetic activity. The NDVI value generally varies between 0 (sparse vegetation) and 0.8 (very dense vegetation), where negative values represent non-vegetated surfaces like snow and water, but also occur due to cloud contamination. Compared to the reflectance of leaves, the reflectance of soil in the infrared doesn't differ much from its visible response (Los 1998) and shows a linear relationship. Nevertheless, spectral response of the soil vary with its type and composition (% organic matter, % saturation, amount of iron) and can have a significant contribution to the NDVI signal (Qi et al. 1994). As measurements of the Sahara desert revealed, remotely-sensed vegetation indices can vary up to 0.1 NDVI with a standard deviation of 0.02 NDVI (Los 1998) - variations, that are attributed to soil effects.

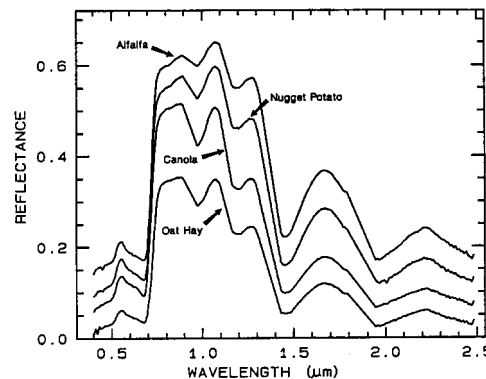


Figure 2.2: Absorption (reflectance) curves of different vegetation types in relation to wavelength. Each curve has been offset by 0.05 from the one below. Data was extracted from airborne AVIRIS measurements. (Clark et al. 1995, USGS)

Several vegetation indices have been proposed to be less sensitive to the soil background effects, such as the Soil-Adjusted Vegetation Index SAVI (Huete and Tucker 1991). The SAVI applies a soil

adjustment factor L to the visible and near-infrared reflectances to account for the soil effects. As the soil background effects only are of significance in sparse vegetated areas, L should vary inversely with the amount of vegetation present (the fraction of vegetation cover). Qi et al. (1994) developed the Modified SAVI (MSAVI), where the vegetation signal to soil noise ratio is significantly enhanced.

2.2 Vegetation Monitoring with the NOAA AVHRR Sensor

As described in section 2.1, vegetation activity can be measured by taking the ratio between near-infrared and the visible (red) wavelengths of the radiation reflected at the earth surface. These spectral bands can be remotely-sensed by ground equipment, airplane carried sensors and satellites. Different satellites have been equipped with passive scanners, which collect data at these spectral bands. Since 1979 the Advanced Very High Resolution Radiometer (AVHRR) onboard the National Oceanic and Atmospheric Administration (NOAA) series of Polar orbiting Operational Environmental Satellites (POES) is collecting data over the entire globe at 4km^1 resolution twice daily in five different spectral bands (Table 2.1).

Table 2.1: NOAA AVHRR instrument channels

Channel	Wavelength [μm]	Description	Detection of
1	0.58-0.68	Visible, red	Clouds, Vegetation, pollution
2	0.73-1.10	Near Infrared	Vegetation
3	3.55-3.93	Shortwave Infrared	Fires, water clouds, fog, SST
4	10.3-11.3	Longwave Infrared	Emitted radiation, clouds
5	11.5-12.5	Longwave Infrared	Emitted radiation, moisture

This frequent global coverage makes AVHRR data suitable for applications, that need high temporal resolution and coverage over large areas. Regional and local vegetation-studies with higher spatial resolutions will rely on remotely-sensed data from Landsat or SPOT (typically acquired only a few times a year) as well as on aerial photography (Los 1998).

For studies of broad scale land surface, satellite remote sensing with the AVHRR instrument has various advantages, like high revisit frequency and very large coverage, which also result in low spatial resolution (Robinson 1996). Furthermore the AVHRR Global Area Coverage (GAC) data is continuously available since 1981, making it a very suitable source for inter-annual and multi-year climate studies in the biosphere field (Prince and Goward 1996).

Nevertheless, vegetation characteristics from satellite-acquired surface data have to be analyzed carefully and need corrections for the various error sources resulting from this technique. Much research has been done in this field during the last two decades (Tucker and Matson 1985, Holben 1986, Gutman and Ignatov 1995, Cihlar et al. 1994) and only the main results will be discussed here as an overview.

-
1. The instantaneous field-of-view (IFOV) of 1.4 milliradians provides a ground resolution of approximately 1.1 km at the satellite nadir from the orbit altitude of 833 km. The data are derived from an onboard sample averaging to 4-km resolution, since off-nadir pixels will cover larger surface areas..

The effects to be considered when working with AVHRR data fall into two categories: atmospheric and geometric. The two principal atmospheric effects are 1) decrease in transmissivity due to absorption (ozone, water vapor) and scattering (aerosols, molecules) present in the atmosphere (Cihlar et al. 1997; Sellers et al. 1996b) and 2) cloud cover, which can seriously affect the usability of satellite sensed surface-data in certain regions like Amazonia (Holben 1986). The key geometric effects are the bidirectional reflectance distribution function (BRDF), which is a function of the relative positions of the satellite, the sun and the remotely-sensed surface and is also dependent of the surface properties as well as of the used wavelength (Cihlar et al. 1994). The knowledge of the BRDF and its effects on the measured surface variable is essential for each remotely-sensed dataset, since albedo and surface radiant fluxes are highly sensitive to incident and reflected radiation properties. This bidirectional reflectance distribution function is usually quantified by semi-empirical models to correct satellite data (Gutman 1991, Cihlar et al. 1994). It is not well understood and hasn't been implemented in most of the currently available remotely-sensed vegetation datasets.

The NOAA POES satellites have an afternoon equatorial crossing time between 14:30 and 17:30. At this time, convective cloud systems over land surfaces have evolved and cover a large part of the ground surface. This poses a heavy limitation to acquire vegetation data in tropical regions. (Prince and Goward 1996)

Instruments also show certain degradations over time, that can be described with empirical functions through in-orbit calibrations (e.g. calibration at time-invariant sites). Orbital drift leads to a gradual shift in the equatorial crossing time and successive replacements of the satellites introduce non-steady transitions in time-series. These effects need to be considered, when several years of AVHRR data are processed.

Combining several AVHRR bands in ratios (like NDVI, SR) reduces atmospheric and geometric effects (Holben 1986). Generating spatial and/or temporal composites has the same effect, whereas several algorithms can be applied.

2.3 The Pathfinder NDVI Dataset

The Pathfinder Program, initiated by NOAA and NASA, produces long-term datasets processed in a consistent manner for global change research. The data covers the period from July 1981 through the present. It is available to the public and provides daily and 10-day composite GAC NDVI data with a resolution of 8x8km.

The Pathfinder datasets have been corrected for the following errors and effects:
(James and Kalluri 1994)

- Sensor degradation
- Atmospheric effects (Rayleigh scattering and Ozone)
- Surface topography
- Solar zenith angle and azimuth
- Cloud contamination

Each scanned pixel is navigated through an orbital model and atmospheric as well as bidirectional reflectance distribution effects are calculated considering surface topography and the relative sun-target-scanner positions. Clouds are detected by the thermal channel (AVHRR channel 4) and flagged

to missing data. To produce the 10-day composite data and reduce cloud contaminations and bad scans, the maximum-value composite (MVC) algorithm is used (Holben 1986).

The use of the AVHRR Pathfinder dataset has several advantages, like its continuous and global coverage, as well as being a public available source. Several studies have discussed the accuracy of this dataset for climatic change research (Anyamba and Eastman 1996, Potter and Brooks 1998, Los et al. 2000), for environmental monitoring (Belward 1992), crop yield estimation (Quarmby et al. 1993) and for land cover classification (DeFries et al. 1998).

It was shown, that the atmospheric correction scheme used by the Pathfinder project science team only tend to reduce the NDVI linearly (Prince and Goward 1996). No corrections for water vapor and aerosols are performed at all in the NOAA Pathfinder dataset. The maximum-value composite algorithm favors off-nadir viewing angles, which results in an overall quality decrease of the NDVI values (Gutman 1991). Other compositing techniques like the “best index slope extraction filtering” (Moulin et al. 1997) and “maximum-value interpolated time-series” (Taddei 1997) have been discussed to solve MVC-induced errors. Deviations of the NDVI values due to soil background reflectances (See section 2.2) are not corrected in the Pathfinder dataset.

2.4 Analysis of Pathfinder NDVI Time-Series

The original Pathfinder NDVI data contain useful information about the temporal and spatial distribution of vegetation. Figure 2.3 shows two examples of one year NDVI timeseries. The above discussed relationship between vegetation activity and spectral vegetation indices is apparent: the observed Sahara site (left) has no vegetation cover throughout the year. It is well reproduced by a low NDVI curve with very little background noise. A deciduous forest in Norway (right) is expected to have a significant seasonal vegetation trend and a very late growing season - these features are obviously reproduced well by the NDVI time-series in a qualitative manner.

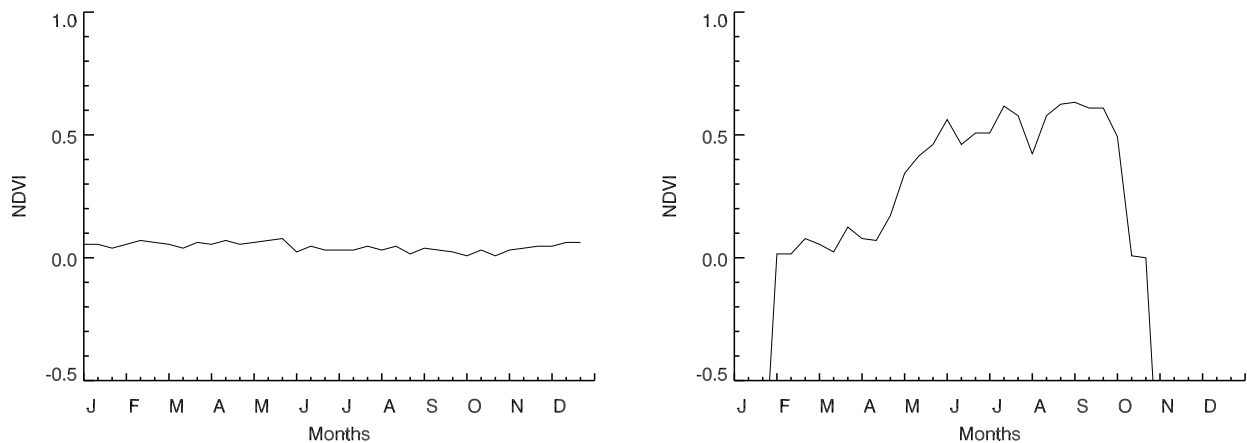


Figure 2.3: Desert in north Africa shows a constant and low NDVI time-signal throughout the year (left, Algerian desert site at 34°N, time-series from 1987), whereas the northern parts of Europe usually have a high seasonality in the vegetation cycle (right, Norway boreal forest site at 65°N, time-series from 1993). Both graphs show time-series for a single pixel.

However, closer analysis of Pathfinder NDVI time-series reveals, that they need further processing to be useful for quantitatively represent the vegetation activity. There are several problems, which can be discovered in the Pathfinder dataset and need to be addressed:

- NDVI series do not vary smoothly with time: high frequency components and missing data are a common phenomena (Figure 2.4 left).
- Positive outliers occur in almost all (except in non-desert) time-series (Figure 2.5)
- High latitudes have extreme solar zenith angles or receive no solar irradiation at all during several months in the winter season. No visible radiation will be measured at these locations for up to six months per year, which leads to missing NDVI data during those months (Figure 2.4 right, Figure 2.6 left).
- Solar illumination, orbital geometry and communication errors have not been corrected successfully in all the datasets. Scan stripes from the satellite's swath-path are visible in certain datasets (Figure 2.6 right).

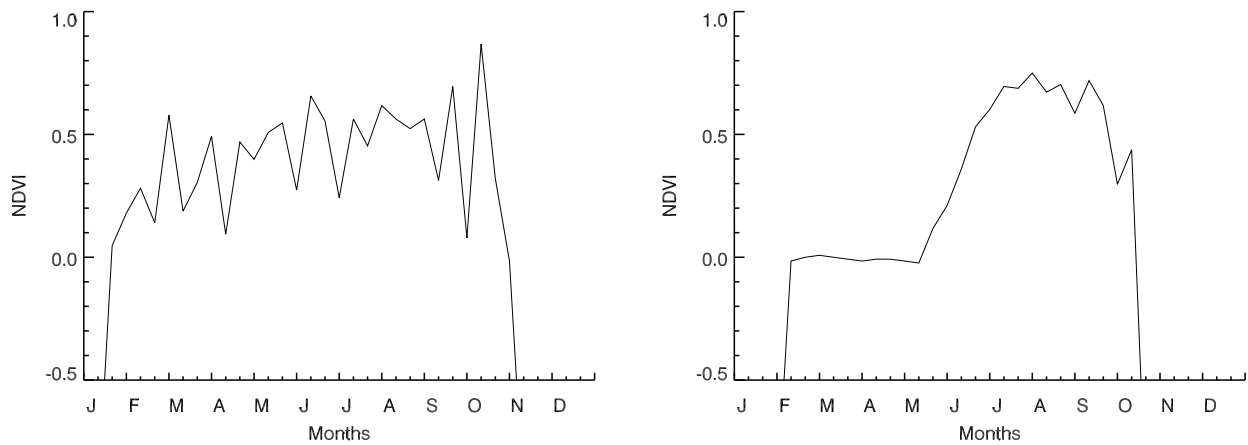


Figure 2.4: High noise observed in the NOAA Land Pathfinder dataset (left: central Scotland time-series from 1993) and data dropout in winter time at high latitudes (right: northern Sweden time-series from 1993). Both graphs show time-series for one single pixel.

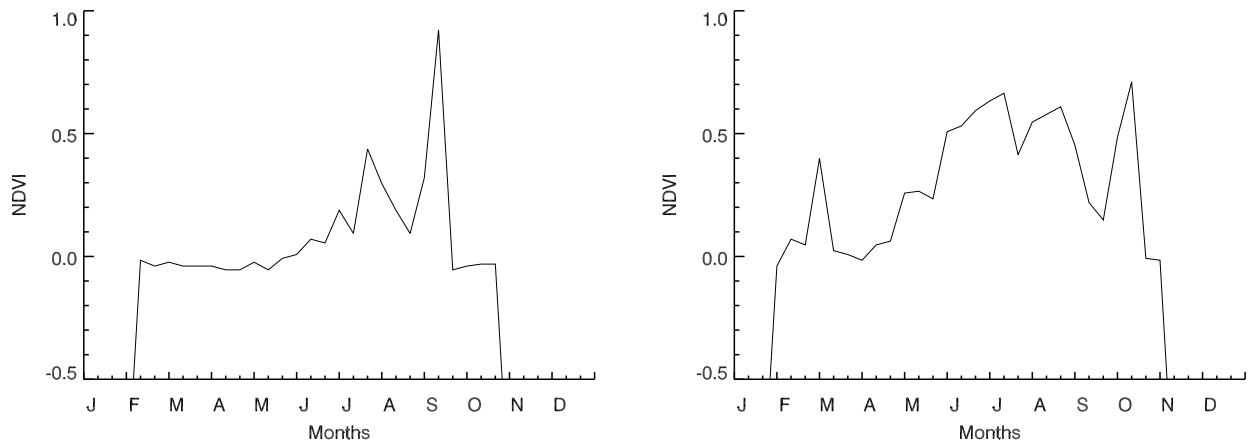


Figure 2.5: No clear vegetation cycle visible. Both graphs illustrate remote sensing errors that have to be corrected prior to use the Pathfinder dataset for deriving biophysical land surface variables. (left: Iceland time-series from 1983, right: Finland time-series from 1983). Both graphs show time-series for one single pixel.

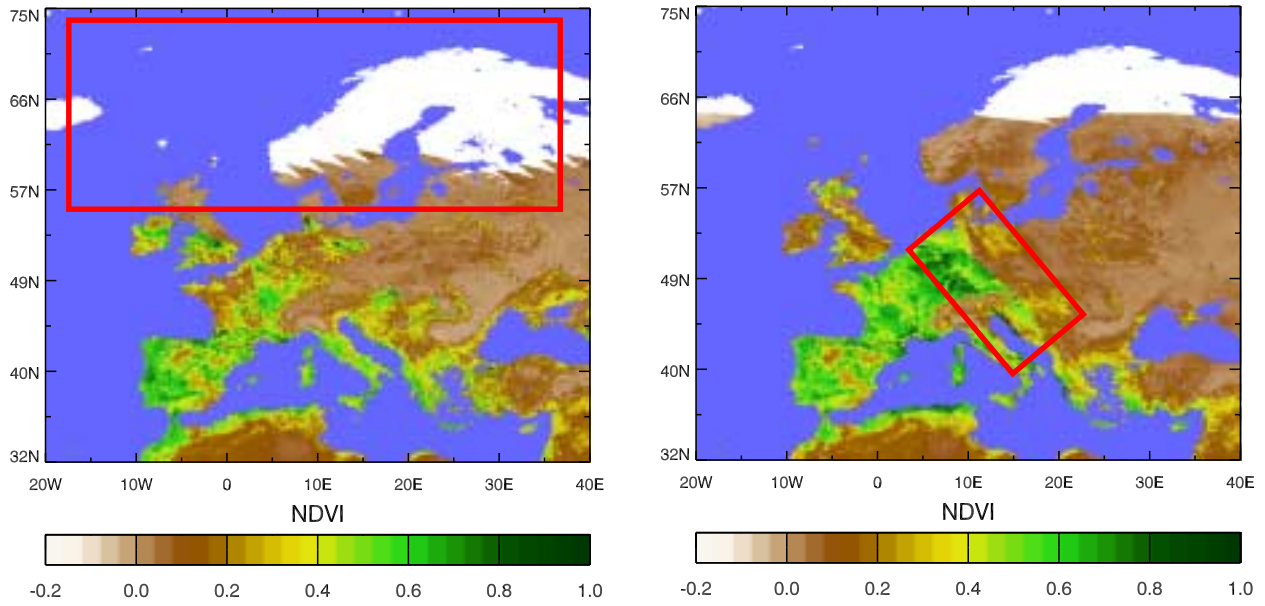


Figure 2.6: No remote sensing of vegetation is possible in winter at high latitudes (left: 1985 January 21-31 composite), what leads to data dropouts during various months. Visually inspecting the Pathfinder datasets shows artifacts like visible swath-paths of the AVHRR sensor (right: 1985 February 1-10 composite).

2.5 Fourier Adjustment of NDVI Time-Series

2.5.1 Introduction

Two assumptions can be made for estimating the quality of a specific NDVI time-series and to apply corrections for the error sources. These assumptions form the basis and main justification for the discussed algorithms in this chapter and their application in chapter 3 to effectively estimate land surface vegetation activity by remote sensing of their spectral absorption characteristics:

- A-1 The vegetation follows a truncated seasonal cycle (Moulin et al. 1991) and NDVI values will vary smoothly with time (Sellers et al. 1996b).**
- A-2 Outliers in NDVI time-series are the result of either cloud cover or sudden changes in atmospheric constituents. These effects tend to decrease NDVI values (Holben 1986, Los 1998).¹**

The objective of the proceeding sections 2.5.2 to 2.5.5 is to describe an algorithm which can successfully reduce errors, data dropouts and other temporal inconsistencies in the Pathfinder NDVI dataset. To achieve this aim the two essential assumptions A-1 and A-2 from above are included in a weighted least squares fit of partial Fourier series.

2.5.2 Fourier series

A Fourier series is suitable to represent the seasonal character of vegetation, because vegetation usually shows a periodic cycle, associated with the seasonal varying environmental conditions. Fourier series arise from the task of representing such a given periodic function by discrete trigonometric series. A Fourier series can be described as follows (partial sum Fourier series; Kreyszig 1988):

$$Y_i = a_0 + \sum_{j=1}^{m-1} a_j \cos(j\phi_i) + b_j \sin(j\phi_i) \quad (2.3)$$

with

$$\phi_i = 2\pi \frac{(i-1)}{n}$$

1. Clouds have already been masked out in the original Pathfinder NDVI dataset. However, not all cloud contamination could be removed with the algorithm, which bases on cloud detection with the thermal infrared AVHRR channel (cirrus and cumulus clouds) and the visible channel (warm, low-level clouds). Clouds and other sources of atmospheric contamination tend to have higher reflections in the visible wavelength and will generally decrease the NDVI values. This assumption also justifies the generation of maximum-value composites (Holben 1986) or maximum-value interpolated composites (Taddei 1997), which implies trusting in high NDVI values when looking at time-series. A few exceptions to this assumption exist, but their effects are an order of magnitude smaller of the general case (Los 1998).

where

- a, b = Fourier coefficients
- n = number of points in the sequence
- i = phase; ranges from 1 to n
- m = number of harmonics

With increasing m, higher harmonics are considered in the Fourier series and short term irregularities (resulting in high frequencies) will be successfully approximated by the series. To represent the seasonal vegetation variation, m=3 was selected¹ (Sellers et al. 1996b). The application of Fourier series to one year of monthly composites results in 12 data points, hence n=12.²

To have a Fourier curve fitted through the measured dataset, the Fourier coefficients have to be calculated. A robust least-squares method is used here to eliminate data, which deviates strongly from the general trend (Weisberg 1985). Sellers et al. (1996b) and Los (1998) have modified the robust least-squares method to take into account the assumption A-2 (see section 2.5.1): errors in the remotely-sensed NDVI mainly result in decreased values.

2.5.3 Fourier adjustment algorithm

The original Fourier adjustment algorithm for NDVI time series was developed by Sellers et al. (1996b), using Global Inventory, Modeling, and Monitoring System (GIMMS) monthly composite global NDVI with a spatial resolution of 1° by 1°. The Fourier adjustment algorithm will be outlined here shortly, to provide a background for its application in the third chapter. A complete description can be found in Sellers et al. (1996b) and Los (1998).

2.5.4 First Fourier adjustment

Using the least-squares method (Weisberg 1985), the Fourier constants c_j are solved for a Fourier series with the degree m-1:

$$(\mathbf{F}^T \cdot \mathbf{F}) \cdot \mathbf{c} = \mathbf{F}^T \cdot \mathbf{Y} \quad (2.4)$$

where $\mathbf{Y} = \begin{bmatrix} Y_1 \\ \dots \\ Y_n \end{bmatrix}$ are the observed data,

-
1. If more terms are taken into consideration, more complex time-series can be approximated by partial Fourier series. m=1 will show a constant function, where m=2 results in a function able to represent annual features. Only Fourier series with m=3 or higher account for halfperiodic features (See Kreyszig 1988).
 2. The correction is limited to less than a sequence of outliers of $n/2^{m-1}$ data points in a row. For m=3, n=12 no more than 3 successive outliers will be corrected.

$$\mathbf{F} = \begin{bmatrix} 1 & \cos(\phi_1) & \sin(\phi_1) & \dots & \cos((m-1)\phi_1) & \sin((m-1)\phi_1) \\ & & & \dots & & \\ 1 & \cos(\phi_n) & \sin(\phi_n) & \dots & \cos((m-1)\phi_n) & \sin((m-1)\phi_n) \end{bmatrix} \text{ is the Fourier matrix}$$

$$\text{and } \mathbf{c} = \begin{bmatrix} c_1 \\ \dots \\ c_{m-1} \end{bmatrix} \text{ are the Fourier constants to be solved for and applied to the Fourier series}$$

Fitting a Fourier curve with this generalized least squares algorithm means, that the sum of the squares of the distances between the fitted and the given data points is minimum (Kreyszig 1988).

2.5.5 Apply a second weighted Fourier adjustment

A second Fourier curve fit is applied through the data, now including the calculated weights W_i :

$$(\mathbf{F}_w^T \cdot \mathbf{F}_w) \cdot \mathbf{c} = \mathbf{F}_w^T \cdot \mathbf{Y}_w \quad (2.5)$$

$$\text{where } \mathbf{Y}_w = \begin{bmatrix} W_1 Y_1 \\ \dots \\ W_n Y_n \end{bmatrix} \text{ are the weighted observed data}$$

$$\text{and } \mathbf{F}_w = \begin{bmatrix} W_1 & W_1 \cos(\phi_1) & W_1 \sin(\phi_1) & \dots & W_1 \cos((m-1)\phi_1) & W_1 \sin((m-1)\phi_1) \\ & & & \dots & & \\ W_n & W_n \cos(\phi_n) & W_n \sin(\phi_n) & \dots & W_n \cos((m-1)\phi_n) & W_n \sin((m-1)\phi_n) \end{bmatrix}$$

is the weighted Fourier matrix.

2.5.6 Weighting the Fourier fit

The weights W_i are calculated according to the distance of the measured data from the fitted curve:

$$W_i = \begin{cases} 0 & U_i \leq -k \\ [1 + ((U_i + r)/k)]^4 & -k < U_i < -r \\ 1 & -r \leq U_i \leq r \\ [1 + ((U_i - r)/k)]^2 & U_i > r \end{cases} \quad \text{if} \quad (2.6)$$

and

$$0 \leq W_{i=1} \leq 1$$

$$0 \leq W_{i=n} \leq 1$$

where

$$\begin{aligned} U &= (Y - \hat{Y})/M \\ M &= \text{Median}(|Y - \hat{Y}|) \\ k &= 2 \\ r &= M / 20 \end{aligned}$$

The errors resulting from the first generalized least squares fit have different variances for each fitted point. Using weights $W_i \geq 0$, certain variances can be taken into account more than others. Cases with high weights will have small variances and are therefore be more important to the regression problem (Weisberg 1985). The minimization of the squared differences used in the first generalized least squares fit thus changes, eliminating the influence of bad data from the regression and interpolating with the reliable points. High weights are applied to all data points lying on, or above the first fitted Fourier curve, where low weights are assigned to the less credible data with low NDVI values (Los 1998).

2.6 Effectiveness of the Fourier Adjustment Technique

The Fourier adjustment procedure discussed above bases on the estimation of a time-series by curve-fitting discrete sample points with least squares and at the same time excluding erroneous points by trusting the higher observed values. Testing of this theory is performed in this section. A simulated (theoretical) vegetation curve is created and discretized with 12 points (Figure 2.7 left, solid line) in comparison to 36 points (Figure 2.7 right, solid line).

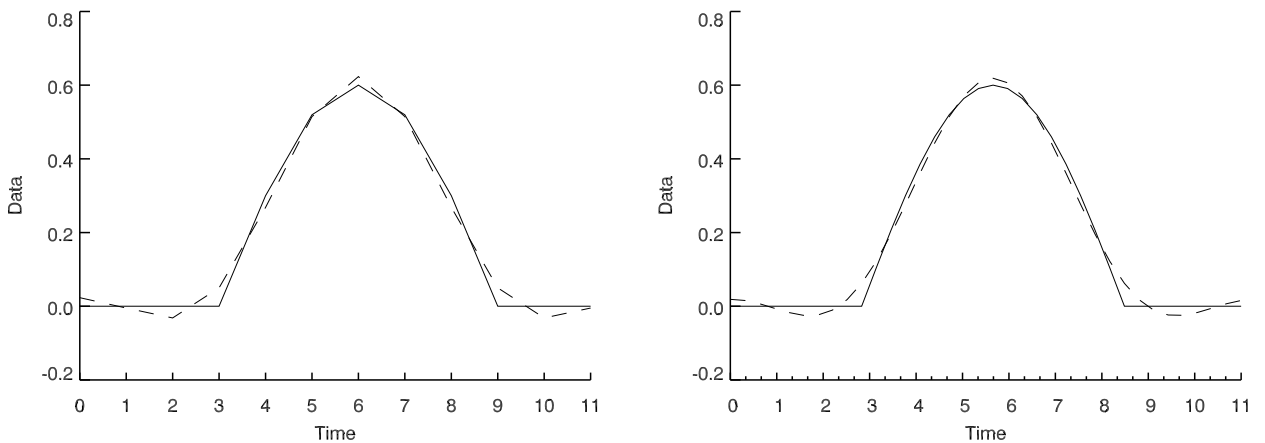


Figure 2.7: Fitting a theoretical vegetation time-series when 12 discrete data points are available (left) and when a higher temporal resolution of 36 values is offered (right). A linear regression of the original and fitted curve results in a better fit ($R^2=0.995$) of the finer scaled data than the coarse time-series ($R^2=0.989$). The original time-series is shown as solid line and the Fourier adjusted curve is dashed.

Figure 2.7 shows that there is a slight improvement in Fourier adjusting 36 points over 12 points. A linear correlation between the original curve and the Fourier adjusted curve show an R^2 of 0.995 for 36 discrete points, whereas an R^2 of 0.989 results, when 12 points are considered.

As soon as erroneous data is included in the simulated time-series, the power of the weighting scheme can be evaluated. Figure 2.8 was created by assuming a data dropout of one month in a yearly time series. A Fourier adjusted time-series with 36 discrete points show a R^2 of 0.990 in comparison to the slightly lower R^2 of 0.986, when 12 points are used. The observed fitting errors do not differ much between the two cases. Considering that the simulated curve is nearly sinusoidal, it can be assumed that the Fourier adjustment will be a very efficient fitting method. The fitting errors are much higher in real data, what will be seen in Chapter 3. Having two months of missing values (Figure 2.9) verifies this assumption, since R^2 decreases from 0.994 with 36 points to 0.922 with a discretization of 12 points.

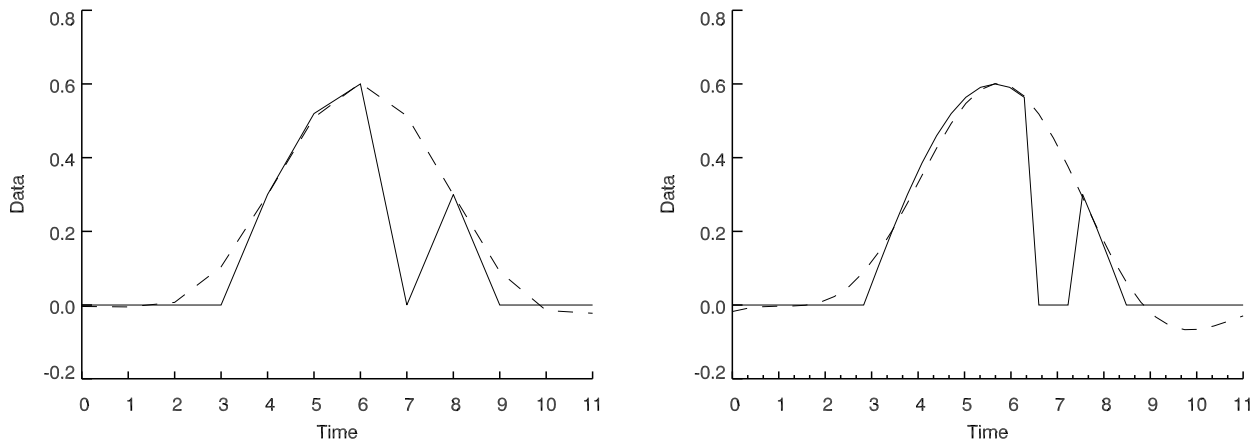


Figure 2.8: Assuming a data dropout of one month in a yearly time-series (solid line), the Fourier adjustment (dashed line) shows better estimation of the original curve (solid line in Figure 2.7), if more discrete values are provided. $R^2=0.986$ for 12 values (left) and 0.990 for 36 values (right).

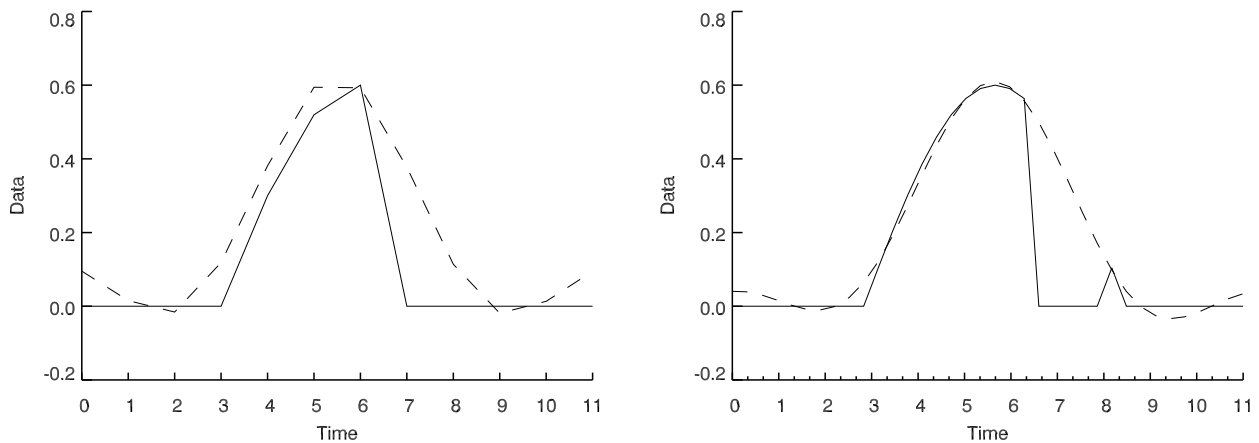


Figure 2.9: Assuming two months of data missing, the Fourier adjusted curve shows a better estimation when more data points are provided. $R^2=0.922$ (left) assuming monthly time-series and $R^2=0.994$ when 10-day composites are considered (right).

The above experiments are all run with a 2nd order Fourier series ($m=3$)¹. Figure 2.10 illustrates, that with 1st order series (a) a yearly anomaly will be simulated by the Fourier adjustment, whereas higher order series (c) will not represent the expected seasonal vegetation cycle anymore. Higher order Fourier series try to adapt to the short-time anomalies instead of extracting the smoothly varying vegetation signal, what is best achieved with 2nd order series as shown in (b).

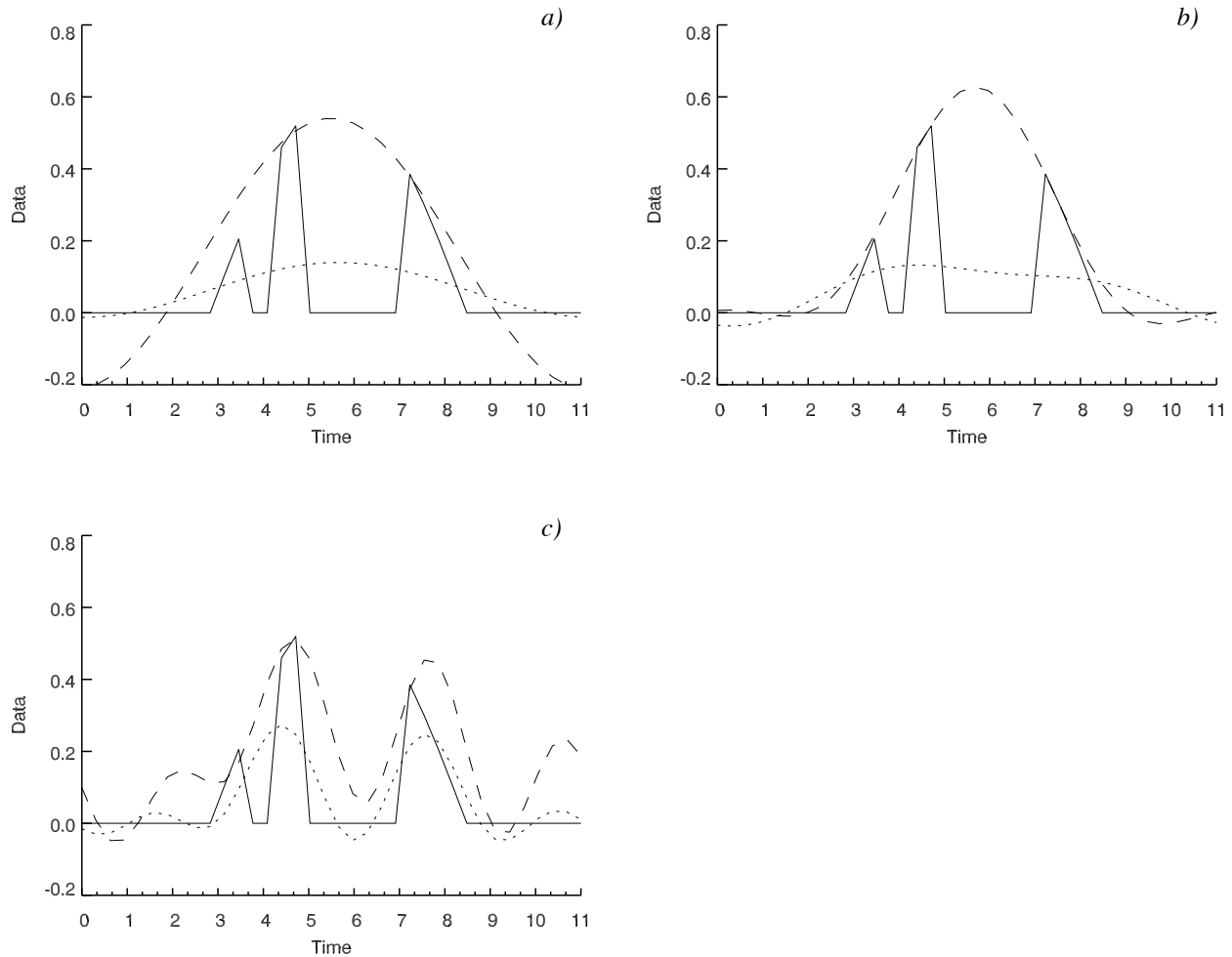


Figure 2.10: It can be demonstrated, that the seasonal feature of vegetation is best approached by a Fourier series with $m=3$ (2nd order) like shown in b). The Fourier adjustment in a) is a 1st order series that can only simulate yearly features. c) shows a 4th order series, that will reproduce high frequency components in the dataset (noise, dropouts). The solid curves represent the raw original data, the dotted line is the first Fourier adjustment and the dashed line show the second weighted Fourier adjustment. A linear regression between the original curve (Figure 2.7 right, solid line) and a) shows a R^2 of 0.912, best fit is achieved by b) with $R^2=0.994$ and c) doesn't represent the seasonal curve at all with $R^2=0.541$.

1. It is discussed and assumed in section 2.5.2, that half-year features like the seasonal variation of vegetation are best approached with $m=3$.

3. The European Fourier-Adjusted and Interpolated NDVI

The main application of the present work leads to a consistent set of biophysical land surface parameters, that will drive the Simple Biosphere Model SiB2 (Sellers et al. 1996a) at the regional scale. These land surface parameters (see section 5.5) show a linear or exponential dependency from vegetation activity, which can be empirically estimated with the remotely-sensed normalized difference vegetation index (NDVI). A correct estimation of NDVI is therefore the most critical task prior to the calculation of the energy, carbon and water fluxes between the surface and the atmosphere. Vegetation varies temporally and spatially at very small scales¹, but it was shown that GCM performance significantly increased by using land surface parameterizations derived from coarse grid, remotely-sensed vegetation data (Los 1998, Los et al. 2000), in comparison to fixed land surface parameterizations (Dickinson 1984, Sellers 1986).

Fung et al. (1997) suggested that subgrid information is essential for a better estimation of land surface processes. Regional climate models (RCMs) have evolved during the recent years and work with finer scales to explore regional climate variations induced either from natural or anthropogenic forcings. Since remotely-sensed vegetation data is available in much higher spatial and temporal resolution than applied in global climate models, this study proposes a 0.1° by 0.1° (approximately 8km pixel size at 47°N) Pathfinder NDVI dataset with a temporal interval of 10 days for regional climate models.

To generate a consistent European vegetation dataset over 11 years, there were several reasons to use the NOAA/NASA Land Pathfinder dataset and to evaluate an own adjustment scheme for making it suitable for regional SVATS:

- GIMMS FASIR-NDVI by Los et al. (1994), which is used by the revised Simple Biosphere model (Sellers et al. 1996a), is only available in coarse resolution and monthly composites. Public availability is restricted to the years 1987 and 1988. A higher spatial and temporal extended FASIR-NDVI is in preparation, but will not be available in final release until Summer 2002 (Collatz 2000, personal communication).
- The NOAA Land Pathfinder NDVI is public domain and can be retrieved by Ftp. It is continuously available from July 1981 until the present.
- The author intends to build local expertise, which can be used to understand the problems associated with and estimate the quality of vegetation products becoming available in the future.

In this chapter the generation of a continuous Fourier-adjusted and interpolated 11-year NDVI dataset for the European domain (EFAI-NDVI) is presented. As illustrated in Figure 3.1, several steps are performed in a linear sequence to create the resulting EFAI-NDVI dataset. The applied procedures will be outlined in the following sections 3.1 - 3.7. Especially sections 3.4 and 3.7 contain the modifications to the original Fourier adjustment procedure provided by Los (1998). Special case studies are presented, where this algorithm does not hold and solutions are evaluated as well as tested.

1. Everyone will confirm this statement by taking a look out of the window (The author assumes here that your window does not offer an outlook to either the Sahara desert or to the dusty red Mars surface).

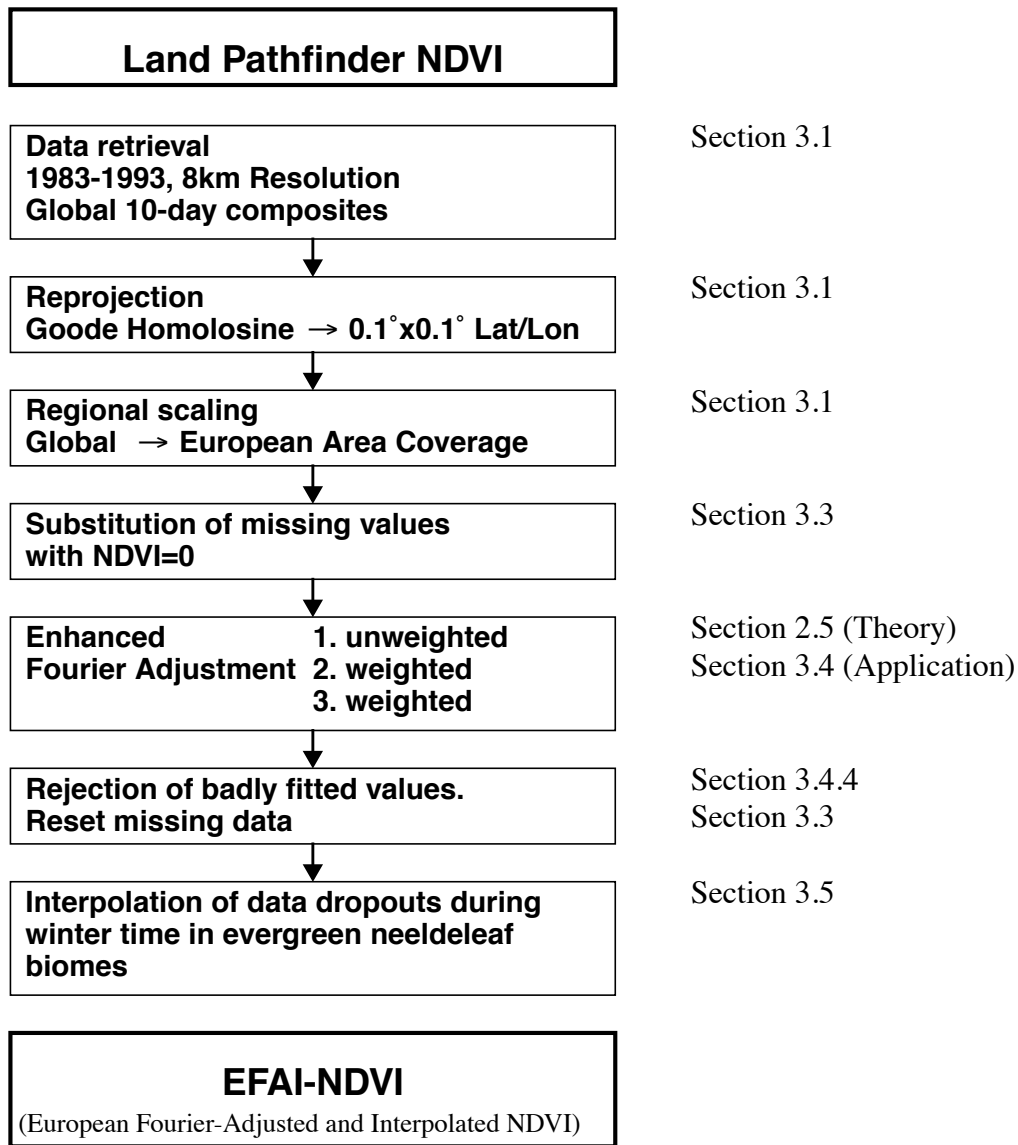


Figure 3.1: A Flow diagram illustrating the processing steps to generate the multi-year EFAI-NDVI.

3.1 Data Retrieval, Reprojection and Scaling

The Pathfinder NDVI dataset was retrieved¹ from the years beginning with 1983 until the end of 1993, divided in 10-day composite files. A practical difficulty in using the data arises from the used equal-area map projection, which is the Goode interrupted homolosine projection introduced to the dataset by Steinwand (1994). Only few image processing tools are able to handle this kind of projection and the interrupted nature produces a data volume 78% larger than is needed to hold the terrestrial data (Prince and Goward 1996).

1. The Land Pathfinder data is available at the GSFC Earth Sciences (GES) Distributed Active Archive Center: ftp://daac.gsfc.nasa.gov/data/avhrr/global_8km

For the entire period, all datasets were reprojected to a regular lat/lon grid with a spatial resolution of $0.1^\circ \times 0.1^\circ$, using an algorithm from Baldwin (1996). This reprojection was implemented in the Interactive Data Language (IDL) by the author in 1998. For all further steps, data amount was reduced by only proceeding with the European domain with a latitude/longitude extent of [32N-75N, 20W-40E]. This domain extent is currently used by the ETH regional climate model (CHRM).

3.2 Programming and Implementation

IDL was chosen as the programming language for the Fourier adjustment, since it includes the advantage of built-in matrix operations, graphic display capabilities and the ability to compile programs as well as testing them at the command line. Disadvantages of IDL include raw processing speed loss and a higher amount of memory use compared to native programming languages like Fortran and C. This is well made up by the very easy use of complex built-in routines and the possibility to interactively display and check datasets through graphical output.

The IDL-program, which applies the Fourier adjustment¹ to the remotely-sensed NDVI dataset, consists of a core routine with the Fourier adjustment algorithm (see 2.5.3) implemented and an input/output routine to handle the 11-year dataset extent.

The core routine applies the Fourier adjustment to a yearly time-series (36 10-day intervals) for each specific latitude/longitude pixel. The pixel neighbors are not considered in the algorithm, i.e. the adjustment technique is applied over time and does not imply any spatial relations among the single pixels. From each corrected yearly time-series, only the middle six months (18 10-day intervals) are used and the next adjustment is shifted by six months to generate a continuous time-series.

This way, “wrap-around” effects from December to January data are avoided and the results show steady transitions between successive years.

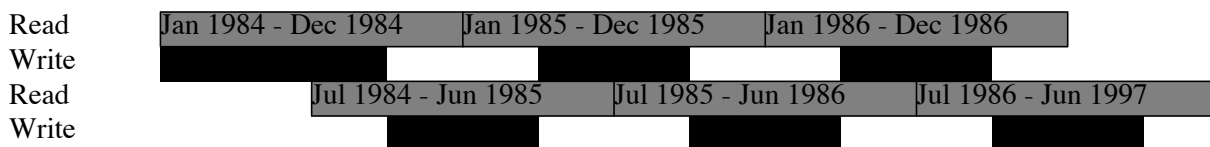


Figure 3.2: The Land Pathfinder data is read and Fourier adjusted for one year, but only the middle six months are written back to avoid wrap-around effects at the sides.

Example:

- 36 NDVI datasets (January 1984 - December 1984) are read and Fourier adjustment is applied.
- 18 NDVI datasets (April 1984 - September 1984) are saved
- 36 NDVI datasets (July 1984 - June 1985) are read and Fourier adjustment is applied
- 18 NDVI datasets (October 1984 - April 1985) are saved

1. The Fourier coefficients were calculated by solving the least squares regression problem described in 2.5.3. Although the common used Fast Fourier Transform (FFT) algorithm has a better performance and is implemented in IDL as a single compiled routine, it does not allow weighted Fourier series.

3.3 Substitution of Missing Data

Prior to any adjustment to the raw data, missing points (cloud flagged and high latitudes during winter season) are substituted with $NDVI = 0$, assuming that the majority of missing values is found during periods of low surface temperature and therefore low vegetation activity (Los 1998).

After the Fourier adjustment algorithm is applied successfully, sequences of eight or more missing values in a row (80 or more days without data) are set to missing again, since they will not be approximated successfully with this configuration¹. Permanent data dropouts either occur in tropical regions (see Los 1998) due to permanent cloud cover or during winter time at high latitudes, where high solar zenith angles - or no sunlight at all - do not allow to use passive remote sensing in the visible wavelength.

Tropical vegetation is not present in the examined region. Whereas high latitudes show substantial data dropouts, what poses a serious limitation on the usability of vegetation data in those regions. A partial solution to this problem is presented in section 3.5.

3.4 An Enhanced Fourier Adjustment Procedure

The Fourier adjustment algorithm by Los (1998) has proved to be reliable and successful to remove cloud interferences and atmospheric effects by identifying outliers in the NDVI time-series². Los has tested the algorithm by removing single months from of his 12-month time-series and reconstructed the incomplete time-series with the Fourier adjustment technique. The results show close agreement between the estimated and the original values. The standard error of all estimates was less than $0.05 * NDVI$. In section 2.6 the author has enhanced this findings by including higher temporal resolution into the theoretical analysis of the Fourier adjustment. Results in section 2.6 indicate that a finer discretization increases the effectiveness of the Fourier adjustment in actually representing the original vegetation signal. This can be seen as a reasonable justification for the use of 10-day intervals in favor of monthly ones.

Nevertheless, the spatial and temporal high resolved NOAA Land Pathfinder NDVI data poses serious problems for a correction with the Fourier adjustment provided by Los.

By visually checking the Fourier adjustments on the Pathfinder NDVI time-series, various special cases are identified, where the original Fourier adjustment algorithm does not hold. In the proceeding sections 3.4.1 - 3.4.4 these cases are extracted and modifications are applied to the original algorithm.

-
1. See section 2.5.2: Only $n/2^{m-1}$ outliers are corrected successfully. For a yearly time-series with 10-day intervals $n=36$ and $m=3$, hence interpolations for more than 9 outliers in a row are suspect to fitting errors.
 2. The Fourier adjustment technique presented in chapter 2 has been successfully applied to the GIMMS NDVI dataset (Los et al. 1994) by Sellers et al. (1996b). The Sellers' group has selected the GIMMS NDVI dataset, because it was the only one available at the early stages of the ISLSCP project. They subsampled global data derived from the AVHRR sensor to monthly composites with a spatial resolution of 1° by 1° , using the maximum-value composite method discussed earlier.

3.4.1 Adjustments to the weighting scheme

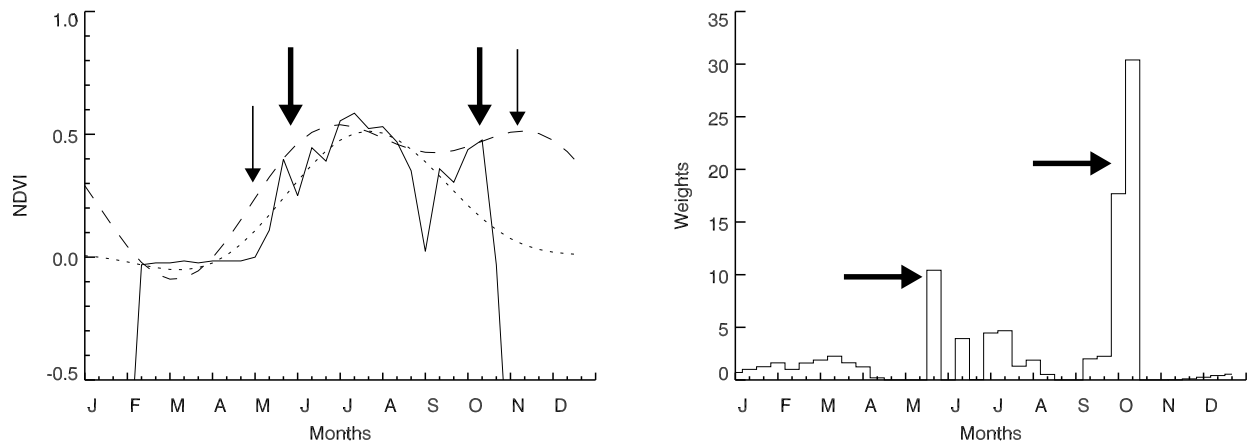


Figure 3.3: Extremely high weights (bold arrows, right) calculated from the first Fourier adjustment (left, bold arrows) lead to an overestimation of the weighted Fourier curve in neighbor locations (fine arrows, left). Solid line = Pathfinder NDVI data, dotted line = first Fourier adjustment, dashed line = second weighted Fourier adjustment. The time-series is from 1987, NE of Norway.

As seen in Figure 3.3, the first Fourier curve (dotted line) is fitted through a measured NDVI time-series (solid line), such that the squares of the errors are minimized. The data-points lying under the Fourier curve receive very low weights; the ones above the Fourier curve are weighted very high (Figure 3.3: bold arrows). The second Fourier fit (dashed line) - now taking in consideration the weights - is running through the high weighted points, but does not have included any more the values with lower weights. This weighted least squares fit, which was originally thought to effectively remove erroneous outliers in the GIMMS-NDVI dataset, does not hold here due to the following reasons:

- the 0.1° by 0.1° Pathfinder NDVI dataset is not spatially subsampled and therefore includes heavier noise. There is no smoothing effect as by taking the average of 100 neighbor values (1° by 1° subsampled NDVI data, Sellers et al. (1996b)).
- our time-series have more temporal variabilities, since the MVC-method is only applied to 10 days in comparison to one month as used in Sellers et al. (1996b).
- the above results in high slopes within the measured NDVI dataset. These slopes either occur due to measurement errors or they represent the vegetation growth period. If these slopes are close to high peaks, which receive high weights, the slopes will be ignored.

Several solutions can be considered to solve this problem:

- a) create monthly composites instead of 10-day composites
- b) decrease the spatial resolution
- c) increase the number of harmonics in the Fourier series
- d) adjust the weighting scheme of Sellers et al. (1996b)

a) and b) Reduce the temporal and spatial resolution

The first two solutions will not help to create a better vegetation dataset than it is available from various sources (e.g. ISLSCP CD-ROM). There is no need to reproduce those datasets for the European domain, since they can be retrieved with ease from global 1° by 1° gridded vegetation data. It was shown in section 2.6, that a higher temporal resolution increases the ability of the Fourier adjustment algorithm to actually represent the real vegetation behaviour and section 4.4 will show evidence for the need of higher spatial resolution when estimating land surface vegetation. Choosing 10-day composites with $0.1^\circ \times 0.1^\circ$ resolution in favor of a monthly ones with a $1^\circ \times 1^\circ$ latitude/longitude grid will help to significantly improve the quality of this dataset.

c) Increasing the number of harmonics

As discussed in section 2.5.2, the number of harmonics was set to two ($m=3$), since half-year features, like the vegetation seasonality, will be included in a periodic function with a one year period. At the same time, the curve is interpolated over time, such, that a sequence of $n/2^{m-1}$ outliers in a row will be corrected by the Fourier adjustment.

If the number of harmonics in the Fourier series is increased, the fitted Fourier curve is expected to have lower squared differences, hence the weights will be less extreme and the weighted fit should better represent the original data. The results with $m=5$ look very promising (Figure 3.4, right), since the weighted Fourier curve adjusts well to the measured dataset. The vegetation leaf-out in spring is not represented by this curve though, and neither is the typical seasonal variation. Two NDVI peaks are visible in Figure 3.4 (right), what does not reflect a natural phenomenon.¹

Reducing the number of harmonics to one ($m=2$) (see Figure 3.4 left), has the effect that the growth period is ignored. Such features are now considered as short term irregularities and will be smoothed out in this configuration.

Neither extending nor shrinking the length of the partial Fourier series provides a better adjustment of the measured NDVI time-series, what is in agreement with the experimental results of section 2.6. The number of harmonics is therefore used as described in Sellers et al. (1996b).

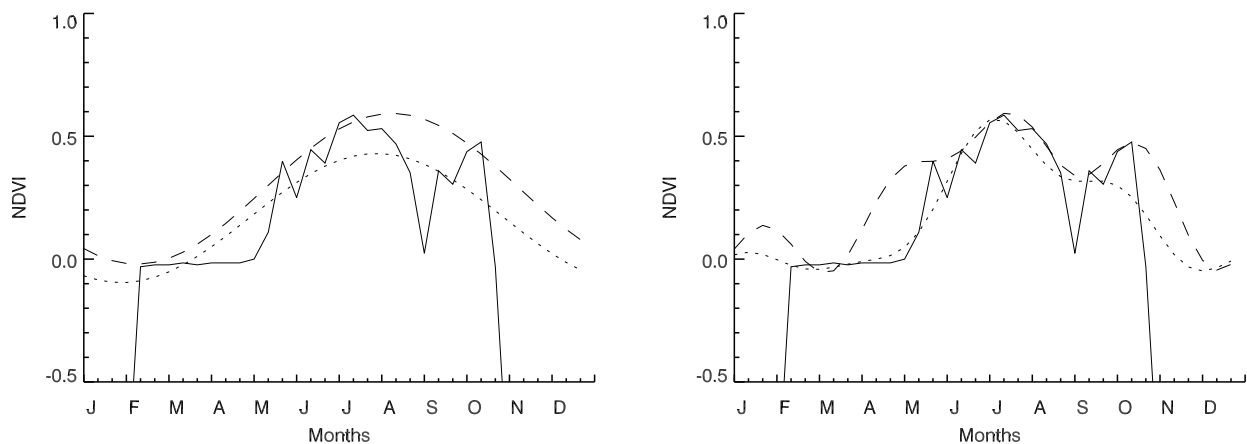


Figure 3.4: Number of harmonics considered in the Fourier series: $m=2$ (left), $m=5$ (right). Solid line = Pathfinder NDVI; dotted line = first Fourier adjustment; dashed line = weighted second Fourier adjustment

1. Vegetation types with more than one growing season exist (e.g. crops, see discussion in section 3.7.3). But this time-series does show a boreal vegetation in high latitude with a late growing season.

d) Adjust the weighting scheme:

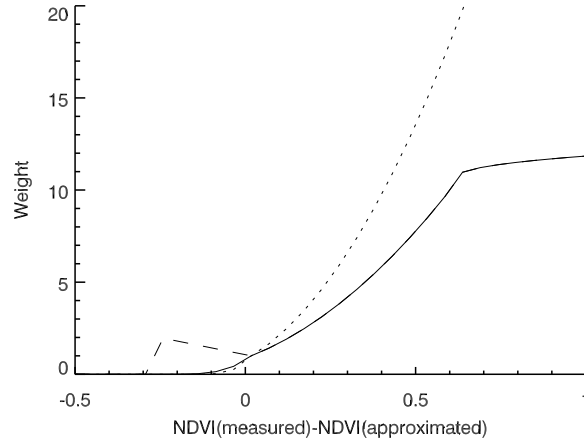


Figure 3.5: Weighting function. The curves represent the weighting applied to the measured time-series in relation to the fitting error of the first Fourier adjustment. dotted line = weighting by Los, solid line = weighting applied, when $NDVI > 0.2$, dashed line = weighting for $NDVI < 0.2$.

Applying weights to the measured data helps to distinguish credible points from erroneous data and is therefore central to the Fourier adjustment algorithm by Sellers et al. (1996b). This part was modified to enhance the quality of the Fourier adjustment. As seen before, high variability in the Pathfinder NDVI time series leads to very high weighted positive outliers and rejects all negative outliers after the first Fourier adjustment. This means, that low NDVI values occurring in spring (days before the growing period) are also treated as unreliable data and are ignored in the weighted Fourier adjustment, since their neighbor points have high NDVI values and therefore receive high weights. This problem was addressed by examining and adjusting the original weighting function (Figure 3.5, dotted line):

$$W_i = \begin{cases} 0 & U_i \leq -k \\ [1 + ((U_i + r)/k)]^4 & (-k < U_i < -r) \wedge NDVI \geq 0.2 \\ 1 - U/k & \text{if } (-k \leq U_i \leq -r) \wedge NDVI < 0.2 \\ 1 & -r \leq U_i \leq r \\ [1 + ((U_i - r)/k)]^2 & U_i > r \end{cases} \quad (3.1)$$

Where

$$k = 4 \text{ (originally 2)}$$

Low NDVI values, which are observed at the start and at the end of the growing season are now receiving higher weights, if their values do not differ too much from the first Fourier fit (U has to be no lower than $-k$). This weighting function, which also applies weights in a quadratic relationship to high and positive fitting errors, is extended with the following condition:

$$W_i(W_i > 10) = 4\sqrt{W_i - 10} + 10 \quad (3.2)$$

Very high weights, which previously resulted from single positive outliers in the time series, are suppressed by this condition and result in the new weighting function (Figure 3.5, solid line). Figure 3.6 shows the example from above, where vegetation growth starts only beginning of May, but with

a steep slope. It shows a better approximation by the modified weighting scheme, as compared to the original adjustment in Figure 3.3.

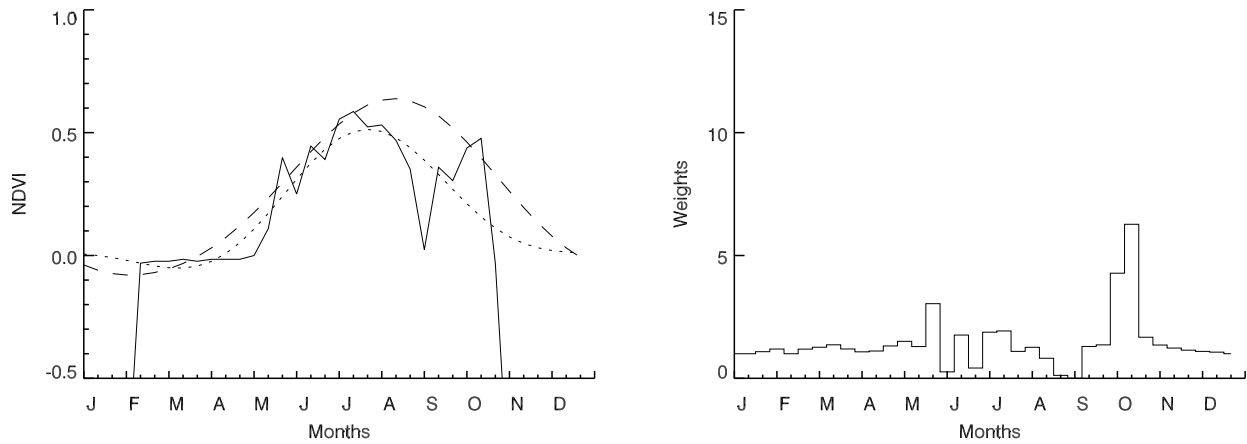


Figure 3.6: NDVI time-series with the modified weighting scheme applied here (right). Solid curve = Pathfinder NDVI; dotted curve = first Fourier adjustment; dashed curve = weighted second Fourier adjustment

3.4.2 Detection of positive outliers

In the years 1983 and 1987 several locations in eastern and northern Europe show unexpected high NDVI values between January and March. These values are identified as reliable data by the Fourier adjustment procedure. Considering vegetation cycles in high latitudes, the date when vegetation begins to grow is shifted in time with a rate of about one week per degree latitude (Moulin et al., 1997). Peaks in the vegetation cycle in latitudes from 40°N to 75°N usually occur end of July; no vegetation peaks are to be expected in the spring season.

The assumption has been made that all single peaks should be ignored if they happen in the first four months. These peaks are not identified by comparison to neighbor values as proposed by Moulin et al. (1997), because this tended to be unreliable when applied to the very noisy Pathfinder NDVI dataset. The following condition to the weighting function is applied instead, so that the early year peaks are identified by comparing them to the seasonal trend:

$$\text{IF } W_i \geq 3 \text{ for all } i < 15 \text{ then set } W_i = 0 \quad (3.3)$$

This condition has the effect that after the first Fourier adjustment, high positive outliers occurring before week 15 will be ignored by setting their weights to 0 for the second Fourier adjustment. Figure 3.7 shows the difference on a sample time-series from 1983 measured in eastern Europe.

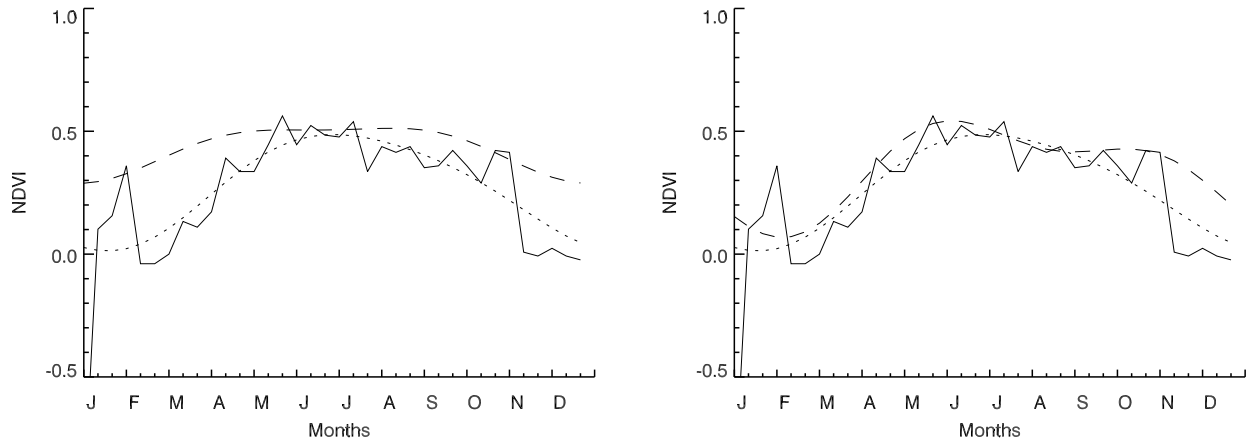


Figure 3.7: Identifying outliers, which happen in the time from January until April. Sample time-series from 1983 showing the effects of the early-peak detection procedure (right). Solid curve = Pathfinder NDVI; dotted curve = first Fourier adjustment; dashed curve = weighted second Fourier adjustment.

3.4.3 A third weighted Fourier fit

For short vegetation growth periods, as they are typical for high latitude deciduous vegetation types, another special case was identified: The vegetation growth period is represented well, but the Fourier curve drops to very low levels in winter times (see Figure 3.8, dotted line). This phenomenon can be explained as follows: the weighting function does not consider well the low NDVI values right before and right after the growing season, because they are usually lying at or below the first unweighted curve. If the low NDVI values outside the growing season are not taken into account as they should, then the Fourier adjustment will generate a curve, which continues its trend (raising or falling, according to growing season start and end).

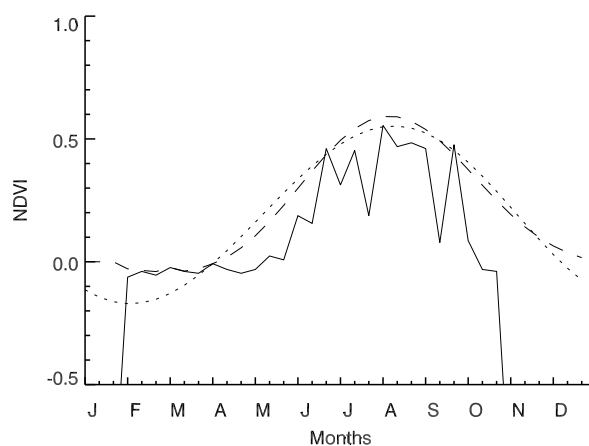


Figure 3.8: A third Fourier adjustment (dashed line) is applied by weighting the second curve (dotted line), 1984 time-series of NW Sweden.

This problem can be addressed as follows: by applying the weighting and fitting procedure again, now to the second Fourier curve. The solution is described here as “approaching” the best fit,

because previously ignored data will be weighted positive and perfect fitted values will keep their weights. This third Fourier adjustment increases the time needed by the routine by more than 50%, what is to be considered, if large multiyear datasets are processed. The quality enhancement of this third Fourier adjustment is higher for latitudes around 50°N-80°N, where large seasonality is part of the vegetation phenology (Moulin et al. 1997).

3.4.4 Repeated main period

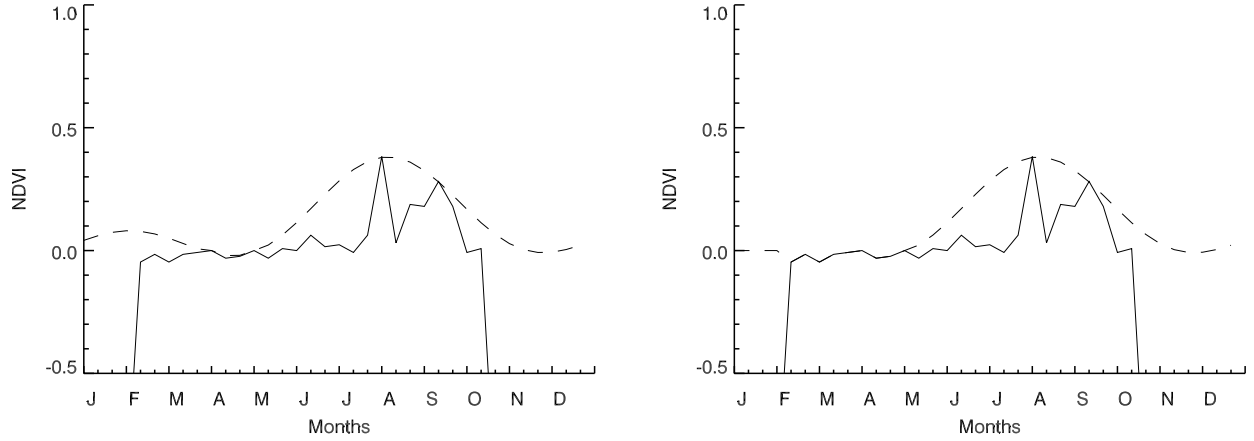


Figure 3.9: Second period of the main Fourier frequency appears at the beginning of the time-series (left). This has been corrected by rejecting Fourier adjusted values at the beginning of the year in time-series, where vegetation activity is happening late.

For vegetation cycles in high latitudes, the proposed Fourier adjustment technique has the following side-effects:

The vegetation-period P_v , where vegetation shows activity, is very short for high latitudes (less than half a year) and shifted into summer and autumn. The Fourier spectrum has its main frequency at $F_0 = (2\pi)/T \approx (2\pi)/(2P_v)$, which means that the length of the main $\sin((2\pi)/T)$ function is less than a year. A second period of this function appears at the beginning of the time-series (See Figure 3.9, left), what can only be suppressed by higher order Fourier series (Kreyszig 1988)¹. This problem has been spotted in a substantial part of the adjustments; therefore the Fourier adjustment algorithm was extended with the following condition:

$$\frac{\sum NDVI_{month > 6}}{\sum NDVI} > 0.9 \quad (3.4)$$

The condition 3.4 estimates, whether the vegetation period is short and only occurs in the second part of the year or if it is distributed equally throughout the year. The Fourier fit of the first three months is rejected and the original pathfinder data are used, if this condition is true (Figure 3.9, right). The threshold value was empirically set to 0.9 after testing the condition with the 1983 Pathfinder NDVI dataset at high latitudes.

1. A Fourier adjustment with higher order partial Fourier series will not hold for the assumption, that vegetation cycles vary smooth over time and have half-year cycles.

3.5 Interpolation of Missing Data

Remote sensing of NDVI strongly depends on the presence of sunlight, since absorption in the green biomass is measured at wavelengths in the visible spectrum. In the winter season at high latitudes, no sunlight is present. This results in missing NDVI data (See Figure 2.6, left). Most vegetation types (deciduous forests, grassland) will be in dormancy or not show activity under the snow cover during that period - this way missing data of these vegetation types will not affect the quality of the vegetation dataset. Evergreen needleleaf trees project through snow cover and influence the surface albedo as well as the roughness length (Betts and Ball 1997). Their presence therefore needs to be taken into account during the time, when no data can be obtained by satellite remote sensing. The applied Fourier adjustment only provides correction for three missing months in a row (see section 2.5.2) and larger data dropouts are set to missing again. Los (1998) is replacing the missing data with the NDVI value at the end of the growing season (usually October) for evergreen needleleaf vegetation types. A linear interpolation from the vegetation senescence to the growing season in the following year was used here instead, to provide steady transitions inbetween single years. An interpolated multi-year dataset can be seen in Figure 3.10.

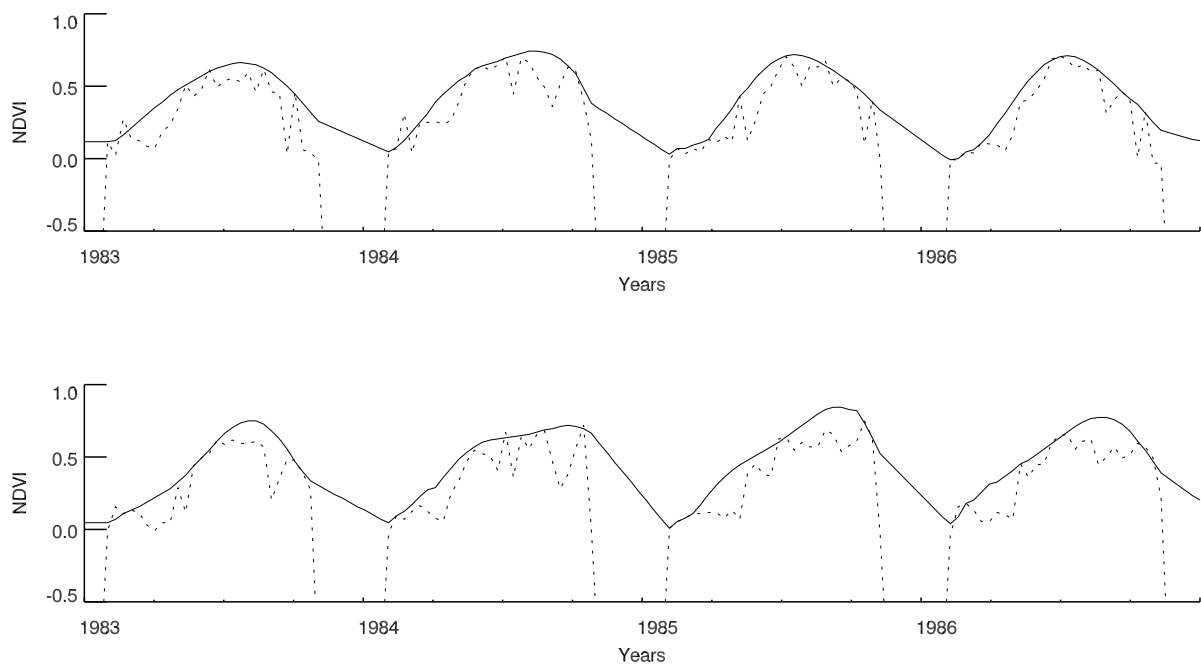
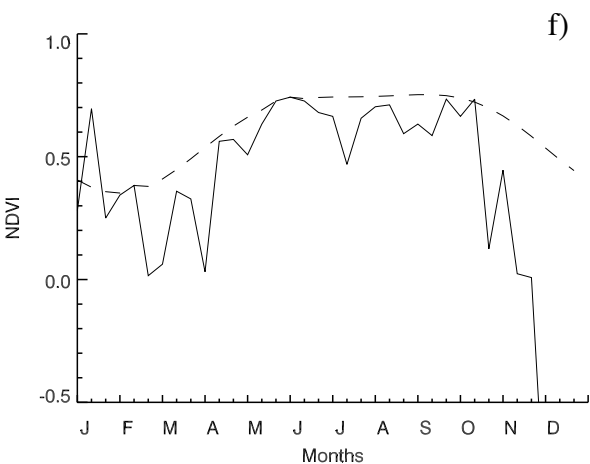
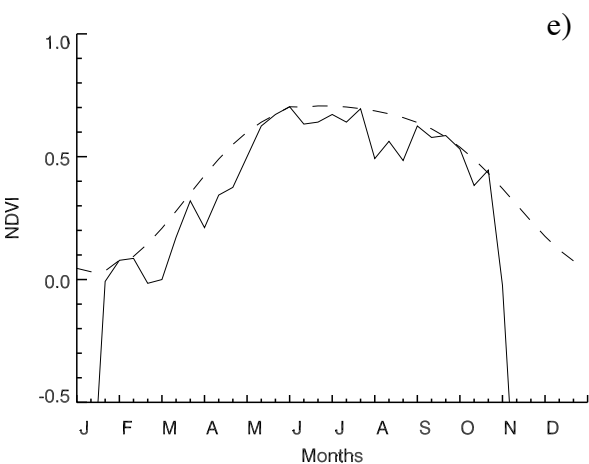
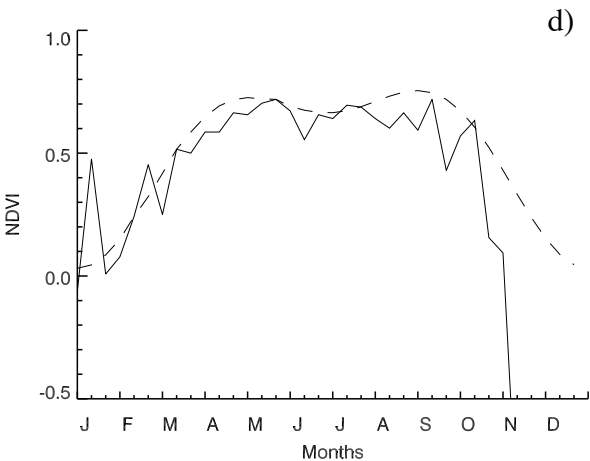
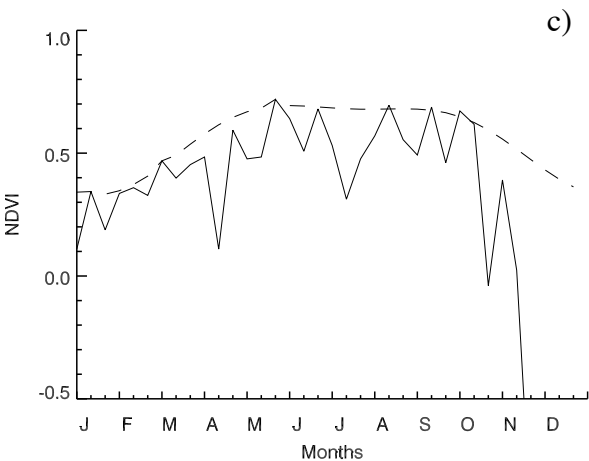
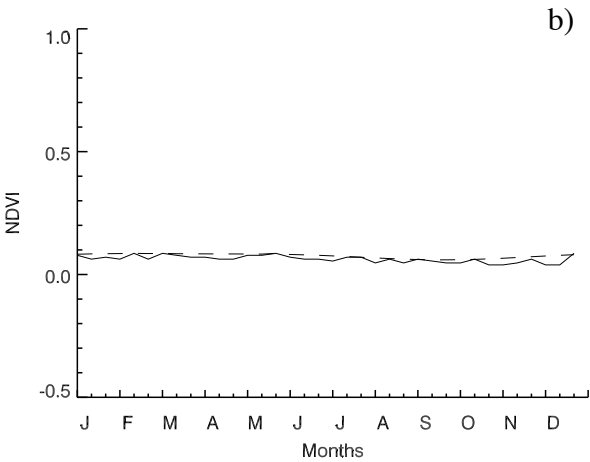
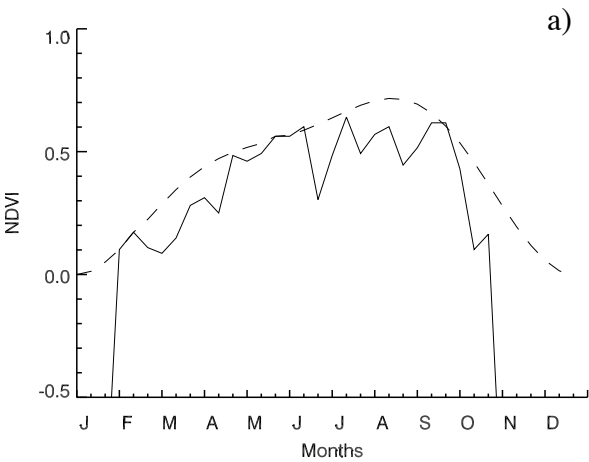


Figure 3.10: *Interpolations of missing data during winter time at high latitudes for evergreen forest in Finland (top) and Norway (bottom). Dashed line=Pathfinder time-series, solid line=Fourier-adjusted and interpolated time-series. Time-series represent one single pixel.*

3.6 Examples of Adjusted NDVI Time-Series



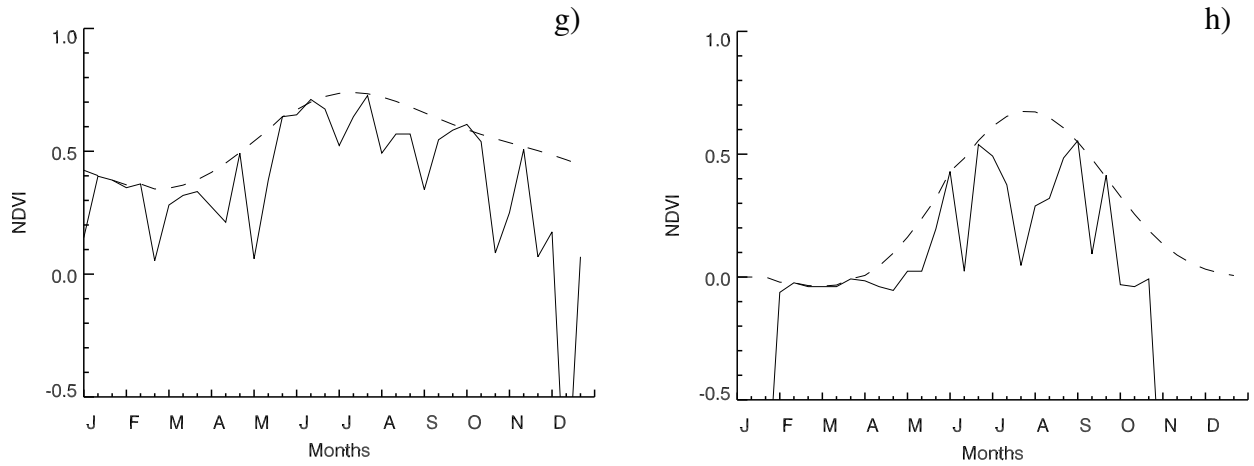


Figure 3.11: Example of time-series after the application of an enhanced Fourier adjustment to the Pathfinder NDVI presented in 3.4: a) Finland, 62°N 25°W, 1993, b) Algeria, 34°N 2°E, 1984, c) Germany, 50°N 7°E, 1993, d) Ireland, 53°N 7°W, 1984, e) Poland, 52°N 16°E, 1993, f) Switzerland, 47°N 8°E, 1993, g) Italy, 44°N 12°E, 1993, h) Iceland, 65°N 17°W, 1984. solid line = The original Pathfinder data, dashed line = adjusted time-series.

3.7 Remaining Anomalies

3.7.1 Positive outliers at the end of the vegetation period

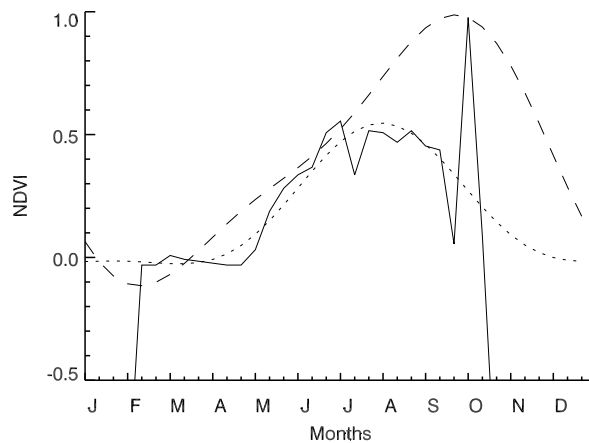


Figure 3.12: Very high NDVI peak at the end of the growing season. After the first Fourier adjustment (dotted line), the weighting scheme considers the high NDVI value as reliable and calculates a weighted Fourier curve (dashed line), that does not represent the vegetation cycle. Time-series from 1983, Finland.

The closer examination of Fourier adjusted time-series reveals another special case, which is showed in Figure 3.12. The very high NDVI value at the end of the growing season does not have any biological backgrounds, if the vegetation growth and senescence period are considered to happen smoothly over time (Moulin et al. 1997). Taking into account that maximum annual NDVI values above 0.7 are

not expected for any vegetation type (DeFries et al. 1998), this anomaly is explained as either a measurement or processing error in the Pathfinder dataset¹. The Fourier adjustment is weighting this anomaly as a reliable NDVI value and calculates a very unusable curve (dashed line). However, this feature has only been spotted in less than 1% of all examined time-series. A correction scheme is not yet included into the Fourier adjustment procedure. A solution to this special case is in preparation and will be presented by the author when available.

3.7.2 Overestimation in boreal forest habitats

Boreal evergreen forests show a steep increase in NDVI values, which corresponds to snowmelt in those regions (Moulin et al. 1997). The sudden variations are detected correctly by the Fourier adjustment (Figure 3.13). Due to the square wave-type shape of this vegetation cycle, the partial Fourier series with $m=3$ will not be able to analytically represent the flat parts of the NDVI time-series. The snow-free NDVI-values from evergreen trees in Boreal regions are overestimated. In Figure 3.13 this effect rises the estimated NDVI value up to 0.2 NDVI above the measured value. Los (1998) is rejecting those overestimates in all time-series - independent of the vegetation type - in favor of the original data. This rejection-procedure has been evaluated here as well, but was not applied, since retaining the original data will lead to underestimates and introduce high frequency components again.

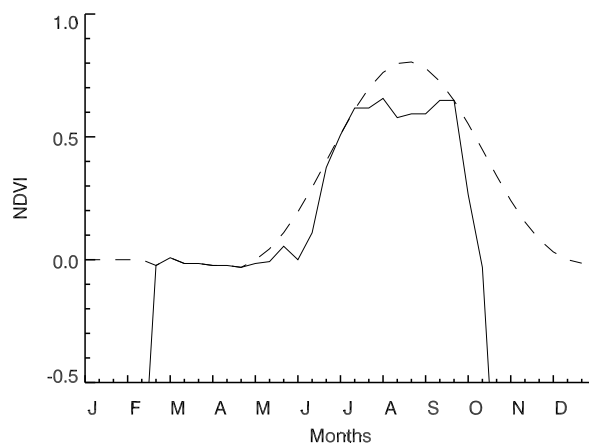


Figure 3.13: Boreal evergreen forests (here: 1993 time-series from the northern part of Sweden) have a square-type shape in the yearly vegetation cycle. The Fourier adjustment can't analytically represent this shape. Note the rejection of adjusted values in the first months of the year.

-
1. The NOAA Pathfinder team has addressed this problem in their documentation found on the web (NOAA/ NASA 1999). The observed anomalous high NDVI values usually occur during times of low vegetation activity and are visible linear features in the dataset. They are explained as transmission errors due to low antenna angles near the horizon, usually occurring in polar regions where only few ground receiving stations are located. Whereas other data dropouts are suppressed by the maximum-value compositing method, these errors are carried throughout the whole processing steps and still appear in the final Land Pathfinder product.

3.7.3 Agriculture

The basic assumption A-1 for the applied Fourier adjustment leads to an optimized detection of seasonal occurring features. Agricultural areas have distinct growth patterns, which usually differ from the ones observed in natural vegetation. Justice et al. (1985) has monitored Chinese crop sites and detected various crop-specific patterns in the temporal development of NDVI at those sites. Rice shows distinct bimodal NDVI time-series in contrast to the seasonal feature observed for other vegetation types. Spring wheat, winter wheat and mixed crop areas will not show the expected patterns for a correct Fourier adjustment. Although most of the crop areas will be correctly represented by the current algorithm, a specific correction scheme has to be evaluated to correctly represent all crop areas, since they are present in about 30% of the European land surface.

3.7.4 Growth-season start and end

A higher temporal resolution was selected for this dataset, because the actual vegetation signal can be extracted more accurately from an error-contaminated time-series if more values are present. This is performed very well for most cases (Figure 3.11 a-h).

However, it needs to be emphasized here that cases with a high seasonality like shown in Figure 3.13 overestimate the vegetation activity when sudden changes at the start and the end of the growth season occur. There are many cases where the discrimination of the correct slope in spring or autumn will not even be possible by visually inspecting the time-series. This problem certainly needs further research. It is expected, that these errors have minor impact, since low ambient temperatures only lead to small photosynthesis assimilation rates in the coupled biosphere - atmosphere model (Los 1998).

4. Results: The 1983-1993 EFAI-NDVI

The Fourier adjustment algorithm presented in section 2.5 and applied to the Pathfinder dataset in sections 3.3 to 3.5 was tested by visually inspecting yearly time-series. This procedure has led to various corrections in the adjustment technique before the final processing of the full dataset from 1983-1993 was performed.

The resulting European Fourier-Adjusted and Interpolated NDVI (EFAI-NDVI) dataset is presented here and undergoes several statistical examinations. The assumption A-2 states that the correction scheme for the Land Pathfinder NDVI is most significant where cloud cover is supposed to occur. This applies to dense green vegetated areas, where impacts of cloud cover and atmospheric water vapor are expected to be strongest. Mid latitude forest vegetation undergoes a strong correction, whereas semi-arid areas and deserts are mainly invariant to the Fourier adjustment. Vegetation varies seasonally, dependent on the type of plant and the environmental conditions. Vegetation class specific seasonal trends are represented well by the EFAI-NDVI. Findings by Myneni et al. (1997) have indicated an increase of vegetation activity in northern latitudes through NDVI time-series for the period from 1981-1991. A positive inter-annual trend of $1.4\% \text{ yr}^{-1}$ is detected here from the multi-year NDVI signal in the decade ranging from 1983 to 1993; a result that needs further verification and should be trusted with care, since multi-year satellite remote sensing datasets can be subject to systematic errors and drifts (Hurrell and Trenberth 1997).

Spatial variations within the examined dataset lead to various findings: Some areas do not show substantial spatial variability up to $1^\circ \times 1^\circ$ grid resolution, where other examined sites strongly justify the use of a higher spatial resolution, because there is significant spatial variability in the land surface vegetation within small areas detectable.

The resulting dataset is also compared to the FASIR-NDVI by Los (1998), what indicates a general consensus in the inter-annual differences. The two datasets show different amplitudes, especially at the end of each growing season. Variations between the two datasets are likely to be caused by different correction algorithms involved or by the maximum-value composite sampling of the FASIR-NDVI.

4.1 Latitudinal Variation

The effect of the Fourier adjustment algorithm varies by latitude (Figure 4.1). The applied corrections are most significant for latitudes between 40° and 55° northern latitude, whereas subtropical regions around 30°N have characteristic low NDVI values and are not subject to heavy cloud cover. This minimizes the effect of the applied correction for low latitudes in the used domain. The Fourier adjustment has the greatest effect where NDVI values are supposed to be most suspect due to cloud interference in mid latitudes. The interpolation of missing values in northern parts of Europe for evergreen needleleaf forest will only show a minor effect here, since all missing NDVI values - also the non interpolated ones - are set to 0 prior to spatial averaging the dataset. The interpolation of data most certainly influences the estimation of biophysical land surface parameters and associated biosphere-atmosphere processes in future applications.

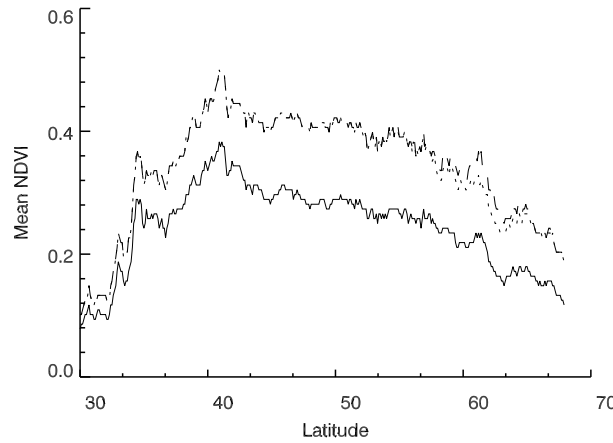


Figure 4.1: Latitude dependent mean NDVI plot: The Fourier adjustment (dashed line) rises the mean NDVI values in mid latitudes to almost 50% over the Pathfinder NDVI values (solid line). The effect of interpolating evergreen needleleaf vegetation values during winter in high latitudes can be clearly identified around 60 degrees (dotted line), where most of this vegetation type occurs. The latitudinal NDVI values were averaged over longitude and time for the year 1987.

4.2 Seasonal Variation

It is assumed that the Fourier adjustment technique is also showing higher effectiveness when a large fraction of dense green vegetation cover is present, since atmospheric disturbance will have a higher effect on the absolute NDVI values. To illustrate this relationship, the European NDVI was averaged by land cover class¹ as shown in Figure 4.2.

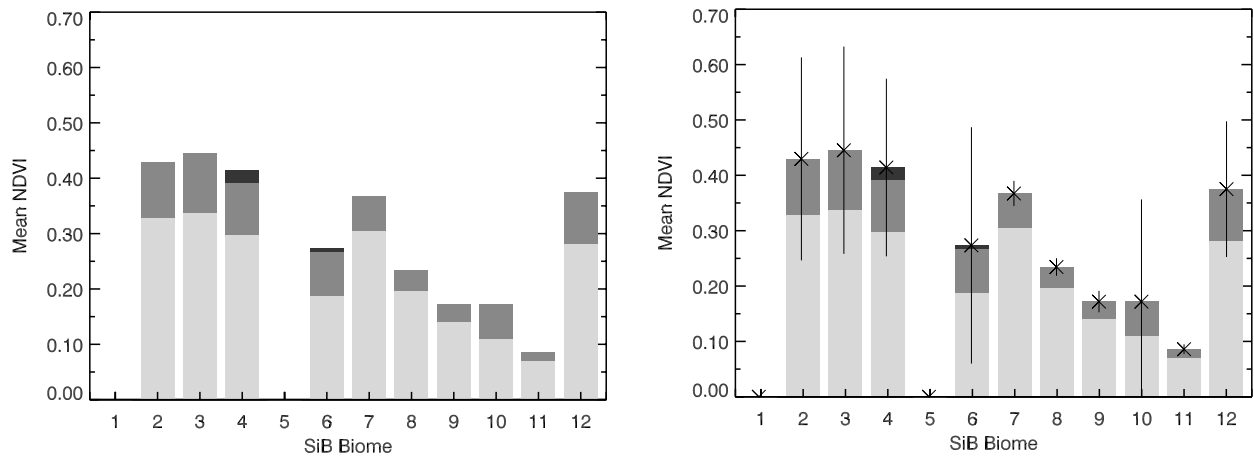


Figure 4.2: NDVI values averaged by SiB biome for the 1992 dataset(classification scheme see section 5.2). [light grey] shows the biome averages of the original Pathfinder dataset, [dark grey] is the Fourier adjusted dataset and [Black] demonstrates the effect of interpolating evergreen needleleaf vegetation values (SiB class 4). Temporal NDVI anomalies are shown by the plotted standard deviation for each SiB vegetation class. Missing values of the other datasets were set to NDVI=0 prior to averaging.

1. A complete description of the used land cover classification can be found in section 5.2

Mid-latitude forest biomes 2,3 and 4 are undergoing a heavy adjustment, whereas shrub cover, bare soil and desert classes 8,9 and 11 are only subject to marginal corrections. Boreal evergreen needleleaf forests (Class 4) are interpolated for missing values during winter time by the Fourier adjustment procedure presented in chapter 3. This effect is clearly visible in Figure 4.2 (black), but does only have minor impact to the overall correction.

Vegetation types can be clearly distinguished by their difference in seasonal variability. This approach was used to generate land cover maps from remotely-sensed vegetation in recent publications (Moulin et al. 1997 and DeFries et al. 1998). Figure 4.2 (right) and Figure 4.3 both plot the temporal standard deviation and the yearly mean NDVI for each of the 12 land cover classes¹. Deserts (class 9) and tundra (class 10) both show low mean NDVI values, where deserts have a characteristic low NDVI throughout the year and tundra exhibits a high inter-annual variation. High mean NDVI values and large inter-annual ranges are found in mid latitude forests (classes 2 and 3). Grassland (class 7) has a high mean NDVI and does not show much seasonal variation. These findings are in close agreement with the results from the global ISLSCP-II initiative presented by Los et al. (2000). Differences are observed for grassland, shrubland and other semi-arid land cover classes. These differences can be explained due to the very coarse grid resolution of the ISLSCP-II data, what will support the theory by Fung et al. (1997), that sub-grid vegetation-components² are essential for the correct estimation of the vegetation properties - which certainly needs further examination. However it is assumed here that these differences in certain biomes result because the ISLSCP-II data was globally sampled, where the present data only includes vegetation data from Europe. Semi-arid vegetation is supposed to differ much across different continents.

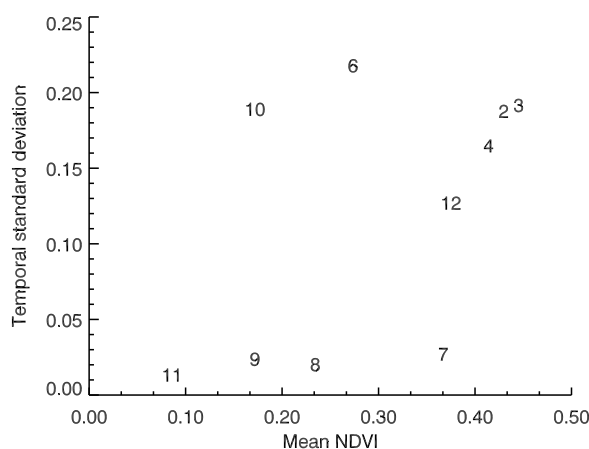
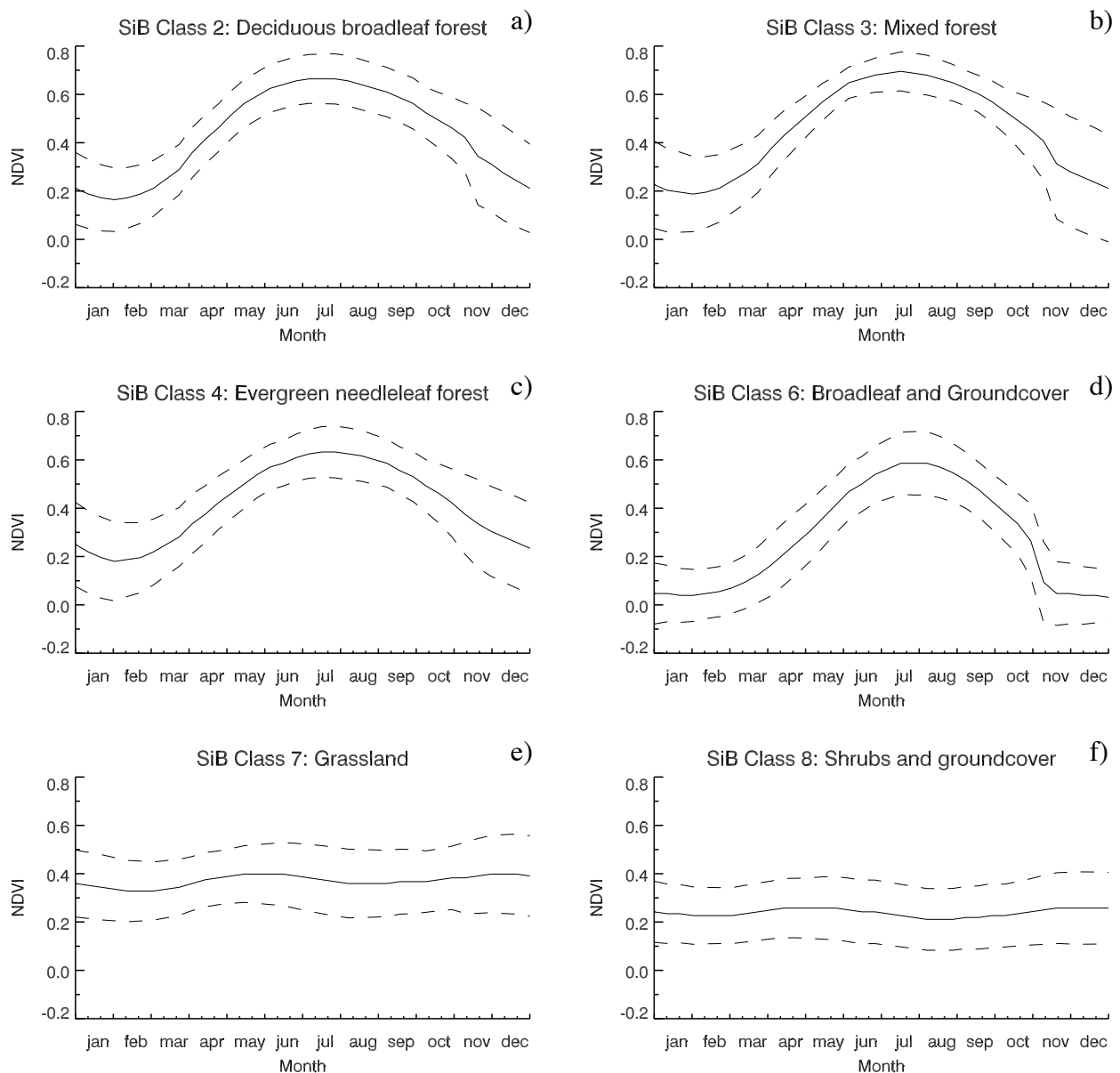


Figure 4.3: Mean annual NDVI, averaged per SiB land cover type is plotted against its temporal variability. Desert, Shrubs and bare soil (Biomes 8, 9, 11) show low vegetation activity and also have low annual variabilities, grassland (Biome 7) has a high mean NDVI and low temporal deviation, whereas forest biomes (2 ,3 ,4) show high NDVI values and also a significant seasonal variability.

Sufficient spatial and temporal accuracy of the NDVI dataset and the actual representation of the seasonal vegetation variability is a main requirement to quantitatively derive the associated land sur-

-
1. The SiB classes 1 (tropical forest) and 5 (evergreen broadleaf forest) are not present in the examined domain.
 2. Certain SVATS work with sub-grid vegetation classes to better estimate fluxes in a heterogeneous, but coarse-scale gridded land surface: a mosaic pattern with various sub-grid classes is used in the RAMS SVATS for instance.

face parameters for use in the Simple Biosphere model. In Figure 4.4 the temporal development of NDVI for the 12 SiB land cover classes is shown. The values are spatially averaged by land cover class and plotted over time for the 1992 EFAI-NDVI product. The mean values are shown together with their \pm spatial standard deviation. Except for the desert class 11, all classes show a substantial standard deviation, what is obvious, because vegetation classes do not reflect a homogeneous vegetation type, but rather include a variety of plant types with similar properties. They also occur in different eco-zones (eg. different latitude or height), such that plant phenological characteristics are variable for the same vegetation type. This finding is important and certainly will be addressed in the future research of LSP's, because land surface properties are NDVI- as well as land cover type dependent.



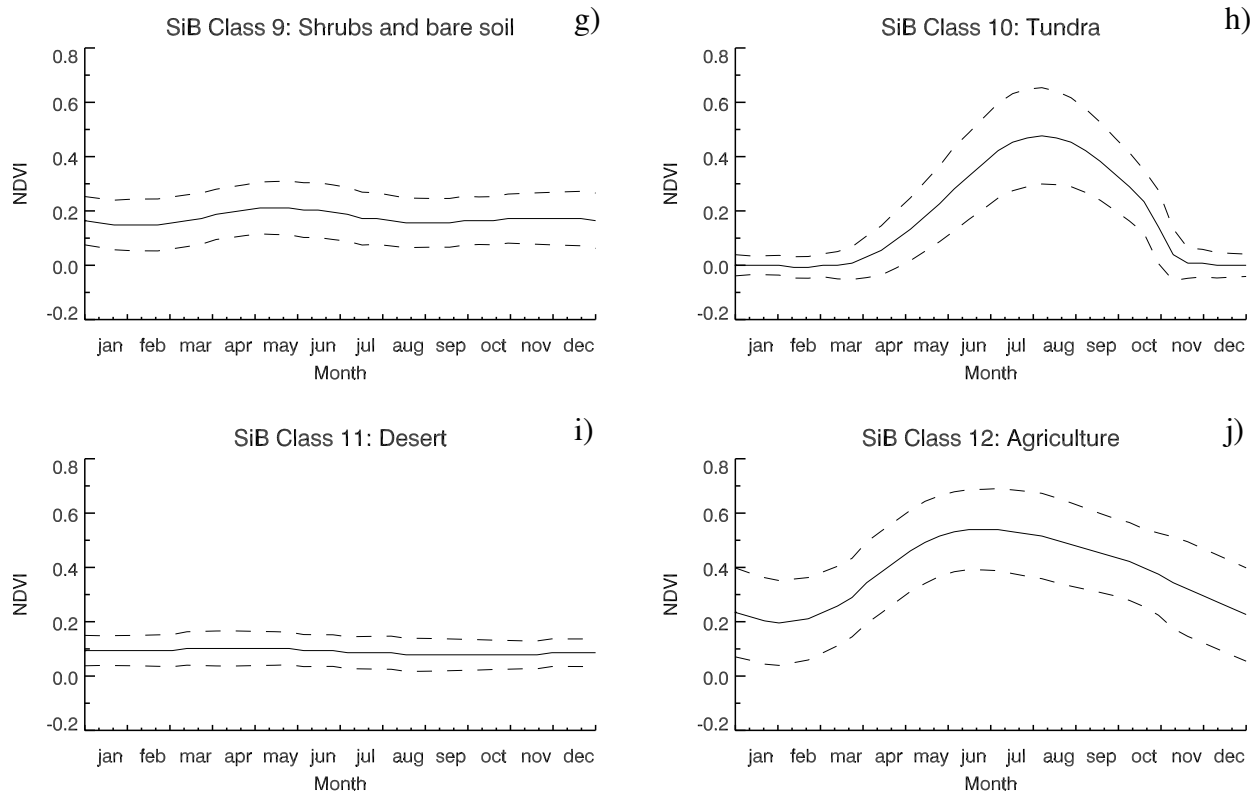


Figure 4.4: Seasonal characteristics from NDVI time-series for the 12 SiB land cover classes (a-j). Classes 1 (tropical rainforest) and 5 (Deciduous needleleaf forest) are not present in the regional European domain. Solid line= averaged NDVI, dashed line = average \pm standard deviation. Missing values were set to NDVI=0 prior to averaging.

4.3 Inter-Annual Variation

The naturally changing environmental conditions affect the timing of the vegetation phenology by definition (Farquhar 1980, Nultsch 1996, Collatz et al. 1998). In addition to the seasonal trends presented in section 4.2, inter-annual signals can be extracted from NDVI datasets. There exist natural phenomena on time scales between 1-4 years like the El Niño/Southern Oscillation (ENSO) or volcanic eruptions on one side and long lasting trends like the anthropogenic carbon dioxide increase in the atmosphere on the other side. Land surface vegetation can serve as an indicator for climate change, because it is influenced through environmental forcings that are exerted by these phenomena. Land surface vegetation also provides several feedback mechanisms with its radiative and physiological characteristics. Bounoua et al. (1999) demonstrated this relationship by a coupled biosphere-atmosphere model with a doubled CO₂ scenario.

Myneni et al. (1997) has found strong evidence from remotely-sensed NDVI datasets, that photosynthetic activity of terrestrial vegetation increased and that the vegetation growing season lengthened by several days in the decade from 1981 to 1991. The regions exhibiting the greatest increase lie between 45°N and 70°N, where the most warming has been observed due to early disappearance of snow. The findings indicate a positive correlation between the increase in amplitude of the seasonal cycle of atmospheric CO₂ since the early 1970s and the terrestrial net primary productivity.

The ENSO phenomenon has a typical period of 2 to 7 years and results in a displacement of rainfall patterns in the global tropics and anomalous warmings in the central and eastern pacific ocean waters. Anyamba and Eastman (1996) have shown that inter-annual variabilities in the South Africa AVHRR NDVI values show a relationship to the typical ENSO variables measured in the pacific. Their findings for the observed years 1986 to 1990 illustrate, that remotely-sensed vegetation products can serve as a good indicator of inter-annual climate variabilities.

As a conclusion, it is possible to detect inter-annual trends in global or regional NDVI patterns, but it needs to be considered that remotely-sensed NDVI is only available for the last 10-15 years, what poses a strong limitation to the analysis of time-series longer than a decade for climatic use. The magnitude of these trends are in between 1-3% (Myneni et al. 1997), what is similar to the estimated margin of errors (see section 4.6). Interannual signals have a stronger amplitude when only 2-4 years are compared. The magnitude of inter-annual variations is about 10-15% of the seasonal amplitude (Los et al. 2000), so that more confidence is put in inter-annual variations than in inter-annual trends.

Although error sources induced by the type of measurement are large as described above, a trend analysis is performed here for the EFAI-NDVI dataset through the whole time-series. It shows a general positive trend of 1.4% per year by using a linear estimate for the examined time period from 1983 to 1993 (Figure 4.5). These findings are consistent with the results published by Myneni et al. (1997), who proposed an increase of 7-10% in a 9 year Pathfinder time-series and 8-14% in 8 a year GIMMS time-series.

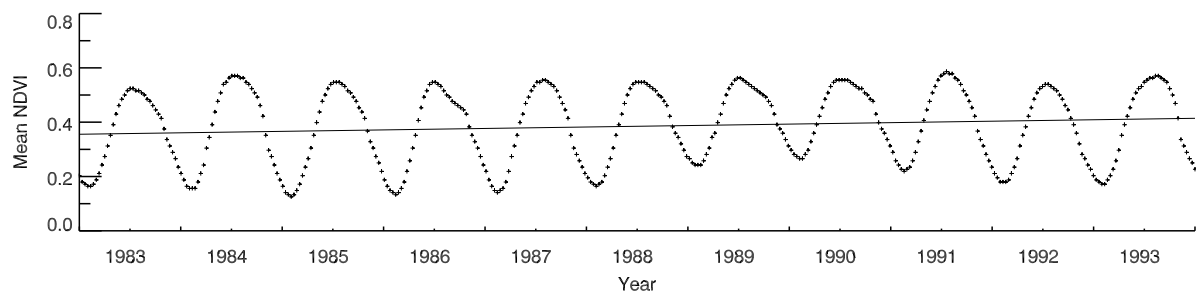


Figure 4.5: Spatially averaged EFAI-NDVI in the period from January 1983 until December 1993 (+). The linear regression (solid line) indicates an increase in vegetation activity of 1.39% per year.

It is evident, that not all vegetation types will show the same response to altered CO₂ conditions in the atmosphere, what is illustrated in Table 4.1. Estimates for the response of net primary production to increased CO₂ are indicated for several of the vegetation classes as presented by Los et al. (2000) for global extents.

However, these results have to be validated by longer time-series and with better corrections to the error-sources in the NDVI dataset. Myneni et al. (1997) shows, that year-to-year variabilities like the unusual global warming of 1990 can strongly influence a decadal trend analysis. There is more credibility in the analysis of vegetation behaviour at specific years.

Table 4.1: A linear trend analysis of the EFAI-NDVI by vegetation type. NPP trend estimates from the global FASIR-NDVI through a biochemical model presented in Los et al. (2000)

	SiB Biomes	EFAI NDVI trend [yr-1]	NPP trend [yr-1]
-	Global	1.39%	1.3%
2	Deciduous broadleaf	1.49%	-
3	Mixed broadleaf & neeleleaf	1.21%	1.5%
4	Evergreen needleleaf	1.50%	2.5%
6	Broadleaf and groundcover	1.39%	1.2%
7	Grassland	0.37%	-
8	Shrubs and groundcover	1.33%	-
9	Shrubs and bare soil	1.27%	-
10	Tundra	1.73%	2.5%
11	Desert	1.45%	-
12	Agriculture	1.42%	1.2%

Inter-annual differences are best analyzed by comparing single years from the multi-year dataset. In Figure 4.6 differences are seen in the start and end of the growing season as well as the amplitude variability between the summer and the winter NDVI over Europe. The year 1990 had a very long growing season in comparison to the year 1992. These results need further validation with measured data from climatic time-series.

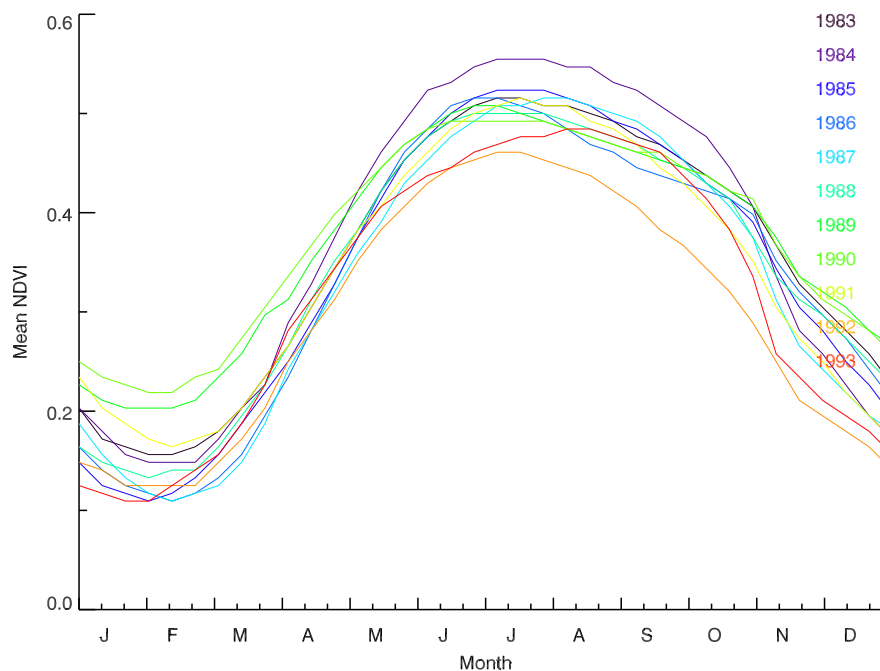


Figure 4.6: An overlay of the mean NDVI yearly time-series for the years 1983-1993 shows differences in the seasonal cycles. The curves are de-trended for the trend found in table 4.1.

4.4 Spatial Variation

The Fourier adjustment discussed in section 2.5 and applied to the NOAA Pathfinder NDVI in section 3.4 is a temporal correction of single pixels and does not account for spatial correlations between the pixels. Eklundh (1995) estimates uncorrelated noise to spatially varying scene information in AVHRR NDVI by measuring inter-pixel variability with a semivariance function. The findings indicate that spatial differences at the instrument pixel level is substantially influenced with noise, such that the usefulness of pixel level (1.1 km) comparisons is to be questioned. Semi arid areas showed lower inter-pixel noise than wetter areas, that are heavily contaminated with cloud cover, what is in agreement with the results in sections 4.1 and 4.2. Inter-pixel noise was furthermore found to be lower when datasets were sampled over time, which strengthens the assumption that the temporal Fourier adjustment applied here is also effective in removing spatial noise from registration errors, instrument malfunction and cloud cover.

Where spatial subsampling with the maximum-value composite method will effectively remove cloud contamination and atmospheric interferences, spatial features at certain scales are likely to be ignored by this process. It needs to be verified if the higher 0.1° resolution dataset allows to better represent biological features on the land surface than 1° datasets as the FASIR-NDVI. Townshend and Justice (1990) analyzed the scene variance dependent on pixel resolution size with Landsat Multispectral Scanner (MSS) data ranging 0.1 to 100 km cell size. They found considerable differences in the scene variance depending on the ecological zone covered by the imagery. It was shown, that although most of the changes in the images occurred at fine scales (0.1-1km), a substantial portion of the spatial variability was detectable at coarser resolutions. They concluded that it is very difficult to select a suitable spatial resolution for all areas.

In comparison to this study, Belward (1992) applied spatial analysis to AVHRR datasets with resolutions from 1 to 24 km and tried to estimate the mean object sizes found in a particular scene. The findings are in close agreement with the results from the study by Townshend and Justice (1990). Some areas inherit increasing spatial NDVI variability with increasing cell size, but areas with high spatial variability at fine resolutions exist as well, dependent on the examined site. Areas were found, where cell sizes of 24 km¹ still were a good approach to represent the main scene features.

The above findings were all performed with either maximum-value composited or raw datasets with no corrections applied. It is assumed by the author that much inter-pixel variance occurs due to temporal inconsistencies in such datasets.

A spatial analysis as described in Richards and Jia (1999) is performed for the 0.1°x0.1° European Fourier-Adjusted and Interpolated NDVI (EFAI-NDVI) dataset of July 11-21, 1987. NDVI distributions at different cell sizes are examined for their spatial standard deviation at different locations within the European domain. The results of this analysis are shown in Figure 4.7. For the interpretation of the plots it is important to remember that when sampled cell size is smaller than the mean size of the scene's dominant objects, then neighboring pixels are highly correlated and inter-pixel standard deviation is low. Where the cell resolution approaches the main feature size, neighboring pixels are less similar and standard deviation is supposed to rise or show a peak. Where the resolution is larger than the main land surface features, standard deviation will again be low. Thus, the spatial standard deviation will show a maximum-value when the size of the sampled area approaches the size of the main features present in the remotely-sensed area.

1. Approximately 0.2° in a latitude/longitude grid at medium latitudes

Results show considerable differences between the observed areas: for mixed forest and grassland vegetation assumed in a), f) and h), the main land surface features seems to occur within larger extents, since spatial variation continuously rises, what is a strong indicator that land surface diversity can be described with a coarse resolution as used by Los et al. (2000) for the ISLSCP-II initiative. In large desert and bare soil areas, spatial standard deviation does not show any significant values, even at 15x15 pixel samples - what corresponds to 135x135 km on the land surface. In contrast, graphs c), f) and g) are strong indicators for scene object sizes below the $1^\circ \times 1^\circ$ grid size, that are only detectable with dataset resolutions like the one used in this work. Plot i) shows randomly chosen test sites over the whole dataset. It is evident that significant spatial NDVI variability occurs below the $1^\circ \times 1^\circ$ and even below the $0.5^\circ \times 0.5^\circ$ resolution for most areas within the examined domain - dependent on the vegetation type present. This finding will need further examination, since the correct spatial and temporal estimation of NDVI is central for the estimation of the land surface energy, mass and momentum fluxes in biosphere models.

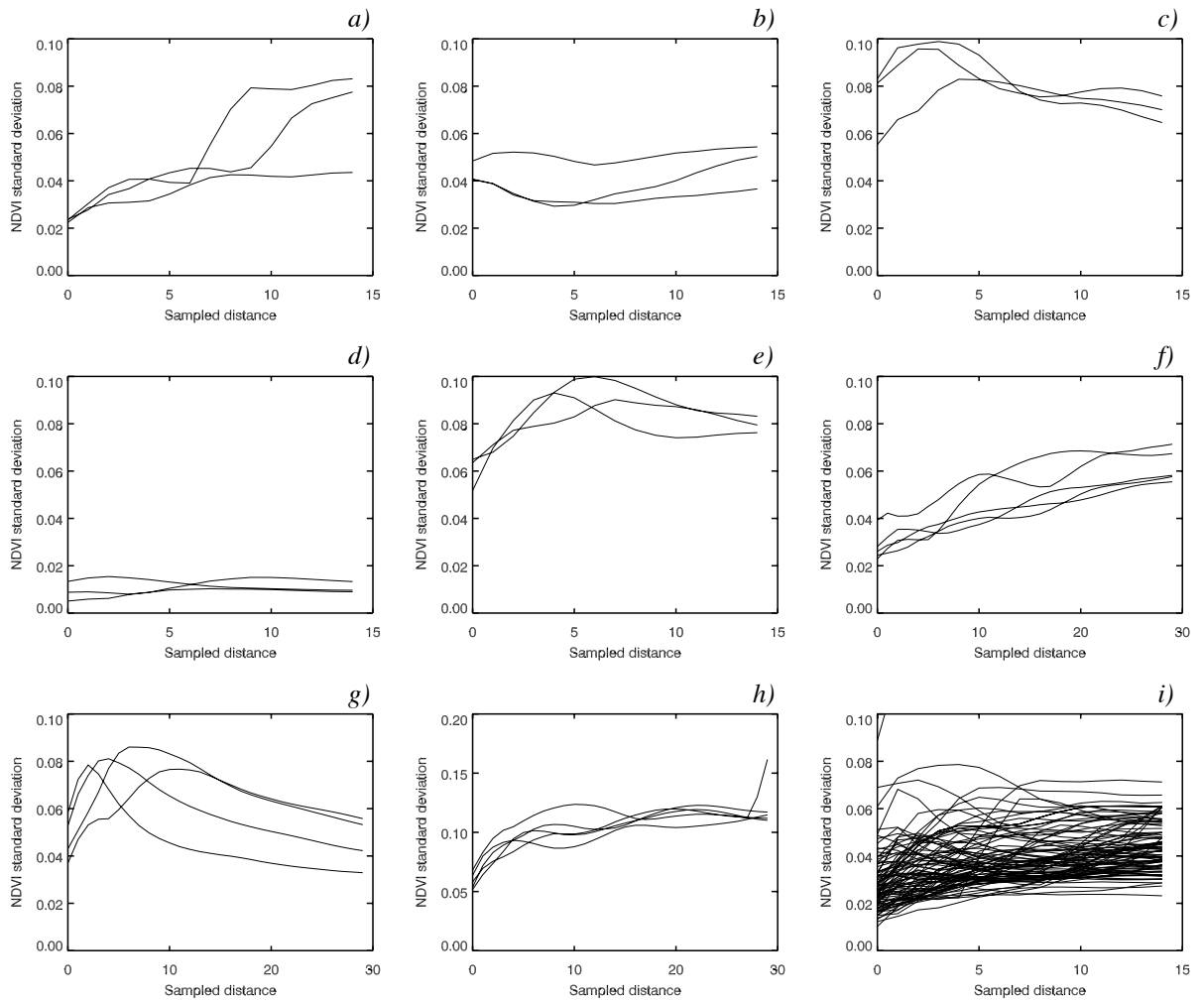


Figure 4.7: NDVI spatial standard deviation in relation to the sampled cell size. The x-axis shows the cell size as the amount of neighbor values included in the sample. High spatial deviations indicate, that dominant features are occurring at that particular cell size. The analysis was performed at different locations within the European domain for the July 11-21, 1987 EFAI-NDVI dataset: a) Germany 50°N 10°E , b) Hungary 47°N 17°E , c) Sweden 58°N 15°E , d) Algeria 35°N 3°E , e) Switzerland 47°N 8°E , f) Poland 52°N 17°E , g) Sweden 58°N 15°E , h) Germany 51°N 10°E , i) Central Europe (multi-plot).

4.5 Comparison to the FASIR NDVI

Among other available NDVI datasets, the ISLSCP-II FASIR-NDVI by Los et al. (2000) is the one most similar to the generated EFAI-NDVI, since it also bases on Fourier adjustment of NOAA NDVI time-series. The FASIR-NDVI is available in a pre-release¹ for monthly composites and the time-period from 1982-1990 with a 1° by 1° global extent. A comparison between the two datasets (where only the European domain was used for Los' dataset) indicates various differences (Figure 4.8). Since the FASIR-NDVI is only a monthly dataset, it is assumed here that the monthly values were set to the 20th day of each month. This is not evident, since the maximum-value composite method does not provide any information about the temporal occurrence of the sampled values.²

The EFAI-NDVI is exerting larger amplitudes in the seasonal vegetation cycle than the FASIR-NDVI dataset. It also behaves differently at the end of each growing season - it takes longer until vegetation activity ceases in autumn. Although, some of the inter-annual variabilities like lower amplitudes in 1989 are visible in both time-series.

The amplitude differences can result from the different processing steps involved in the two datasets. The FASIR - NDVI was undergoing different corrections for sensor degradation, aerosols and has a compensation for viewing angle and solar zenith angle effects applied, that differs from the one used by the NOAA Pathfinder team.

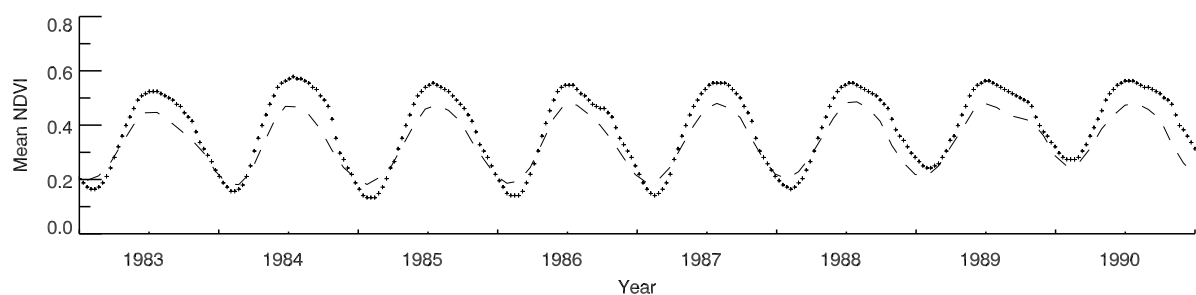


Figure 4.8: Spatially averaged time-series from 1983-1990 for the European domain. The (+) represents the EFAI-NDVI, where (-) is the FASIR-NDVI. Interannual variations show agreement, but differences exist in the seasonal amplitude. The end of the vegetation season is subject large differences between the two datasets.

The Fourier adjustment described by Los (1998) rejects Fourier adjusted values when they are 2% above the maximum of four nearest neighbor values. This procedure leads to the paradox situation, that faulty measured values are retained in favor to the adjusted Fourier curve. Since outliers in NDVI time-series are mostly low values, a smaller amplitude is expected as a result if the NDVI data is spatially averaged like shown in Figure 4.8. Another difference between the processing steps is, that the FASIR-NDVI was spatially and temporally subsampled by the maximum-value composite method discussed earlier. This method favors high NDVI values over lower ones, what has its justification, when assumption A-2 is considered to be reliable. In sections 3.4.2 and 3.7.1 it was shown, that this is not always the case and positive outliers can happen, especially during times of low vege-

1. The ISLSCP-II FASIR-NDVI dataset as well as other satellite derived global land surface parameters are available at: [ftp://islscp2.gsfc.nasa.gov/islscp2_data/] for public use. The final versions of the datasets will be published by 2002.
2. An earlier approach, where both datasets had time-series starting at the beginning of the first month, showed a substantial phase-shift of half a month.

tation activity. It is therefore expected, that the FASIR-NDVI will contain these outliers in the highly averaged dataset and overestimate the NDVI outside of the vegetation season¹.

The high values at the end of the growing season in the EFAI-NDVI curve most likely results from the inability of the current adjustment algorithm to detect the vegetation curve drop at that time in many cases. This issue needs to be addressed in the future to have more evidence for the explanations above.

4.6 Error Estimation

Satellite remote sensing of vegetation has to deal with the various error sources as discussed in section 2.2. Most problems are associated with atmospheric absorption, cloud cover, soil effects, spatial variability, sun-target-sensor geometry estimation, spatial registration errors and sensor degradation during operation. These effects are partially accounted for with empirical correction techniques and allow to estimate spatial and temporal distribution of land surface vegetation. Depending on the use of these datasets, the errors will show lower or higher impact. It is not the intention here to present a mathematical error estimate, but rather to provide an insight into the quality of the presented dataset in relation to their scientific application.

Table 4.2: *Error sources in the remotely-sensed AVHRR NDVI*

Error source	Δ NDVI	Investigator
Ozone absorption	0.0 - 0.03	Townshend et al. (1994)
Water vapor absorption	- (0.01 - 0.10)	Los (1998)
Rayleigh scattering	- (0.02 - 0.04)	Townshend et al. (1994)
Aerosol scattering	- (0.1 - 0.2)	Los (1998)
Overall atmospheric error	- (0.1 - 0.3)	Los et al. (1994)
Soil background reflectance	0.05	Qi et al. (1994)
In-flight sensor degradation	0.07 - 0.10	Rao and Chen (1996)
Clouds	- (0.0 - 0.7)	Los et al. (1994)
Registration error (for 8km resolution)	0.00 - 0.03	Los et al. (1994)
Bidirectional reflectance	-0.10 - 0.05	Los et al. (1994)
Overall errors prior to corrections	0.1 - 0.2	Los (1998)

Errors that produce interferences in the NDVI signal are either resulting from the AVHRR instrument design and the satellite itself or they are related to external sources. The NOAA satellite's instrument degradation is well described by various sources (Rao and Chen 1996, Los 1998) and empirical relationships can be effectively applied to account for this error. The overall satellite instrument degradation is compensated for up to 90-95% with those post-launch calibrations.

-
1. The FASIR-NDVI is sampled 300 times heavier than the EFAI-NDVI dataset presented in this work. The maximum-value composite method has the following effect in this resampling: If one positive outlier occurs in 300 pixels, it will be included in the FASIR-NDVI and trusted as reliable. In comparison, this single outlier was effectively removed here, if it occurs in in early spring. The removal of positive outliers was not performed, but is an issue for future updates (see section 3.7.1).

Atmospheric contamination and cloud cover tend to lower the NDVI value (Table 4.2), which corresponds to the basic assumption A-2 explained in section 2.5. Figures 3.11 a-h indicate that atmospheric effects are successfully eliminated by the Fourier adjustment technique and also by the maximum-value composite method applied to the raw Pathfinder NDVI. The Pathfinder dataset itself is corrected for ozone disturbance and rayleigh scattering (James and Calluri 1994).

The bidirectional reflectance distribution and registration errors are two of the remaining inconsistencies not corrected in the present EFAI-NDVI dataset. These are supposed to appear as non-systematic errors. Partially corrections for external influences (like solar zenith angle or aerosols) will introduce general overestimations or underestimations of the NDVI. Spatially and temporally averaging NDVI-data for climatological use is assumed to significantly reduce those uncertainties. For spatially averaged NDVI, Los et al. (2000) estimates that the inter-annual signal can be detected with an uncertainty of 20-30%, a finding that strongly stands for the usability of remotely-sensed NDVI for climatological purposes. Interannual trends have to be trusted with care though, since the overall error in the dataset is about the same magnitude as the trend signal itself, when applied to a decadal time-series only.

Missing data were not listed as an error source here, but they also pose a serious limitation to the estimation of land surface properties at high latitudes and in tropical regions. The maximum-value compositing method and the Fourier adjustment technique significantly improve the quality of remotely-sensed NDVI datasets, but they can also introduce overestimates as described in sections 2.4 and 3.7.

5. Land Surface Parameterization

5.1 Introduction

Parameterizations of the land surface processes formulate radiative and turbulent transfers between the vegetation, the soil and the lower boundary layer of the atmosphere. Early GCMs demonstrated a high impact of albedo, surface roughness and surface wetness on the atmospheric fields. But first implementations of these surface properties in circulation models were unrealistic, because the fluxes of radiation, sensible and latent heat and momentum were formulated as independent and mostly static processes.

In the mid 1980s, the vegetation-soil system was set up as an independent model (BATS: Dickinson 1984, SiB: Sellers 1986), so that the system could determine its own way how to interact with the atmosphere. These biosphere models have been successively improved in order to be coupled to regional and global circulation models. The main processes include the exchanges of radiation (Figure 5.1, left), the fluxes of sensible and latent heat (Figure 5.1, right) and the frictional forces of the “rough” land surface. A theoretical review of these processes can be found in Sellers et al. (1997).

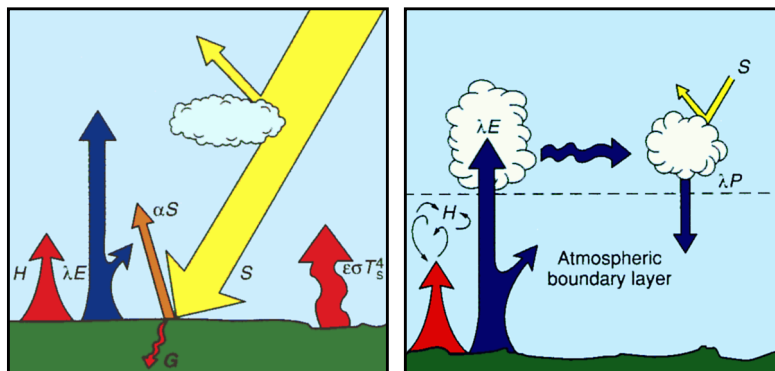


Figure 5.1: Land surface - atmosphere interactions: The global radiation budget, describing the amount of incoming radiant energy absorbed by the land surface (left) and the outgoing heat fluxes (right). (Figures by Sellers et al. 1997)

Biophysical properties of the land surface, especially vegetation itself, determine a substantial amount of the interactive biosphere-atmosphere system. The interactions can be summarized as follows:

Radiation absorption:

Spectral properties of leaves makes vegetation highly absorbent in the visible wavelength interval (PAR, 0.4-0.72 μm) and moderate reflective in the near-infrared region (0.72-4.0 μm). Bare soil shows no such contrast.

Momentum transfer:

Vegetation presents a rough, porous surface to the planetary boundary layer airflow. The resulting turbulence enhances the transport of sensible and latent heat away from the ground

Transpiration:

CO₂ assimilation in plant leaves requires an open pathway between the atmosphere and the water saturated leaf interior and therefore leads to an inevitable loss of water vapor over the same route. These fluxes are driven by the plant stomates, which are regulated by the incident photosynthetical active radiation PAR and various environmental stress factors (temperature, air humidity, leaf water potential) and are characteristic for each plant type. Transpiration is highly sensitive to various atmospheric and surface properties.

Precipitation interception and evaporation:

Soil and plants intercept, store and evaporate precipitation - which results in a decreased precipitation input to the soil, decreases sensible heat flux and increases evaporation.

CO₂ flux

Leaf photosynthesis leads to a CO₂ flux. It is dependent on light (PAR) intensity, CO₂ concentration in the atmosphere and water stress.

Soil moisture availability:

Soil water content (water stress) is significantly altered by vegetation and vice versa.

Insulation:

The soil surface under a dense vegetation canopy intercepts less radiation and the soil energy budget (evaporation, sensible heat flux, ground heat flux) under vegetated canopies is reduced.

All these processes are simulated in the current LSP schemes like SiB2 with individual modules driven by land surface parameters. To calculate land surface radiant and turbulent fluxes for large spatial scales, land surface parameters can be estimated by (satellite) remote sensing techniques described in chapter 2. Especially the surface radiant and turbulence properties are estimated very well by satellite remote sensing through simple relationships to NDVI. Many of the surface processes like carbon, turbulent, latent heat or sensible heat fluxes are determined through these properties and can therefore be spatially and temporally calculated through remotely-sensed surface parameter fields. Sellers et al. (1997) have shown by comparisons to field data that uncertainties in the surface properties are in between 10 to 20%. The calculation of the most important land surface parameters from the presented EFAI-NDVI dataset is performed with the MAPPER program in this chapter. This program was provided by the SiB team and is applied here with several changes included. An overview of the MAPPER program, the soil type and land cover classifications and the relations of the land surface parameters to the NDVI is given in this chapter. A complete theoretical approach to the generation of land surface parameters from remotely sensed NDVI can be found in Sellers et al. (1996a and 1996b).

5.2 Land Cover Classification

In order to generate global fields of land surface parameters (explained in section 5.5), biophysical vegetation and soil properties are used in conjunction with the satellite derived EFAI-NDVI time-series. These properties are available for certain vegetation types from surveys of the ecological literature. The very diverse global flora can be merged into simplified classification schemes. Land cover classifications are found in a variety of classification schemes derived by conventional ground surveys. Sellers et al. (1996a) has presented a global land cover classification with 12 biomes¹, composited from two classifications by Kuchler (1983) and Matthews (1985). Over recent years, researchers have turned to remotely-sensed data to improve the accuracy of datasets that describe the geographic distribution of land cover. DeFries et al. (1998) has developed a classification decision tree to acquire land cover maps by satellite remote sensing. The classification consists of 13 classes as shown in Table 5.1.

Seasonal properties like start, length, end and amplitude of the growing season in combination with surface temperature measurements served to distinguish the different vegetation types. Visible and thermal band time-series of the AVHRR sensor were used and the results were checked at training sites by analyzing high resolution Landsat images. The overall accuracy of this remotely-sensed land cover classification was estimated between 81-90% (dependent on land cover type). The heterogeneous nature of the earth surface at small scales is represented more accurately with the 8km dataset than with the previous 1° by 1° global map, although multi-year datasets will have to be used for more robust classifications in high spatial resolutions (DeFries et al. 1998).

Table 5.1: *Land cover Classes by DeFries et al. (1998)*

Class	Description	% of land surface (Europe)
0	Water	
1	Evergreen needleleaf forest	19.5
2	Evergreen broadleaf forest	0.0
3	Deciduous needleleaf forest	0.0
4	Deciduous broadleaf forest	2.6
5	Mixed forest	7.3
6	Woodland	7.5
7	Wooded grassland	2.8
8	Closed bushland or shrubland	0.6
9	Open shrubland	2.8
10	Grassland	17.2
11	Cropland	31.9
12	Bare Ground	4.9
13	Urban and built-up ^a	2.8

a. Mosses and lichens are merged into the same land cover class, since reflective properties are very similar.

1. SiB2 uses 9 land cover classes. However, for the calculation of land surface parameters, 12 land cover classes are used, since all physical properties are still keyed to the (old) SiB 1 land cover classification .

The DeFries landcover map is reclassified here to be used with SiB2. The 12 SiB vegetation type dependent properties are provided by Sellers et al. (1996b) for each of the 12 SiB biomes (See Appendix A). The Sellers and DeFries classifications do not match for all land cover types, since they were derived for different vegetation criterias, but a reasonable solution for this reclassification is found by the author and shown in Table 5.2. (Map: see appendix B)

A problem results from the fact, that the DeFries does not include a class for Ice. It is suggested here that for a proper estimation of surface properties over glacier and permanent snow cover, this class is patched in from the soil map presented in section 5.3 This problem also becomes obsolete, when a circulation model is used, which simulates snow cover effects independently of the prescribed snow values by land cover class. A way more serious inconsistency is expected from merging C₃/C₄ plant types by this reclassification. The DeFries land cover classes do not distinguish between the two different photosynthesis types. It is this differentiation that makes the SiB classification very sophisticated for use in biosphere models. Collatz et al. (1992) has presented a C₄ photosynthesis model that shows unique responses to environmental conditions when compared to C₃ plants. The carbon assimilation process is more efficient in C₄ plants than in C₃ plants at low CO₂ concentrations and C₄ plants are also favored over C₃ plants in warm, humid climates. Collatz et al. (1998) predicts a serious degradatation in the global distribution of C₄-grasses with a doubled CO₂ scenario.

Table 5.2: Land cover classes used in SiB, matched with corresponding classes from the DeFries classification (% of land surface are calculated within the European domain)

SiB Class	DeFries Class	Description	% of Land surface
0	0	Water	
1	2	Evergreen broadleaf forest (tropical rain forest)	0.0
2	4	Deciduous broadleaf forest (temperate forest)	2.6
3	5	Mixed deciduous broadleaf and evergreen needleleaf forest	7.3
4	1	Evergreen needleleaf forest (boreal forest)	19.5
5	3	Deciduous needleleaf forest (larix)	0.0
6	6	Broadleaf and C ₄ -groundcover (drought woodland or savannah)	7.5
7	7	Grassland and shrub cover	2.8
8	8	Shrubs and groundcover (evergreen broadleaf woodland)	0.6
9	9	Shrubs and bare soil (evergreen broadleaf shrubland)	2.8
10	13	Tundra (Dwarf trees and shrubs)	2.8
11	12	Desert, bare soil	4.9
12	11, 10	Agriculture or C ₃ - grassland	49.2
13	no class	Ice	

C₄ plants are usually of short stature and do not fall in any of the tree classifications. C₄ plants are thought to be represented by the SiB classes 6, 7, 8 and 11. They each correspond to a similar land cover class in the DeFries classification. But it is expected, that grass land cover from the DeFries class 10 will contain a mixture of C₃/C₄ grasses and lead to a missclassification when simply merged into the SiB vegetation class 12 (C₃ grasses only!). However, Collatz et al. (1998) has shown that the relative proportions of C₄ - grasses will not be high in the European continent due to unfavorable environmental conditions. Furthermore it is expected that changes caused by in land cover classification are small, since land surface properties are first order dependent on NDVI and second order dependent on the classification.

5.3 Soil Type Classification

Land surface parameterization schemes like the Simple Biosphere Model (SiB) consider different soil layers for the calculation of the surface albedo and latent as well as sensible heat fluxes between the ground surface and the atmosphere. In SiB2 there are three soil layers present, each one with different properties and thickness (Sellers et al. 1996b): the topmost layer, the root zone layer and the recharge zone layer.

In an early version of the Simple Biosphere model, soil properties were calculated from the distribution of the present vegetation (Sellers et al. 1996a), but it was found, that soil properties have regional variations, which can be independent from vegetation type. Since 1978, a global soil map is available from the Food and Agriculture Organization of the United Nations (FAO) in printed form. It has been continuously updated and is available as digitized GIS data on CDROM (FAO 1995 - Digital soil map of the world).

The digital soil map has the original FAO soil legend assigned, which defines 4930 different soil classes, grouped in 26 major soils. The classification represents the soils found in the upper 30cm of each grid cell. Each grid cell contains one dominant soil type and can have several subdominant soil types also present in the same grid cell. The accuracy of the time-invariant 5' by 5' soil type dataset is unknown, since no error estimates are indicated in the FAO documentation.

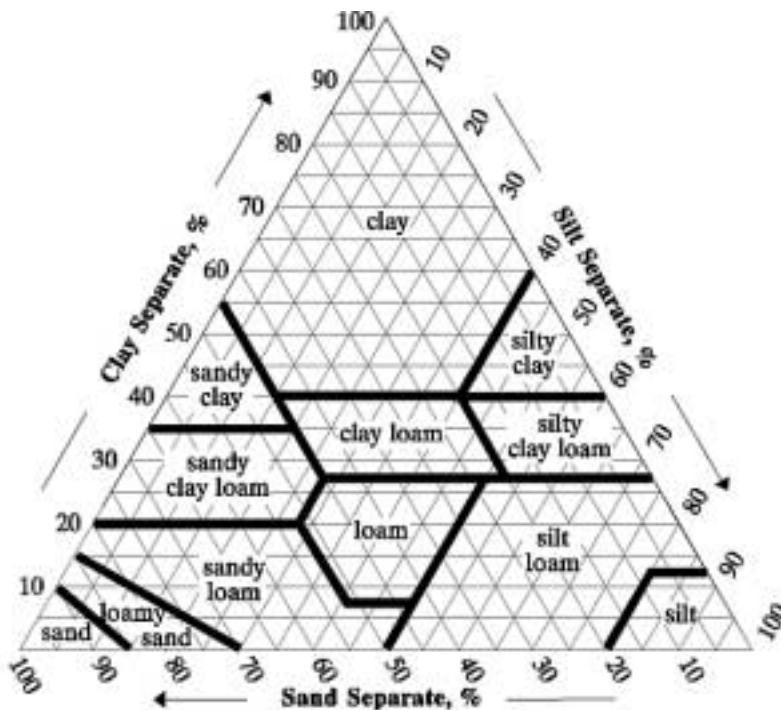


Table 5.3: *SiB soil texture classes*

Nr.	Texture Type
1	Sand
2	Loamy sand
3	Sandy loam
4	Loam
5	Silt Loam
6	Silt
7	Silty clay loam
8	Clay loam
9	Sandy clay loam
10	Silty clay
11	Sandy clay
12	Clay

Figure 5.2: *U.S.D.A. soil texture triangle, where dominant soil types are classified from the percentage of present sand, silt and clay. (U.S.D.A. 1951)*

Soil tension ψ_s , soil wetness exponent B , hydraulic conductivity K_s and the soil porosity Θ_s need to be known by the Simple Biosphere Model in order to calculate the soil dependent energy and mass fluxes (Theoretical background by Sellers et al. 1996a). Los (1998) has matched the seven soil texture types defined by Zobler (1986) with the physical parameters from Clapp and Hornberger (1978).

This classification isn't used here, instead the percentage of clay (fine texture, $d < 0.002\text{mm}$), silt (medium texture, $0.002 < d < 0.05\text{mm}$) and sand (coarse texture, $0.05 < d < 2\text{mm}$) for each of the 4930 texture classes is taken to reclassify the FAO soil type map based on the U.S. Dept. of Agriculture texture triangle (U.S.D.A. 1951). The resulting soil type map (see appendix B) and the physical parameters from Clapp and Hornberger (1978) (appendix A) are combined to produce time-independent European fields of soil tension ψ_s , soil wetness exponent B , hydraulic conductivity K_s and the soil porosity Θ_s . These soil properties vary spatially and the associated maps can be directly used to drive the SiB2 soil scheme.

The 26 main soil types defined by the FAO are extended by the following land surface features. These do not fit in the clay/silt/sand classification used by the U.S.D.A. texture triangle (Figure 5.2):

- Dunes and shifting sand
- Salt flats
- Rock debris or desert detritus
- Glaciers
- Inland water

As a solution, dunes and shifting sand are set to texture class 1 (sand), whereas the other surface features are assigned to no texture class and soil properties are therefore missing at these sites. It needs to be evaluated, how non-soil features are included in the calculation of surface properties.

5.4 Estimation of Net Primary Production through NDVI

Photosynthesis in terrestrial vegetation occurs in chloroplasts, which are cells contained in plant leaves. The reaction process of photosynthesis can be summarized as follows:



Driven by light absorption, carbohydrates are synthesized and molecular oxygen is released during the photosynthesis process. To provide this reaction with a continuous flux of the reactands, leaves have to regulate the transport of air (with a fractional content of CO_2) to the chloroplasts, such that CO_2 can be absorbed and O_2 is released. The transport of these gases occurs through stomata in the upper and lower epidermal surfaces of plant leaves, whereas H_2O is contained in the hydrated mesophyll cells and is uptaken by the plant roots from the soil. A more precise description of the photosynthesis process and leaf physiology can be found in Tucker and Sellers (1986) and Nultsch (1996).

The incident Photosynthetically Active Radiation (PAR) is able to penetrate the upper epidermal layer of leaves, which is largely transparent to PAR (Tucker and Sellers 1986). The PAR spectrum reaches from 0.4 - 0.72 μ m and contains several absorption maximums corresponding to the different types of chlorophyll present in plant leaves (see section 2.1, Figure 2.1). Only a part of the incident PAR is used by the plant leaves to drive the photosynthesis process, also known as Fraction of Photosynthetically Active Radiation absorbed by the plant leaves (FPAR). This distinct spectral absorption behaviour of plant leaves for incident sunlight is measured through the remotely sensed NDVI. NDVI can therefore be used to estimate FPAR. FPAR directly relates to the photosynthesis process and serves to calculate vegetation parameters and the net primary productivity NPP (Sellers et al. 1997).

$$NDVI \propto FPAR \propto NPP \quad (5.2)$$

5.5 Calculation of the Surface Parameters from NDVI

5.5.1 SR - FPAR

The remotely-sensed NDVI represents the amount of green material in the vegetation canopy. For horizontally homogeneous, closed vegetation, which consists almost exclusively out of green material with a dark soil background, it was shown, that FPAR increases almost linearly with the Simple Ratio (SR), as described in Hall et al. (1990) and Sellers et al. (1996b). SR can be transformed from NDVI by the following relationship:

$$SR = \frac{1 + NDVI}{1 - NDVI} \quad (5.3)$$

The linear relationship between the measured SR and FPAR can be derived, if two endpoints are known. These two points are characteristic for each plant type; corresponding NDVI-SR values are listed by vegetation class in Los (1998). The 98 percentile of the NDVI frequency distribution in a specific vegetation class is assumed to represent vegetation with a maximum FPAR value ($FPAR_{max}=0.95$) and the 2% NDVI is assumed to represent no vegetation with the minimum FPAR value ($FPAR_{min}=0.01$). Sellers et al. (1996b) describes the empirical relationship between SR and FPAR (SR-FPAR model):

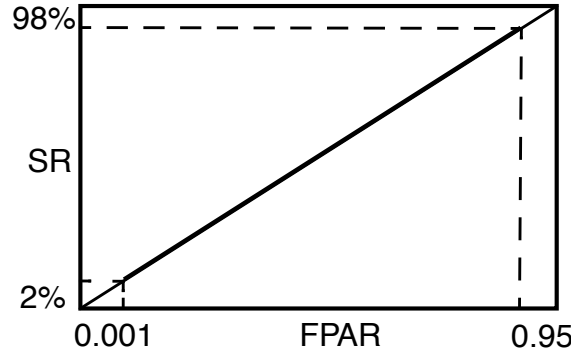


Figure 5.3: Linear relationship between the single ratio (SR) and the fraction of absorbed photosynthetically active radiation (FPAR)

$$FPAR_{SR} = \frac{(SR - SR_{min})(FPAR_{max} - FPAR_{min})}{SR_{max} - SR_{min}} + FPAR_{min} \quad (5.4)$$

SR_{max}, SR_{min} = Land cover dependent, corresponding to the 98% and 2% NDVI intervals

5.5.2 NDVI - FPAR

Goward et al. (1992) and Dye and Goward (1993) use another linear relationship for the FPAR calculation, which bases on NDVI instead of the simple ratio. Coincident NDVI values of FPAR were predicted for a 1984 global AVHRR NDVI dataset according to the equation:

$$FPAR = 1.06 \times NDVI - 0.07 \quad (5.5)$$

Equation 5.5 was modified as follows by Sellers et al. (2000) to have land cover dependent characteristics (referred to as NDVI-FPAR model):

$$FPAR_{NDVI} = \frac{(NDVI - NDVI_{min})(FPAR_{max} - FPAR_{min})}{NDVI_{max} - NDVI_{min}} + FPAR_{min} \quad (5.6)$$

where

$NDVI_{max/min}$ = Land cover dependent, correspond to the 98% and 2% NDVI intervals

5.5.3 SR-NDVI intermediate model

Both models were tested in comparison with ground surveys from FIFE (Hall et al. 1992), OTTER (Angelici et al. 1991) BOREAS (Hall et al. 1993, Chen et al. 1997) and HAPEX-Sahel (Hanan et al. 1997). FPAR values were calculated with the NDVI-FPAR and the SR-FPAR model and compared with the ground measurements and it was shown, that the NDVI-FPAR model performed worst with significant overestimates, whereas the SR-FPAR model significantly underestimated FPAR (Los et al. 2000). Different approaches for directly measuring FPAR at the ground and by estimating FPAR through remote sensing can be found in Gower et al. (1999).

An intermediate model, which calculates the average of both models, shows a good fit to the ground surveys (no significance at the 1% level) and is used in this work for the calculation of FPAR from the remotely-sensed vegetation index.

$$FPAR = \frac{FPAR_{SR} + FPAR_{NDVI}}{2} \quad (5.7)$$

5.5.4 Leaf area index and canopy greenness

Leaf area index (LAI) is another key variable, which is linked to remotely-sensed vegetation indices. LAI stands for the size of one-sided area of leaves per unit of ground area. Dense vegetation has an LAI of around 5 (5m² leaves per m² ground surface). generally vegetation indices approach a saturation level asymptotically for LAI, ranging from 2 to 8, depending on the type of vegetation. Several empirical relationships between have been discussed to estimate LAI from the various vegetation indices (Tucker and Sellers 1986, Baret and Guyot 1991, Sellers et al. 1996b).

Sellers et al. (1996b) has used linear and logarithmic relationships, depending on vegetation type. Comparisons with field experiments (FIFE, BOREAS) have indicated, that linear relationships were invalid at the canopy level (since PAR extincts exponentially when passing through the canopy), but did perform well spatially.

By calculating the vegetation cover fraction f_V of a particular grid cell, the logarithmic FPAR-LAI relationship will hold for the vegetated area (Los et al. 2000):

$$f_V = \max(FPAR_t)/0.95 \quad (5.8)$$

with

f_V = Vegetation cover fraction

$\max(FPAR_t)$ = maximum measured FPAR over the year

Then, FPAR values for the vegetated fraction only are:

$$FPAR_{fv} = FPAR/f_v \quad (5.9)$$

Green leaf area ($L_{G,fv}$) for the vegetation cover fraction is expressed by the following logarithmic relationship:

$$L_{G,fv} = L_{G,max}(i) \frac{\log(1 - FPAR_{fv})}{\log(1 - FPAR_{max})} \quad (5.10)$$

where

$L_{G,max}(i)$ = vegetation type dependent maximum green leaf area index (see appendix A)

$FPAR_{max}$ = Maximum FPAR value = 0.95

Therefore the green leaf area index (LG) for the total grid cell is:

$$L_G = L_{G,fv} \cdot f_V \quad (5.11)$$

It needs to be emphasized, that this FPAR- L_G relationship only accounts for the green portion of the vegetation canopy. A fraction of the vegetation canopy consists of nongreen (supportive) material like stems (stem area index L_S) and dead leaves (dead leaf area index L_D). L_S is assumed to be 0.076

for forests and 0.05 for grasslands and crops (see appendix A). During the vegetation growth period, all leaves are assumed to be green. After the maximum leaf area is reached, dead leaves usually stay within the canopy for one month, before falling off or being eaten. This assumption can be used to calculate the fraction of dead leaves L_D during the senescence period:

$$L_D = 0.5(L_{G, f_V, t-1} - L_{G, f_V, t}) \cdot f_V \quad (5.12)$$

The total leaf area index L_T for one grid cell is the sum of L_G , L_S and L_D :

$$L_T = L_G + L_S + L_D \quad (5.13)$$

And the greenness fraction N^1 , which describes the amount of green phytoelements in the canopy in relation to the total LAI, is then calculated as follows:

$$N = \frac{L_G}{L_T} \quad (5.14)$$

FPAR, LAI, L_G , L_T , N are therefore all time dependent variables calculated from remotely-sensed NDVI. They are used in conjunction with soil and vegetation properties (by soil type and land cover class) to calculate radiative transfers and stomatal resistances in the Simple Biosphere model.

5.5.5 Roughness length

Vegetation accounts for a large part of the turbulent exchange of momentum between the surface and the atmosphere and also influences the transport of sensible and latent heat. The surface roughness length² is altered by vegetation type and activity and therefore also shows a seasonal variation. The L_T fields described above define the roughness length z_0 by this simple exponential expression (Sellers et al. 1996a):

$$z_0 = z_2(1 - b_z e^{-h_z L_T}) \quad (5.15)$$

with

$$\begin{aligned} z_2 &= \text{canopy height [m], vegetation type dependent} \\ b_z &= 0.91 \text{ (empirical)} \\ h_z &= 0.0075 \text{ (empirical)} \end{aligned}$$

1. N is independent of f_V .

2. The roughness length z_0 has a strong influence on the aerodynamic transfer coefficient and shear stress. Details of the aerodynamic transfer model can be found in Sellers et al. (1996a)

5.6 The MAPPER Program

The Simple Biosphere model by Sellers et al. (1996a) uses 37 surface parameters, of which eight are time varying and the others remain static. These parameter fields are calculated by combining a digital land cover map (section 5.2) and soil type map (section 5.3) and remotely-sensed time-series of EFAI-NDVI (chapter 4.) with tables of vegetation and soil properties (appendix A). The relationships between the most important surface properties and NDVI are used as described in section 5.5.

The “preprocessing”¹ of these fields is done with the program MAPPER, which was written in Fortran 90 and provided by the SiB team. It was originally intended to produce global gridded 1° by 1° surface parameters in monthly time-steps. MAPPER routines are adjusted here, so that they now suit for a correct estimation of land surface parameters in the European domain and can be driven with the presented EFAI-NDVI and the carefully derived land cover and soil type maps described in previous sections.

The following programming has been applied to the original MAPPER procedure, now called EUMAPPER:

Table 5.4: *Applied programming to MAPPER.*

Type	Feature
Changed	0.1° by 0.1° resolution (600 by 430 grid points)
Changed	10-day time-steps (36 steps per year)
Changed	latitude-dependent calculation for the European domain
Fixed	Y2k date problem for future time-series
Fixed	leaf-transmission/reflectance parameter handshake between subroutines
Fixed	soil properties calculation with 12 SiB soil classes
Added	masking oceans and missing data pixels
Added	binary interface for all time-varying parameters

The new EUMAPPER program was applied to the full NDVI time-series from 1983 until 1993, what resulted in a set of biophysical land surface properties, that will now have to be validated and tested in the Simple Biosphere model (or any other similar model) to check their accuracy and usability for regional climate studies in the European domain.

The main land surface properties like *NDVI*, *FPAR*, *N*, *L_T* and *z₀* are presented in the appendix B and C without further verification. For *NDVI* and *FPAR*, regional representations of the ISLSCP-II datasets are shown for comparison.

1. The task is described as “preprocessing”, because the real surface-atmosphere interactions like mass, energy and radiant fluxes will only be calculated interactively in the Simple Biosphere model itself, where MAPPER generates the land surface properties as an offline program.

6. Summary and Discussion

This study directly estimated land surface vegetation from satellite observations for the years 1983 until 1993 and associated biophysical parameters were calculated from the generated EFAI-NDVI dataset. Since remote sensing data is subject to large error sources, these have to be accounted for prior to their use for climatological purposes. Simple assumptions can be made for discriminating suspect NDVI values from more reliable data, what leads to the application of a Fourier-series adjustment scheme presented in section 2.5. This adjustment scheme was applied to the GIMMS-NDVI by Los (1998) and is not very effective here. The present dataset has a high spatial and temporal resolution for regional use and is not subsampled, thus includes more unreliable data, what renders the detection of the correct vegetation signal very difficult. It is shown in sections 3.4 to 3.6 that a few modifications can be made to the original Fourier adjustment algorithm to adapt its use to this regional NDVI dataset. Especially of importance is the finding that outliers in NDVI time-series do not always result in decreased values. Positive outliers mainly occur outside the growing seasons, where NDVI is supposed to be low and they therefore have a high impact. This special case has not been considered before, when coarse-resolution vegetation datasets were generated by Los (1998), Los et al. (2000).

However, several anomalies in NDVI time-series are not covered well with the enhanced Fourier adjustment algorithm (section 3.7). High latitude, boreal forests and crops have distinct time-series, that do not represent the assumed seasonal trend serving as the basis for the applied correction scheme. It needs to be evaluated in the future whether a vegetation- and latitude-dependent correction algorithm will better allow to detect the true seasonal behavior of vegetation for all landscapes in Europe. With the current state of the correction scheme, an overestimation of the NDVI in summer months and at the end of the growing season is possible, especially for vegetation with a high seasonality in high latitudes. Visual inspections of time-series show that the seasonal behavior of land surface vegetation in winter and spring is represented well.

Vegetation is very heterogeneous at canopy level, but recent findings by Justice et al. (1985), Los (1998) and DeFries et al. (1998) support the use of a few land cover classes that differ in their seasonal behavior. With the present EFAI-NDVI, seasonal characteristics are detectable for the 12 SiB land cover classes (section 4.2) and can be trusted, since their magnitude is high in comparison to possible error sources. In addition to seasonal variability, inter-annual change in the vegetation signal occurs in response to natural and anthropogenic forcings through changing environmental conditions. The year-to-year variations have a higher confidence than a detectable (possible CO₂-induced¹) trend in the overall dataset (section 4.3), because naturally occurring environmental changes of 2-4 years in length are of reasonable magnitude. The hypothesis at the beginning of this report suggested a high resolution vegetation dataset for regional climate modelling in favor of the currently available coarse resolution vegetation parameterizations. Although finer temporal scales certainly allow to better discriminate a proper vegetation signal from an error-contaminated time-series (section 2.6), this assumption does not necessarily justify a finer spatial grid. Some regions show substantial vegetation variability over short distances (less than 50km), where other areas are successfully approached by the conventional 1° by 1° resolution. This finding does need further examination, because SVATS for regional climate modelling will have to deal with either coarse or high resolution datasets in the future, what substantially influences their need for computing power.

1. See Myneni et al. (1997)

As presented in section 5.5, land surface parameters are calculated through simple relationships from NDVI. This connection is central to all currently used large-scale LSPs. The NDVI-FPAR, FPAR-LAI and the LAI- z_0 models have been validated recently in field experiments and are proved to be reliable if NDVI is provided accurately. The correct estimation of NDVI is therefore a very important task in the generation consistent fields of land surface properties for climatological research. Sophisticated parameterizations of land surface processes in SVATS like SiB would not be possible without satellite remote sensing. Ground measurements of land surface properties are either not available for wide areas, show inconsistencies or lack of temporal resolution.

Satellite data will continue to be the only feasible way to obtain updates about the state of global vegetation in a timely fashion. Almost all remotely-sensed vegetation datasets base on measurements by the AVHRR instrument (section 2.2), which was used on several NOAA polar orbiting satellites during the last 20 years. An improvement is expected from the Moderate-resolution Imaging Spectroradiometer (MODIS) - one of five state-of-the-art instruments onboard the TERRA satellite, that was launched in December 1999 by NASA. MODIS is supposed to expand the time-series of global vegetation monitoring and will improve the ability to better detect vegetation on the land surface as well as phytoplankton activity in oceans with its 36 sharp bands in the visible and infrared spectrum. It has better ability to detect atmospheric disturbances like aerosols and water vapor and also has an improved resolution of 250-500m for local vegetation studies. Future biospheric research will profit from the quality of MODIS vegetation data and the sources and sinks of carbon dioxide in response to natural and anthropogenic climate changes are expected to be better understood and estimated.

The EFAI-NDVI dataset presented in this study should be seen as a basis for specifying land surface processes and for improving coupled biosphere - atmosphere models in the European domain. As a possible future application, the presented land surface parameter fields will have to be tested in sensitivity experiments which show regional climate response to altered vegetation properties. The inter-annual variations within the presented vegetation dataset are reflecting actual responses to environmental changes. Experiments involving this dataset should be run to better understand hydrological, radiative and chemical feedback mechanisms in the soil-vegetation-atmosphere system. Furthermore, the presented multi-year dataset needs to be compared to actual climate records as temperature and precipitation data and also needs to be validated with other vegetation datasets.

References

- Angelici, G.L., Skiles, J.W. and Popovici, L.Z. (1991). *OTTER Pilot Land Data System, Volume 1, Version 1*: Satellite, Aircraft, and Ground Measurements. Published on CD-ROM by NASA. Available on CD-ROM and online [<http://www-eosdis.ornl.gov>] from the ORNL Distributed Active Archive Center.
- Anyamba, A. and Eastman J.R. (1996). Interannual variability of NDVI over Africa and its relation to El Nino Southern Oscillation. *Int. J. Remote Sens.*, 17: 2533-2548.
- Avissar, R. (1995). Recent advances in the representation of land-atmosphere interactions in general circulation models. *Rev. Geophys., Supplement to Volume 33, Part 2*, US National Report 1991-1994, 1005-1010.
- Baldwin, D. (1996). AVHRR Pathfinder Goode's projection to equal angular projection. Navigation and projection routine provided through the *AVHRR Pathfinder Group at the NASA/ Goddard Space Flight Center*, Greenbelt, MD, USA.
- Baret, F. and Guyot, G. (1991). Potentials and Limits of Vegetation Indexes for LAI and APAR assessment. *Remote Sens. Environ.*, 35: 161-173
- Belward, A.S. (1992). Spatial attributes of AVHRR imagery for environmental monitoring. *Int. J. Remote Sens.*, 13: 193-208.
- Betts, A.K. and Ball, J.H. (1997). Albedo over the boreal forest. *J. Geophys. Res.*, 102: 28901-28909.
- Bounoua, L., Collatz, G.J., Sellers, P.J., Randall, D.A., Dazlich, D.A., Los, S.O., Berry, J.A., Fung, I., Tucker, C.J., Field, C.B., Jensen, T.G. (1999). Interactions between vegetation and climate: Radiative and physiological effects of doubled atmospheric CO₂. *J. Climate*, 12: 309-324.
- Chen, J.M., Rich, P.M., Gower, S.T., Norman, J.M. and Plummer, S. (1997). Leaf area index of boreal forests: theory, techniques, and measurements. *J. Geophys. Res.*, 102: 29429-29444.
- Cihlar, J., Manak, D. and Voisin, N. (1994). AVHRR Bidirectional Reflectance Effects and Compositing. *Remote Sens. Environ.*, 48: 77-88.
- Cihlar, J., Ly, H., Li, Z., Chen, J., Pokrant, H. and Huang, F. (1997). Multitemporal, Multichannel AVHRR Data Sets for Land Biosphere Studies - Artifacts and Corrections. *Remote Sens. Environ.*, 60: 35-57.
- Clapp, R.B. and Hornberger, G.M. (1978). Empirical equations for some soil hydraulic properties. *Water Resour. Res.*, 14.4: 601-604
- Clark, R.N. and Roush, T.L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res.*, 89B: 6329-6340.
- Collatz, G.J., Ribas-Carbo, M. and Berry, J.A. (1992). Coupled photosynthesis stomatal conductance for leaves of C₄ plants. *Aust. J. Plant Physiol.*, 19: 519-538.
- Collatz, G.J., Berry, J.A. and Clark, J.S. (1998). Effects of climate and atmospheric CO₂ partial pressure on the global distribution of C₄ grasses: present, past, and future. *Oecologia*, 114: 441-454.
- DeFries, R.S., Hansen, M., Townshend, J.R.G. and Sohlberg, R. (1998). Global land cover classification at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers. *Int. J. Remote Sens.*, 19: 3141-3168.

- Dickinson, R.E. (1984). Modeling evapotranspiration for three-dimensional global climate models. *Climate Processes and Climate Sensitivity, Geophysical Monograph Amer. Geophys. Union*, 29: 58-72.
- Dye, D.G. and Goward, S.N. (1993). Photosynthetically active radiation absorbed by global land vegetation in August 1984. *Int. J. Remote Sens.*, 14: 3361-3364.
- Eklundh, L.R. (1995). Noise estimation in NOAA AVHRR maximum-value composite NDVI images. *Int. J. Remote Sens.*, 16: 2955-2962.
- Fanning, D. (1999). *IDL Programming Techniques*. Fanning Software Consulting, Bradbury Court, 336 pp.
- Farquhar, G.D., Caemmerer, S.v. and Berry J.A. (1980). A Biochemical Model of Photosynthetic CO₂ Assimilation in Leaves of C₃ Species. *Planta*, 149: 78-90.
- Food and Agriculture Organization (1995). *A Digital Soil Map of the World*. Published on CD-ROM by FAO, Roma, Italy.
- Fournier, R.A., Rich, P.M. and Landry, R. (1997). Hierarchical characterization of canopy architecture for boreal forest. *J. Geophys. Res.*, 102: 29445-29454.
- Fung, I.Y., Field, C.B., Berry, J.A., Thompson, M.V., Randerson, J.T., Malmström, C.M., Vitousek, P.M., Collatz, G.J., Sellers, P.J., Randall, D.A., Denning, A.S., Badeck, F. and John, J. (1997). Carbon 13 exchanges between the atmosphere and biosphere. *Glob. Biochem. Cycles*, 11:507-533.
- Goward, S.N., Markham, B., Dye, D.G., Dulaney, W. and Yang, J. (1991). Normalized Difference Vegetation Index measurements from the Advanced Very High Resolution Radiometer. *Remote Sens. Environ.*, 35: 257-277.
- Goward, S.N., Dye, D.G., Turner, S. and Yang, J. (1993). Objective assessment of the NOAA global vegetation index data product. *Int. J. Remote Sens.*, 14: 3365-3394.
- Goward, S.N., Turner, S., Dye, D.G. and Liang, S. (1994). The University of Maryland improved Global Vegetation Index product. *Int. J. Remote Sens.*, 15: 3365-3395.
- Gower, S.T., Kucharik, C.J. and Norman J.M. (1999). Direct and Indirect Estimation of Leaf Area Index, f_{APAR} and Net Primary Production of Terrestrial Ecosystems. *Remote Sens. Environ.*, 70: 29-51.
- Gutman, G.G. (1991). Vegetation Indices from AVHRR: An Update and Future Prospects. *Remote Sens. Environ.*, 35: 121-136.
- Gutman, G.G. and Ignatov, A. (1995). Global land monitoring from AVHRR: potentials and limitations. *Int. J. Remote Sens.*, 16: 2301-2309.
- Hall, F.G., Huemmrich, K.F., Goward, S.N. (1990). Use of narrow band spectra to estimate the fraction of absorbed photosynthetically active radiation. *Remote Sens Environ.*, 32: 47-54.
- Hall, F.G., Huemmrich, K.F., Goetz, S.J., Sellers, P.J. and Nickeson J.E. (1992): Satellite remote sensing of surface energy balance: success, failures, and unresolved issues in FIFE. *J. Geophys. Res.*, 97: 19061-19089.
- Hall, F.G., Sellers, P.J., Apps, M., Baldocchi, D., Cihlar, J., Goodison, B., Margolis, H. (1993). BOREAS: Boreal Ecosystem-Atmosphere Study. *IEEE Geosci. Remote Sens. Soc. Newslet.*, March: 9-17
- Hanan, N.P., Bégué, A. and Prince S.D. (1997). Errors in remote sensing of intercepted photosynthetically active radiation: an example from HAPEX-Sahel. *J. Hydrol.*, 188-189: 676-696.
- Heck, P. (1999). European-Scale Vegetation-Climate Feedbacks Since the Time of the Romans. A Sensitivity Study Using a Regional Climate Model. *Zürcher Klima-Schriften 77*, Geographisches Institut ETH, Zürich, Switzerland.

- Holben, B.N. (1986). Characteristics of maximum-value composite images for temporal AVHRR data. *Int. J. Remote Sens.*, 7: 1435-1445.
- Huete, A.R. and Tucker, C.J. (1991). Investigation of soil influences in AVHRR vegetation imagery. *Int. J. Remote Sens.*, 12: 1223-1242.
- Hurrell, J.W. and Trenberth K.E. (1997). Spurious trends in satellite MSU temperatures from merging different satellite records. *Nature*, 386: 164-167.
- James, M.E. and Kalluri, S.N.V. (1994). The Pathfinder AVHRR land dataset: An improved coarse resolution dataset for terrestrial monitoring. *Int J. Remote Sens.*, 15: 3347-3363.
- Justice, C.O., Townshend, J.R.G., Holben, B.N. and Tucker, C.J. (1985). Analysis of the phenology of global vegetation using meteorological satellite data. *Int J. Remote Sens.*, 6: 1271-1318.
- Kreyszig, E. (1988). *Advanced Engineering Mathematics*. John Wiley & Sons, New York, 1293 pp.
- Kuchler, W.A. (1983). World map of natural vegetation. *Goode's World Atlas*, 16th ed., Rand McNally, 16-17.
- Los, S.O., Justice, C.O. and Tucker C.J. (1994). A global 1° by 1° NDVI dataset for climate studies derived from the GIMMS continental NDVI. *Int J. Remote Sens.*, 15: 3493-3518.
- Los, S.O. (1998). Linkages between Global Vegetation and Climate: An Analysis Based on NOAA-Advanced Very High Resolution Radiometer Data. *Ph.D. Dissertation*, Vrije Universiteit Amsterdam, 199 pp.
- Los, S.O., Collatz, G.J., Sellers, P.J., Malmström, C.M., Pollack, N.H., DeFries, R.S., Bounoua, L., Parris, M.T., Tucker, C.J. and Dazlich, D.A. (2000). A global 9-Year Biophysical Land Surface Dataset From NOAA AVHRR Data, J. Hydrometeor., Accepted for publication in *J. Hydrometeor.*.
- Lüthi, D., C. Frei, C. Schär and H. C. Davies (1997): A regional climate model for the Alpine region. Final report NFP, Institute for Atmospheric Sciences LAPETH, ETH Zürich, 31, 53pp.
- Matthews, E., (1983). Global Vegetation and Land Use: New High-Resolution Data Bases for Climate Studies. *J. Climate App. Meteor.*, 22: 474-487.
- Meeson, B.W., Corprew, F.E., McManus, J.M.P., Myers, D.M., Closs, J.W., Sun, K.-J., Sunday, D.J. and Sellers, P.J. (1995). *ISLSCP Initiative I: Global Data Sets for Land-Atmosphere Models, 1987-1988, Volumes 1-5*, NASA, Published on CDROM
- Moulin, S., Kergoat, L., Vioviy, N. and Dedieu, G. (1997). Global-Scale Assessment of Vegetation Phenology Using NOAA/AVHRR Satellite Measurements. *J. Climate*, 10: 1154-1170.
- Myneni, R.B., Keeling, C.D., Tucker, C.J., Asrar, G. and Nemani, R.R. (1997). Increased plant growth in the northern high latitudes from 1981-1991. *Nature*, 386: 698-702.
- National Oceanic and Atmospheric Administration (NOAA) and National Aeronautics and Space Administration (NASA) Pathfinder Program (1999), *NOAA/NASA Pathfinder AVHRR Land Data Sets*. Available online [http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/LAND_BIO/GLBDST_Documentation.html].
- Nultsch, W. (1996). *Allgemeine Botanik*. Thieme, Stuttgart.
- Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H. and Sorooshian, S. (1994). A Modified Soil Adjusted Vegetation Index. *Remote Sens. Environ.*, 48: 119-126
- Quarmby, N.A., Milnes, M., Hindle, T.L. and Silleos, N. (1993). The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. *Int. J. Remote Sens.*, 14: 199-210.
- Pielke, R.A., Walko, R.L., Steyaert, L.T., Vidale, P.L., Liston, G.E., Lyons, W.A. and Chase, T.N. (1999). The influence of anthropogenic landscape changes on weather in south Florida. *Mon. Weather Rev.*, 127: 1663-1673

- Potter, C.S. and Brooks, V. (1998). Global analysis of empirical relations between annual climate and seasonality of NDVI. *Int. J. Remote Sens.*, 19: 2921-2948.
- Prince, S.D. and Goward, S.N. (1996). Evaluation of the NOAA/NASA Pathfinder Land Data Set for global primary production modelling. *Int. J. Remote Sens.*, 17: 217-221.
- Rao, C.R. and Chen, J. (1996). Post-launch calibration of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-14 spacecraft. *Int. J. Remote Sens.*, 17: 2743-2847.
- Richards, J.A. and Jia, X (1999). *Remote Sensing Digital Image Analysis - An Introduction*. Springer, Berlin. 355 pp.
- Robinson, T.P. (1996). Spatial and temporal accuracy of coarse resolution products of NOAA-AVHRR NDVI data. *Int. J. Remote Sens.*, 17: 2303-2321.
- Sellers, P.J., Mintz, Y, Sud, Y.C. and Dalcher, A. (1986). A simple biosphere model (SiB) for use with general circulation models. *J. Atmos. Sci.*, 43: 505-531.
- Sellers, P.J., Hall, F.G., Asrar, G., Strebel., D.E. and Murphy, R.E. (1988). The First ISLSCP Field Experiment (FIFE). *Bull. Amer. Meteor. Soc.*, 69: 22-27.
- Sellers, P.J., Los, S.O., Tucker, C.J., Justice, C.O., Dazlich, D.A., Collatz, G.J. and Randall, D.A. (1994). A global 1° by 1° NDVI dataset for climate studies. Part 2: The generation of global fields of terrestrial biophysical parameters from the NDVI. *Int. J. Remote Sens.*, 15: 3519-3545.
- Sellers, P.J., Randall, D.A., Collatz, G.J., Berry, J.A., Field, C.B., Dazlich, D.A., Zhang, C. and Bounoua, L. (1996a). A revised land surface parameterization (SiB2) for GCMs. Part 1: Model Formulation. *J. Climate.*, 9: 676-705.
- Sellers, P.J., Los, S.O., Tucker, C.J., Justice, C.O., Dazlich, D.A., Collatz, G.J. and Randall, D.A. (1996b). A revised land surface parameterization (SiB2) for atmospheric GCMs. Part 2: The generation of global fields of terrestrial biophysical parameters from satellite data. *J. Climate.*, 9: 706-737.
- Sellers, P.J., Dickinson, R.E., Randall, D.A., Betts, A.K., Hall, F.G., Berry, J.A., Collatz, G.J., Denning, A.S., Mooney, H.A., Nobre, C.A., Sato, N, Field, C.B. and Henderson-Sellers, A. (1997). Modeling the Exchanges of Energy, Water, and Carbon Between Continents and the Atmosphere. *Science*, 275: 502-509.
- Steinwand, D.R. (1994). Mapping raster imagery to the Interrupted Goode Homolosine projection. *Int. J. Remote Sens.*, 15: 3463-3471.
- Taddei, R. (1997). Maximum Value Interpolated (MVI): a Maximum Value Composite method improvement in vegetation index profiles analysis. *Int. J. Remote Sens.*, 18: 2365-2370.
- Townshend, J.R.G. and Justice, C.O. (1990). The spatial variation of vegetation changes at very coarse scales. *Int. J. Remote Sens.*, 11: 149-157.
- Tucker, C.J. and Matson, M. (1985). Determination of volcanic dust deposition of El Chichon from ground and satellite data. *Int. J. Remote Sens.*, 6: 619-627.
- Tucker, C.J. and Sellers, P.J. (1986). Satellite remote sensing of primary production. *Int. J. Remote Sens.*, 7: 1395-1416.
- U.S. Dept. of Agriculture (1951). Soil Survey Manual. U.S. Dept. of Agriculture Agricultural Handbook, 18: 503pp.
- Vermaas, W. (1998), An Introduction to Photosynthesis and Its Applications. *The World & I*, March Issue: 158-165.
- Weisberg, S. (1985), *Applied Linear Regression*. John Wiley & Sons, New York, 324 pp.
- Zobler, L. (1986). A world soil file for global climate modeling. NASA Technical Memorandum 87802, NASA, Washington D.C., 35 pp.

Abbreviations

AGCM	Atmospheric General Circulation Model
AVHRR	Advanced Very High Resolution Radiometer
BATS	Biosphere Atmosphere Transfer Scheme
BOREAS	Boreal Ecosystem-Atmosphere Study
BRDF	Bidirectional Reflectance Distribution Function
CHRM	Climate HRM
EFAI-NDVI	European Fourier-Adjusted and Interpolated NDVI
ENSO	El Niño-Southern Oscillation
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System
ETH	Eidgenössische Technische Hochschule
FAO	United Nations Food and Agriculture Organisation
FASIR	Fourier Adjustment, Solar zenith angle correction, Interpolation and Reconstruction of NDVI
FFT	Fast Fourier Transform
FIFE	First ISLSCP Field Experiment
FPAR	Fraction of PAR absorbed by the green part of the vegetation
GAC	Global Area Coverage
GCM	General Circulation Model
GIMMS	Global Inventory, Monitoring and Modeling System
GIS	Geographic Information Systems
GSFC	Goddard Space Flight Center
GVI	Global Vegetation Index
HRM	High Resolution Model
IDL	Interactive Data Language
IFOV	Instantaneous field-of-view
ISLSCP	International Satellite Land Surface Climatology Project
LAI	Leaf Area Index
LSP	Land Surface Parameterization
MODIS	Moderate Resolution Imaging Spectroradiometer
MSAVI	Modified SAVI
MSS	Landsat Multispectral Scanner
MVC	Maximum-value composite
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared (wavelength interval)
NOAA	National Oceanic and Atmospheric Administration
NPP	Net Primary Production
OTTER	Oregon Transect Ecosystem Research

PAR	Photosynthetically Active Radiation
POES	Polar-orbiting Operational Environmental Satellite
RAMS	Regional Atmospheric Modeling System
RCM	Regional Climate Model
SAVI	Soil Adjusted Vegetation Index
SiB	Simple Biosphere model
SPOT	Satellite pour l'Observation de la Terre
SR	Simple Ratio
SST	Sea Surface Temperature
SVATS	Soil-Vegetation-Atmosphere Transfer Scheme
TOMS	Total Ozone Mapping Spectrometer
U.S.D.A.	United States Department of Agriculture
VIS	Visible (wavelength interval)

Appendix

A. Tables

Table A.1: Physical soil properties: soil tension ψ_s , soil wetness exponent B , hydraulic conductivity K_s and the soil porosity Θ_s for each of the 12 USDA soil texture classes (see chapter 5). Soil class dependent properties after Clapp and Hornberger (1978).

Nr.	Class	Texture Type	B	ψ_s	K_s	Θ_s
1	S	Sand	5.39	-0.15	7.0	0.45
2	LS	Loamy sand	5.30	-0.57	7.2	0.49
3	SL	Sandy loam	5.30	-0.57	7.2	0.49
4	L	Loam	4.90	-0.07	34.7	0.44
5	SiL	Silt loam	4.38	-0.02	156	0.41
6	Si	Silt	4.05	-0.04	176	0.40
7	SiCL	Silty clay loam	7.12	-0.09	6.3	0.42
8	CL	Clay loam	10.4	-0.06	2.2	0.43
9	SCL	Sandy clay loam	8.52	-0.36	2.5	0.48
10	SiC	Silty clay	7.75	-0.15	1.7	0.48
11	SC	Sandy clay	10.4	-0.17	1.0	0.49
12	C	Clay	11.4	-0.19	1.3	0.48

Table A.2: Vegetation class dependent morphological parameters used in MAPPER. (Section 7.4). Class= SiB vegetation class, z_c =inflection height for leaf-area density, l_w =leaf width, l_l =leaf length, $L_{G,max}$ =maximum Leaf Area Index, L_S =Stem Area Index, $NDVI_{max}$ =maximum NDVI in the 98% interval, $NDVI_{min}$ =minimum NDVI in the 2% interval, $a_{s,v}$ =soil visible reflectance, $a_{s,n}$ = soil near-infrared reflectance. Class 13 is the ice class and was not used here.

Class	z_c	l_w	l_l	$L_{G,max}$	L_S	$NDVI_{max}$	$NDVI_{min}$	$a_{s,v}$	$a_{s,n}$
1	28.2	0.050	0.100	7.0	0.08	0.6181	0.03385	0.110	0.225
2	17.0	0.080	0.150	7.0	0.08	0.6868	0.03385	0.110	0.225
3	15.0	0.040	0.100	7.5	0.08	0.6868	0.03385	0.110	0.225
4	10.0	0.001	0.055	8.0	0.08	0.6868	0.03385	0.110	0.225
5	10.0	0.001	0.040	8.0	0.08	0.6868	0.03385	0.110	0.225
6	0.55	0.010	0.300	5.0	0.05	0.6181	0.03385	0.110	0.225
7	0.55	0.010	0.300	5.0	0.05	0.6302	0.03385	0.110	0.225
8	0.55	0.010	0.300	5.0	0.05	0.6181	0.03385	0.150	0.250
9	0.3	0.003	0.030	5.0	0.05	0.6302	0.03385	0.300	0.350
10	0.35	0.010	0.300	5.0	0.05	0.6302	0.03385	0.110	0.225
11	0.55	0.010	0.300	5.0	0.05	0.6302	0.03385	0.300	0.350
12	0.55	0.010	0.300	6.0	0.05	0.6302	0.03385	0.100	0.150
13	0.55	0.010	0.300	5.0	0.01	0.6302	0.03385	0.300	0.350

B. Classifications

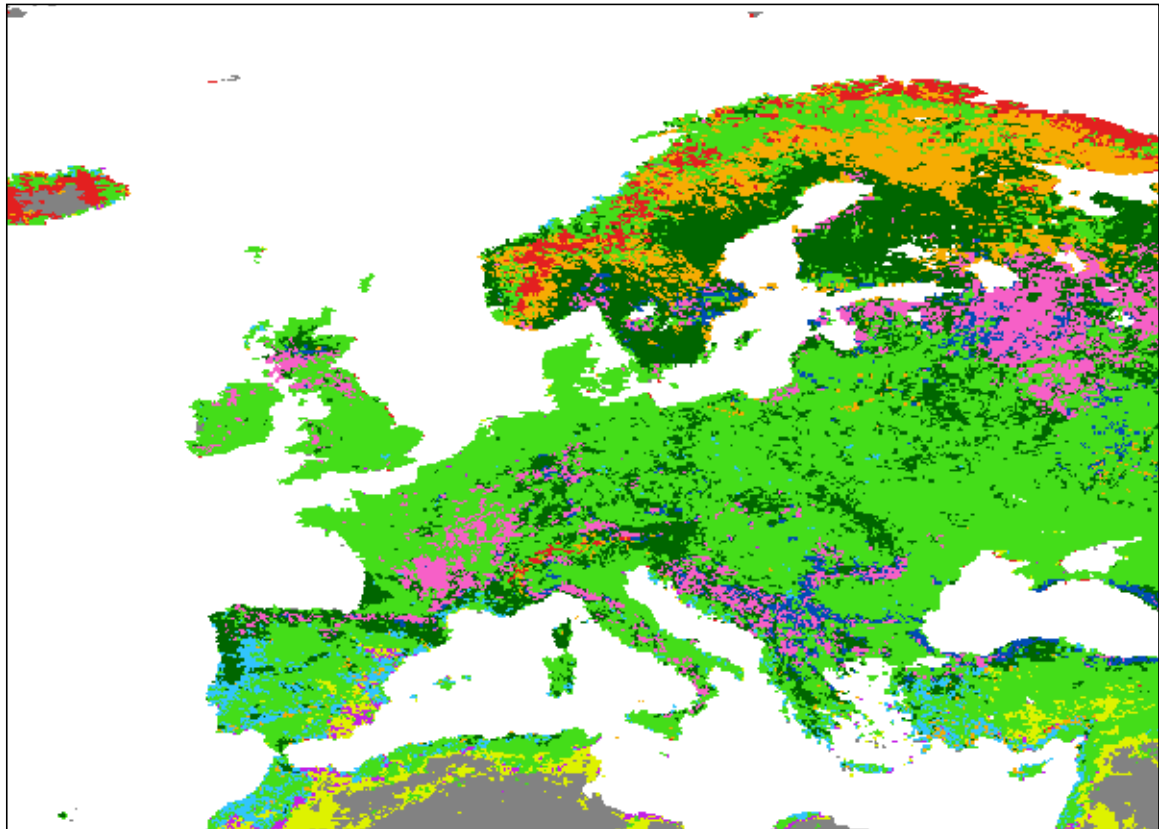


Figure B.1: Land cover classification map derived for the 12 SiB classes. Ice class 13 is not derived in this map. The map has a spatial resolution of 0.1° by 0.1°.

	Water
1	Evergreen broadleaf forest
2	Deciduous broadleaf forest
3	Mixed broadleaf & needleleaf forest
4	Evergreen needleleaf forest
5	Deciduous needleleaf forest
6	Broadleaf and C4-groundcover
7	Grassland and shrub cover
8	Shrubs and groundcover
9	Shrubs and bare soil
10	Tundra
11	Desert, bare soil
12	Agriculture or C3 - grassland

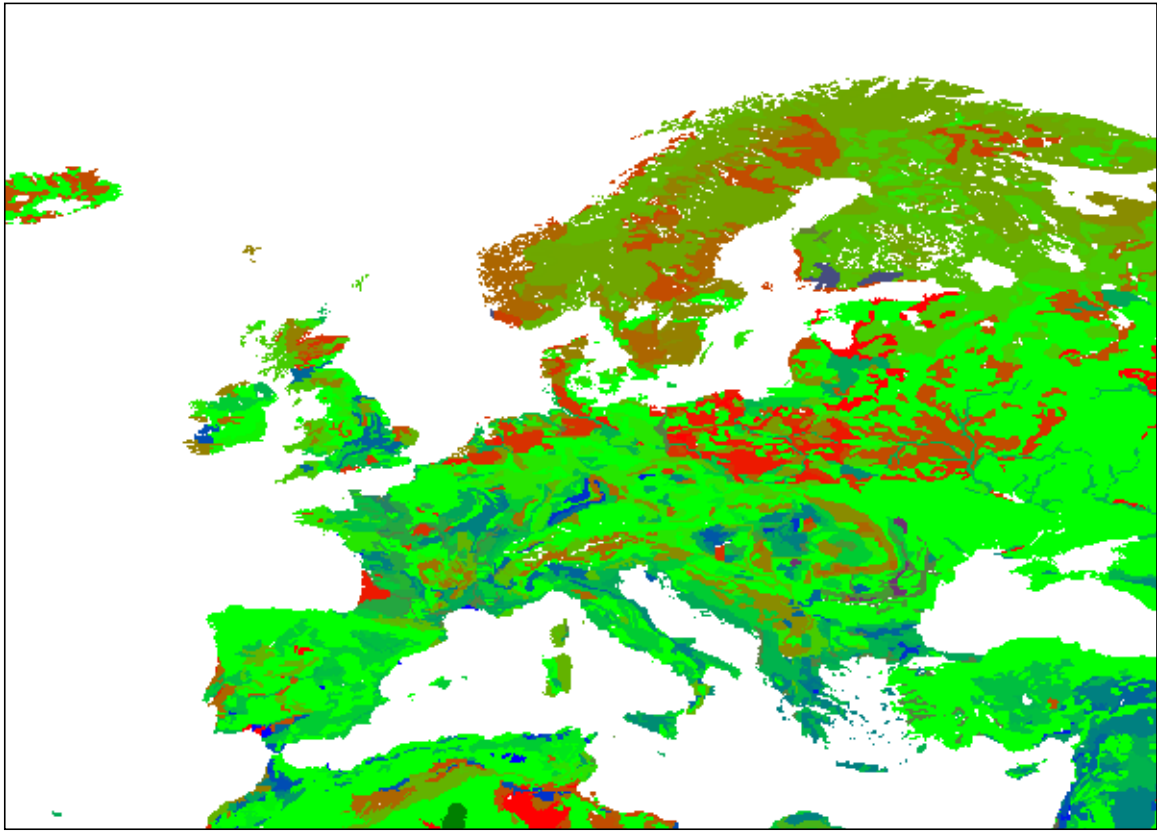


Figure B.2: The topmost dominant soil type described as %sand (red), %silt (green) and %clay(blue). The white areas within the landsurface are either rock, salt, ice or inland water and are therefore not described in this classification. The map is shown in a 0.1° by 0.1° resolution.



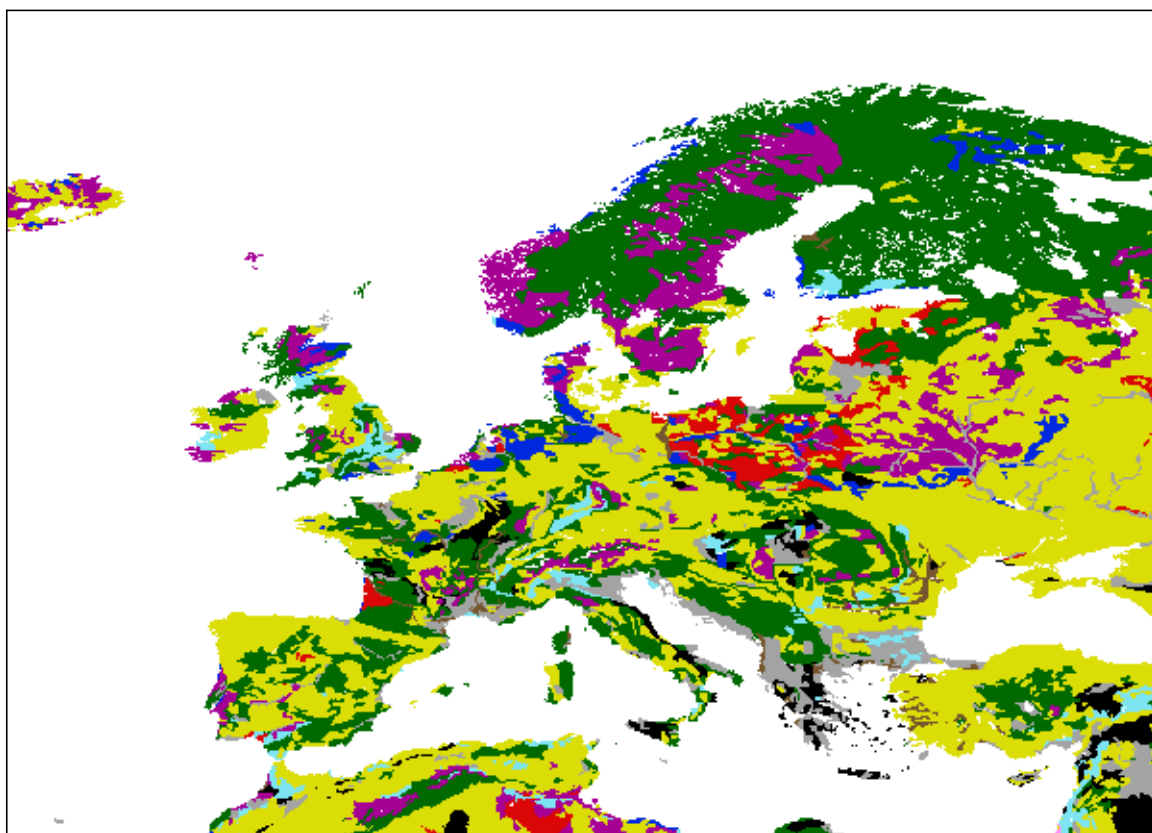


Figure B.3: The European soil map is derived from the FAO digital soil map for the 12 SiB soil classes. Ice, Salt, Rock and inland water areas are set to missing (white) and are not considered at this time. The spatial resolution is 0.1° by 0.1° .

		Water or missing
	1	Sand
	2	Loamy sand
	3	Sandy loam
	4	Loam
	5	Silt loam
	6	Silt
	7	Silty clay loam
	8	Clay loam
	9	Sandy clay loam
	10	Silty clay
	11	Sandy clay
	12	Clay

C. Datasets

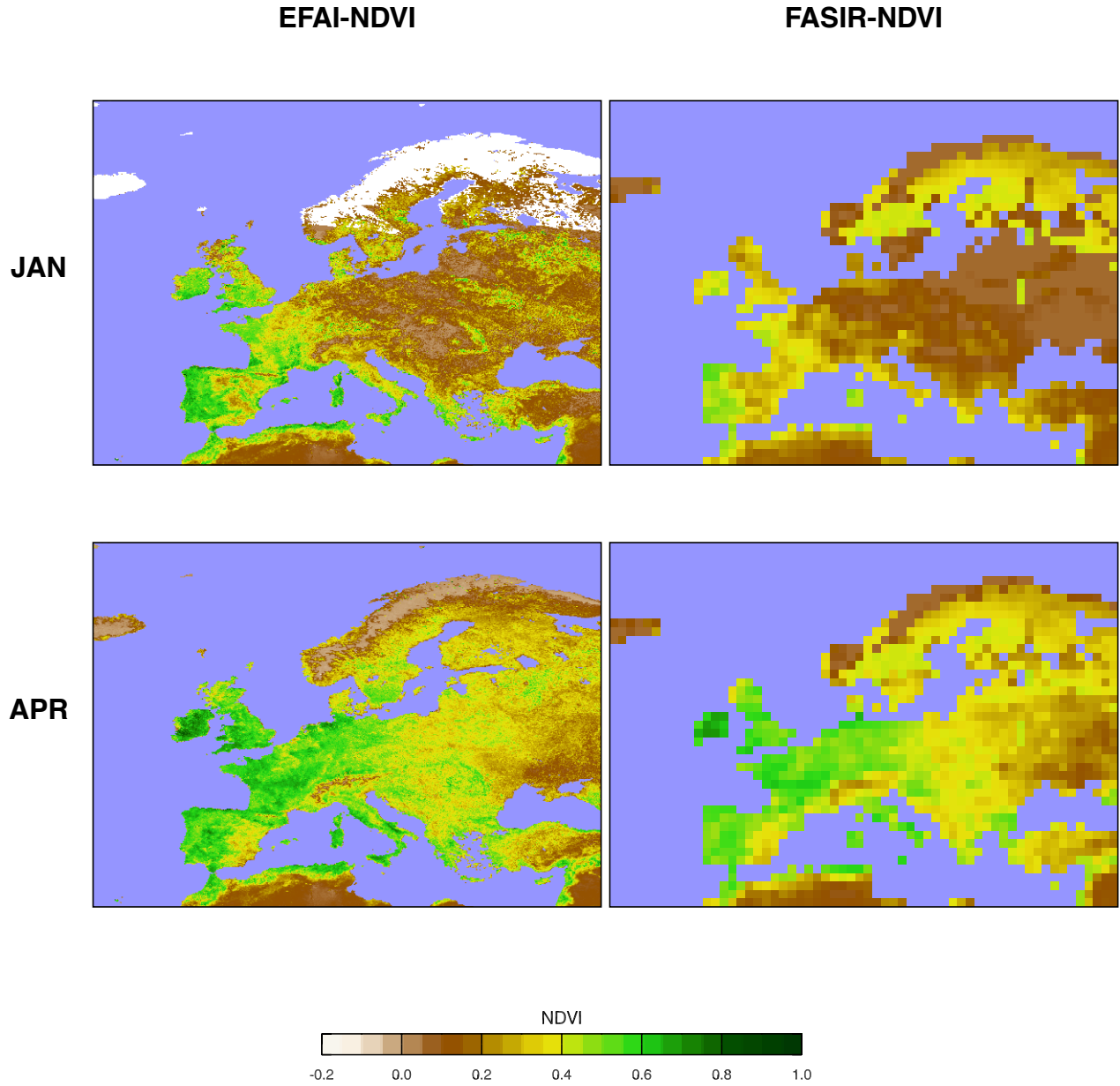


Figure C.1: 0.1° by 0.1° EFAI-NDVI (left) shown in comparison to the pre-released 1° by 1° FASIR-NDVI (right). The EFAI-NDVI is sampled for one month, since it has a temporal resolution of 10 days. top = January 1987, bottom = April 1987.

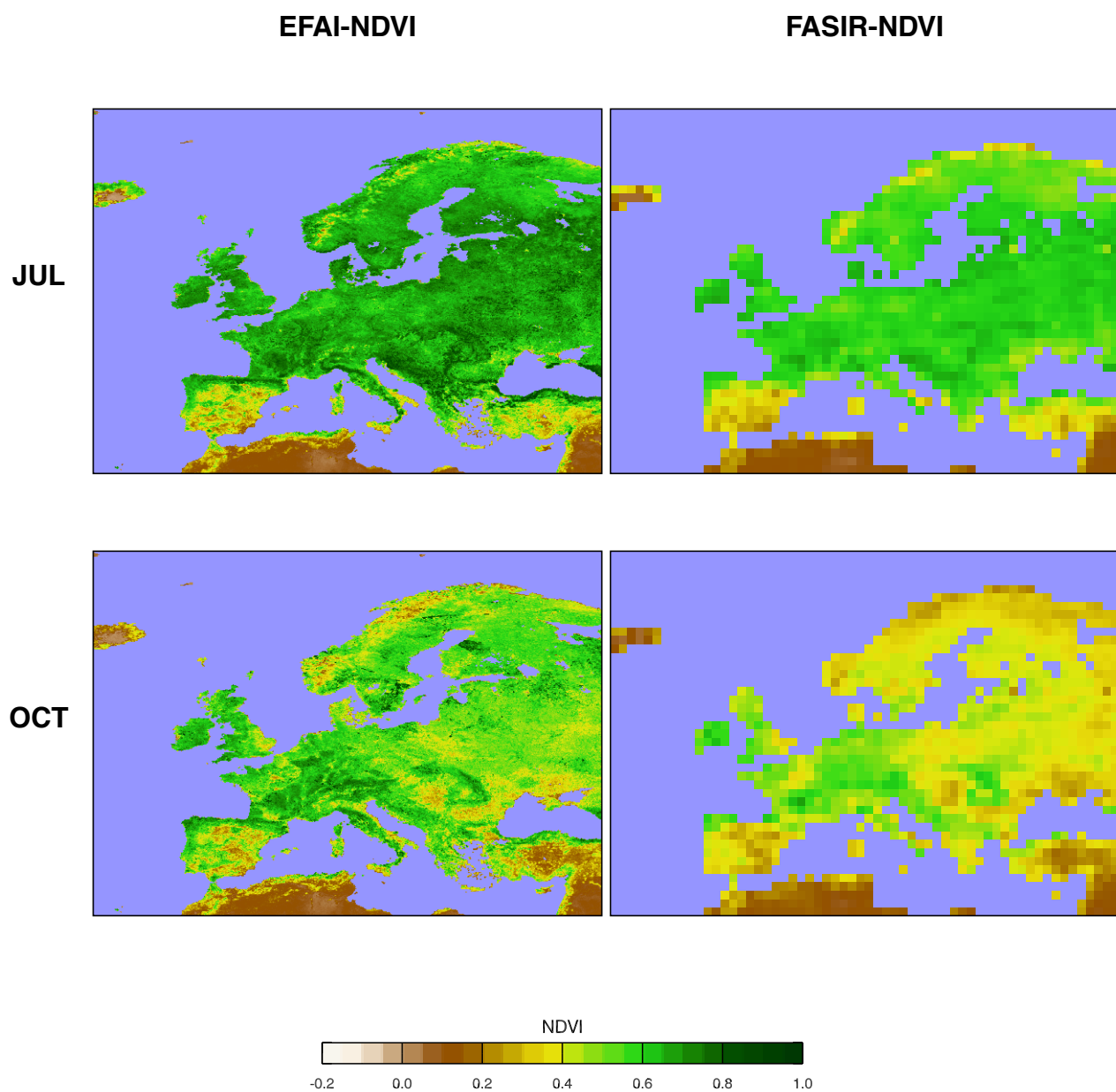


Figure C.2: 0.1° by 0.1° EFAI-NDVI (left) shown in comparison to the pre-released 1° by 1° FASIR-NDVI (right). The EFAI-NDVI is sampled for one month, since it has a temporal resolution of 10 days. top = July 1987, bottom = October 1987

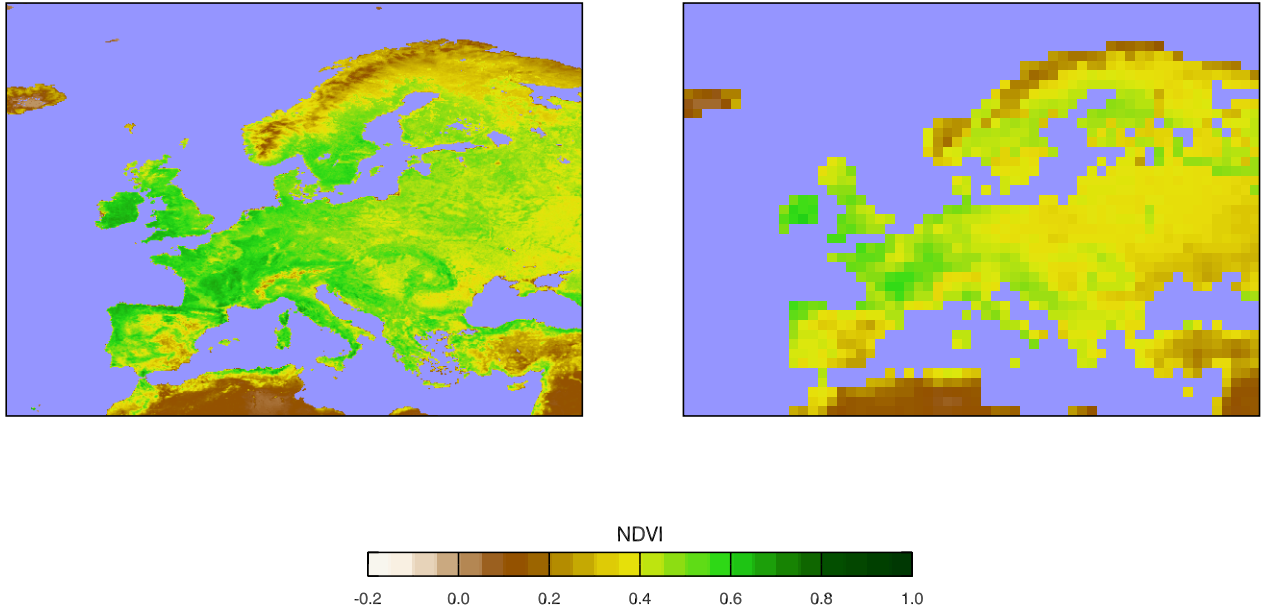


Figure C.3: 0.1° by 0.1° mean EFAI-NDVI (left) shown in comparison to the pre-released 1° by 1° mean FASIR-NDVI (right). The data was temporally averaged for the full time-series.

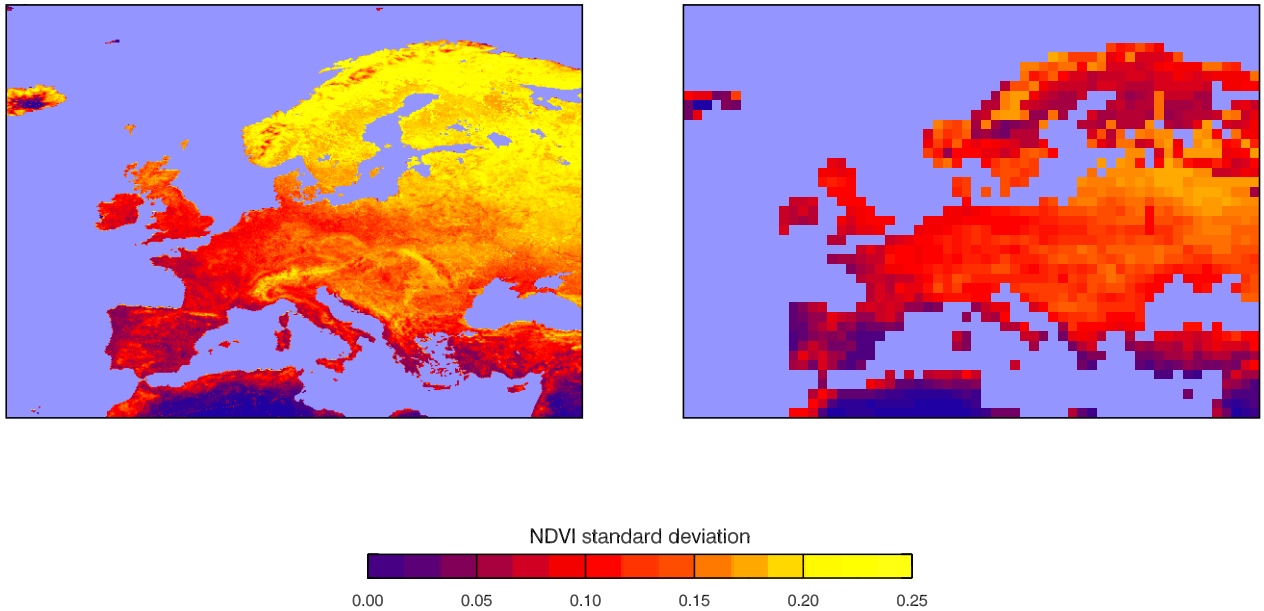


Figure C.4: Temporal standard deviation for the 0.1° by 0.1° EFAI-NDVI (left) and the 1° by 1° FASIR-NDVI (right). The full time-series was considered for both datasets.

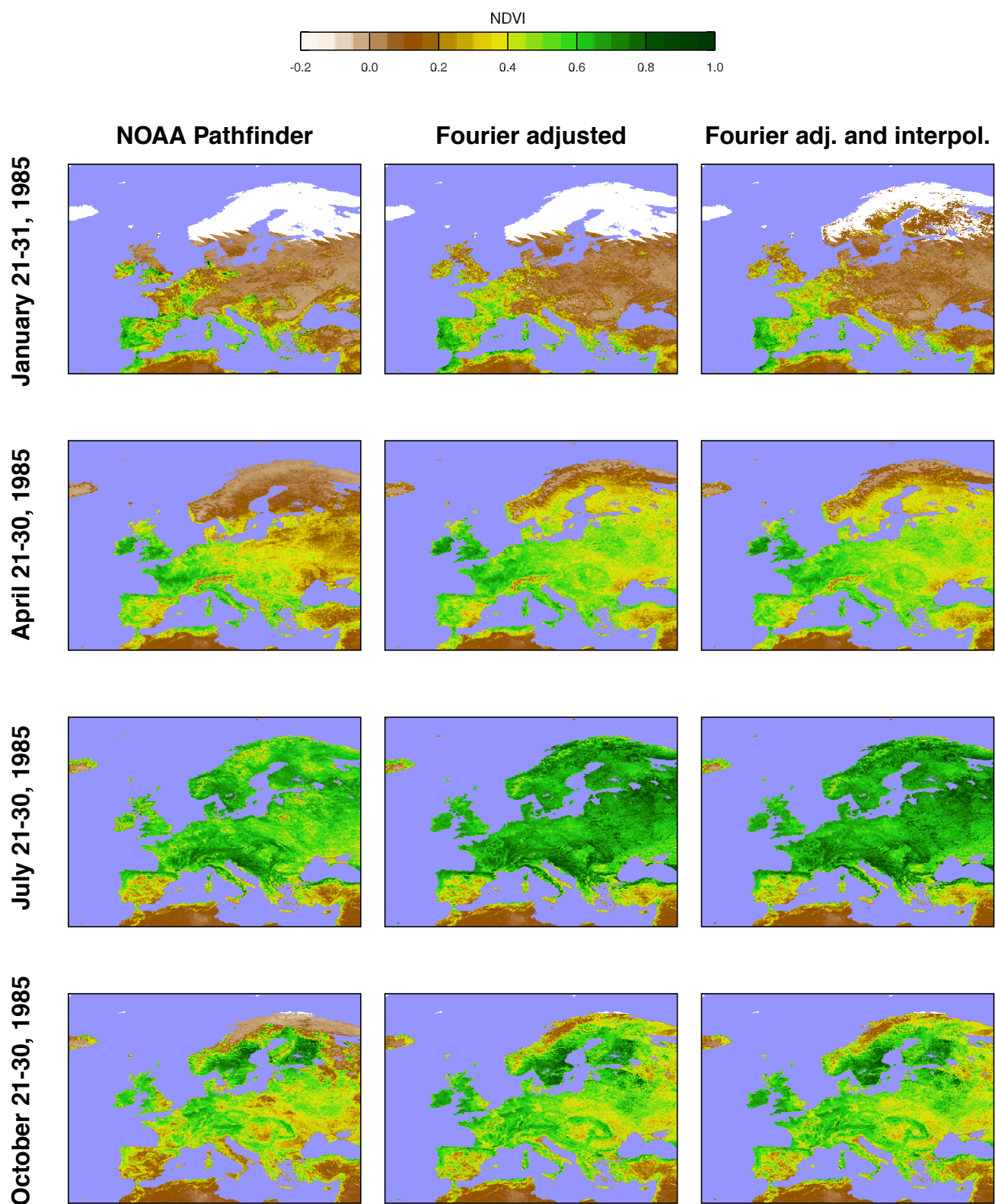


Figure C.5: Successive corrections are applied to the 1985 Land Pathfinder NDVI dataset. The Fourier adjustment shows most effect in spring and summer for mid and high latitudes. The interpolation of missing data during winter can be seen the January dataset.

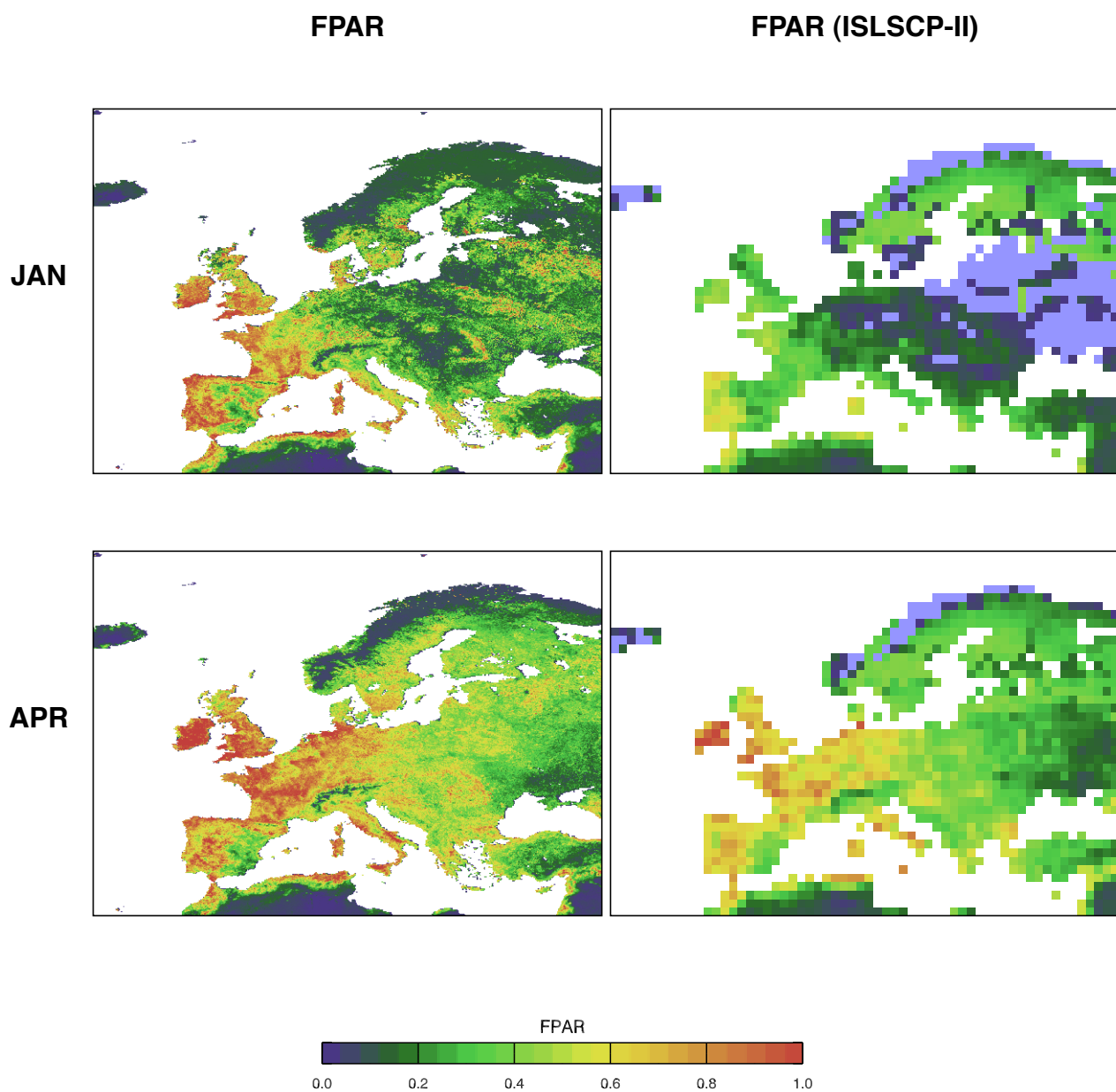


Figure C.6: European fields of FPAR calculated from EFAI-NDVI (left) in comparison to pre-released ISLSCP-II FPAR fields from the FASIR-NDVI (right). The top datasets show January 1987, the bottom datasets show April 1987.

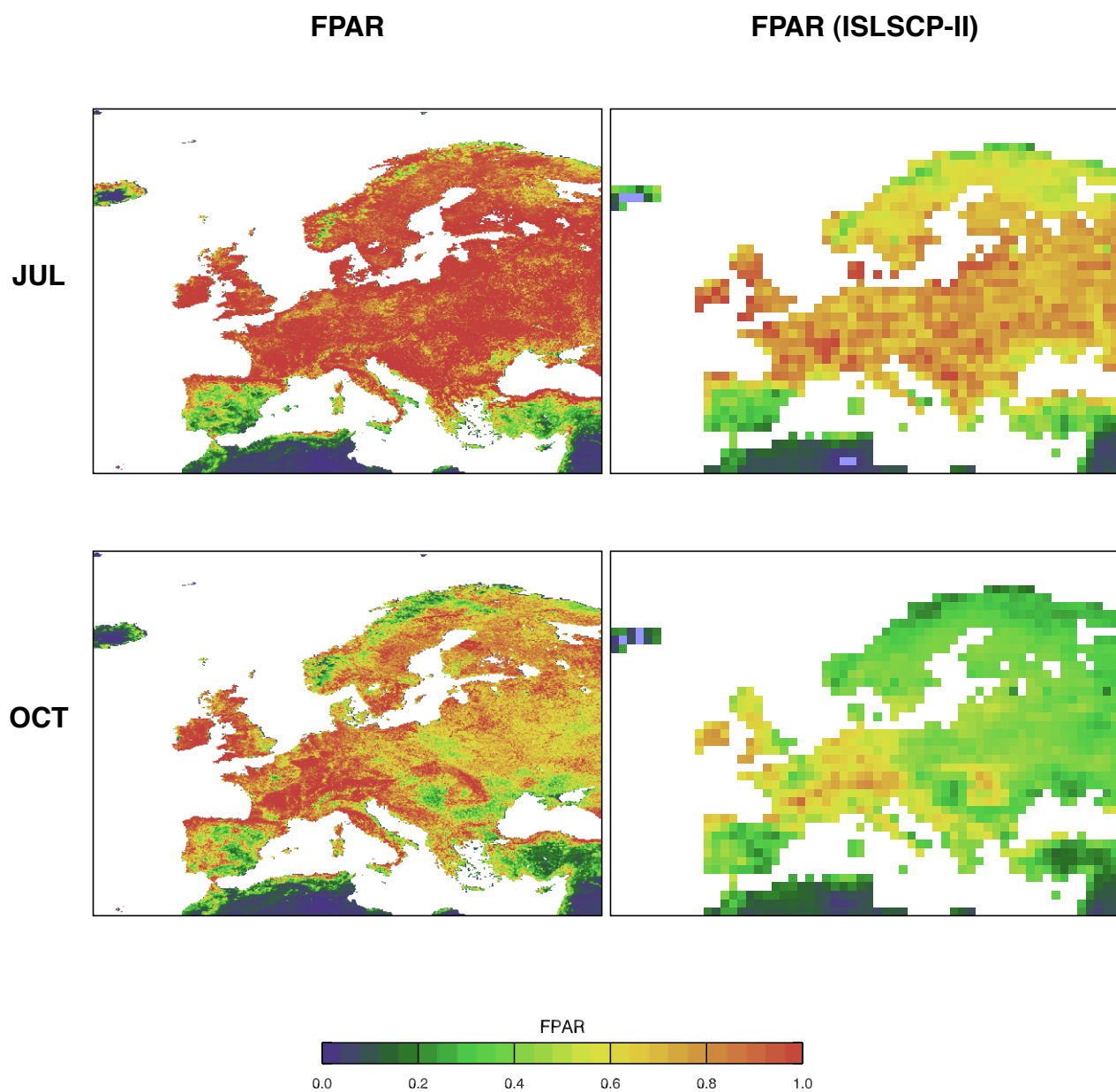


Figure C.7: European fields of FPAR calculated from EFAI-NDVI (left) in comparison to pre-released ISLSCP-II FPAR fields from the FASIR-NDVI (right). The top datasets show July 1987, the bottom datasets show October 1987.

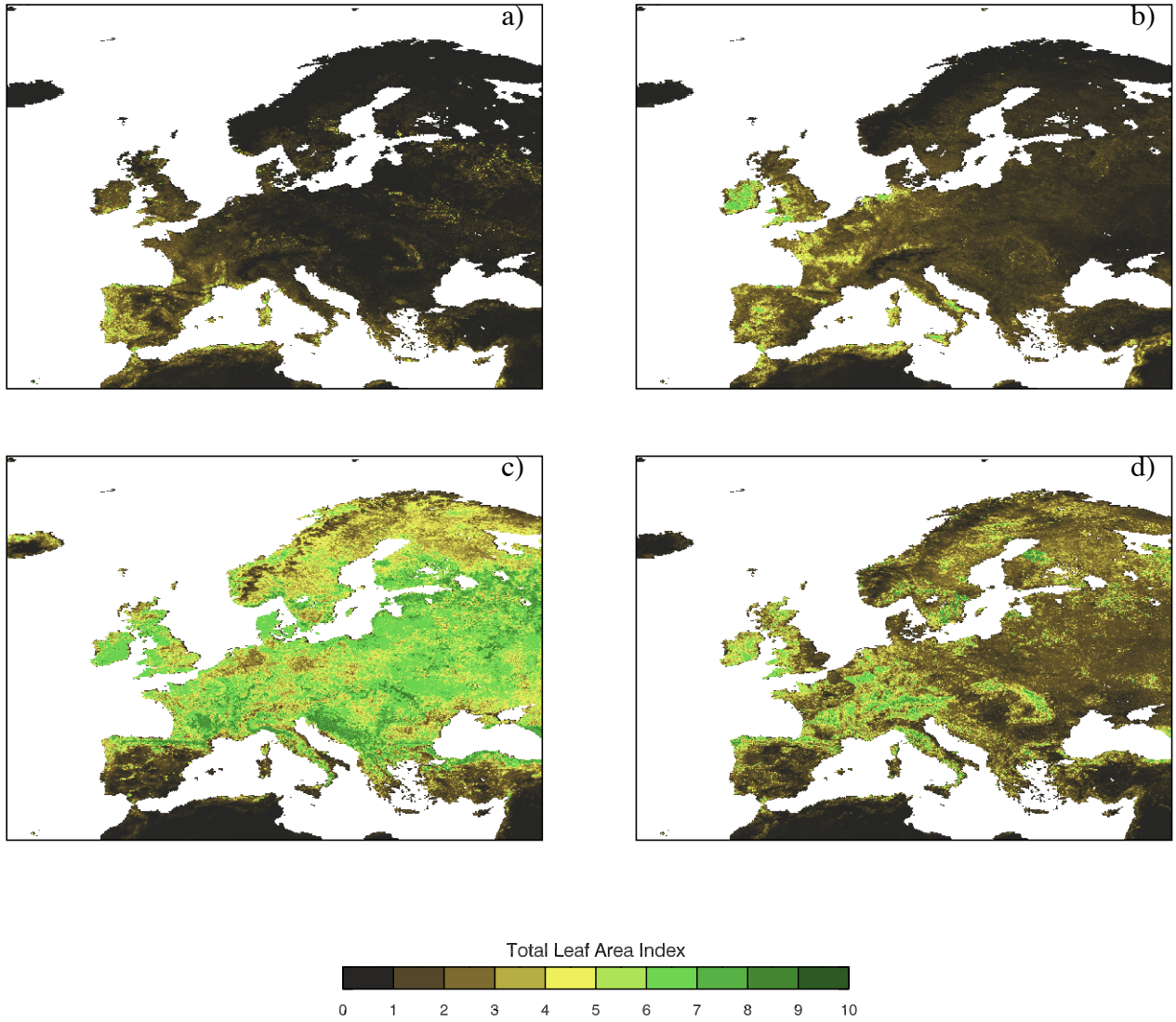


Figure C.8: Fields of Total Leaf Area Index L_T for the months a) January, b) April, c) July and d) October calculated from the 1987 EFAI-NDVI dataset. The L_T fields were assigned differently for each land cover class as described in chapter 7.

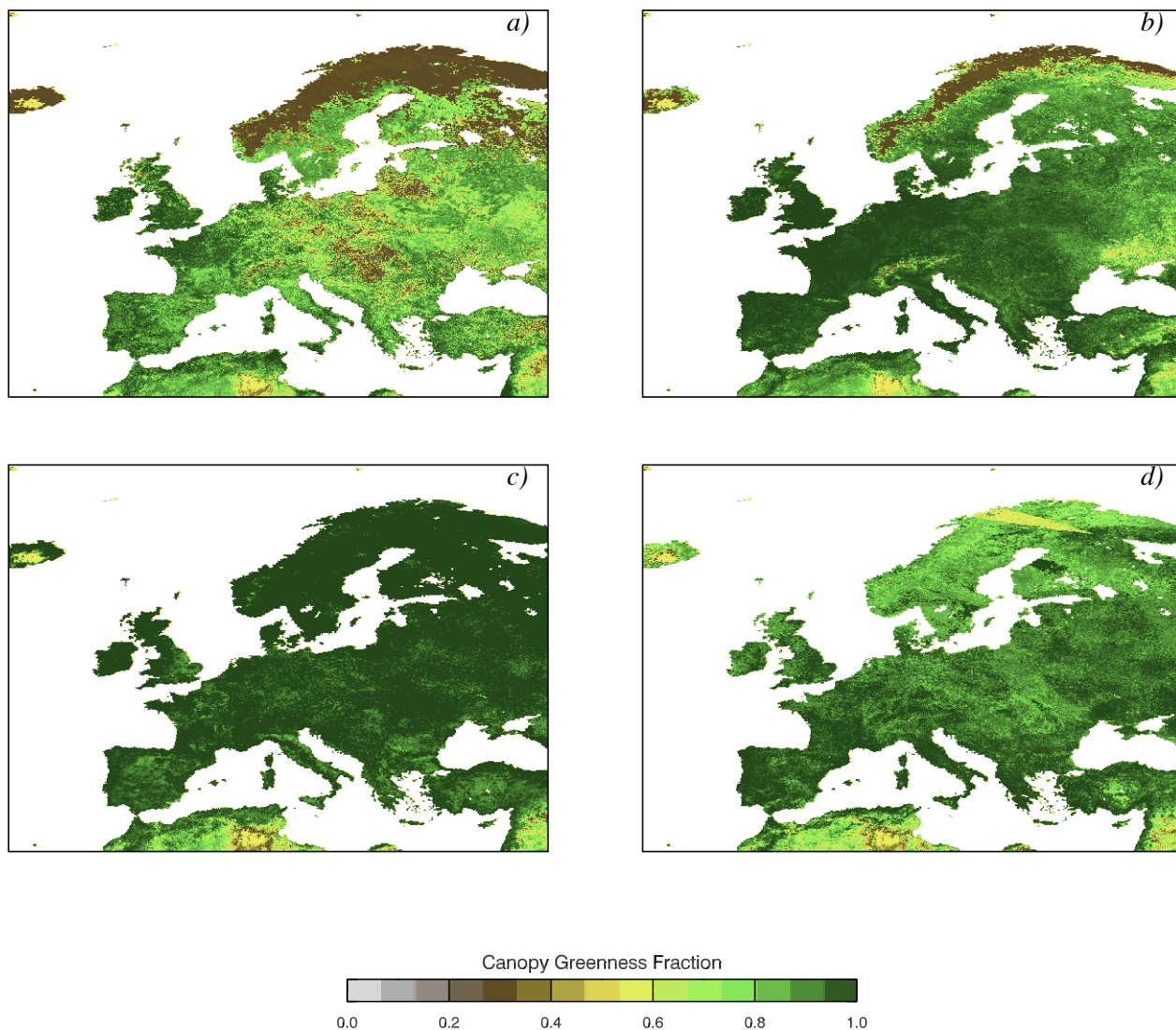


Figure C.9: Canopy Greenness Fraction derived from the 1987 EFAI-NDVI dataset as described in chapter 7. a) January, b) April, c) July and d) October.

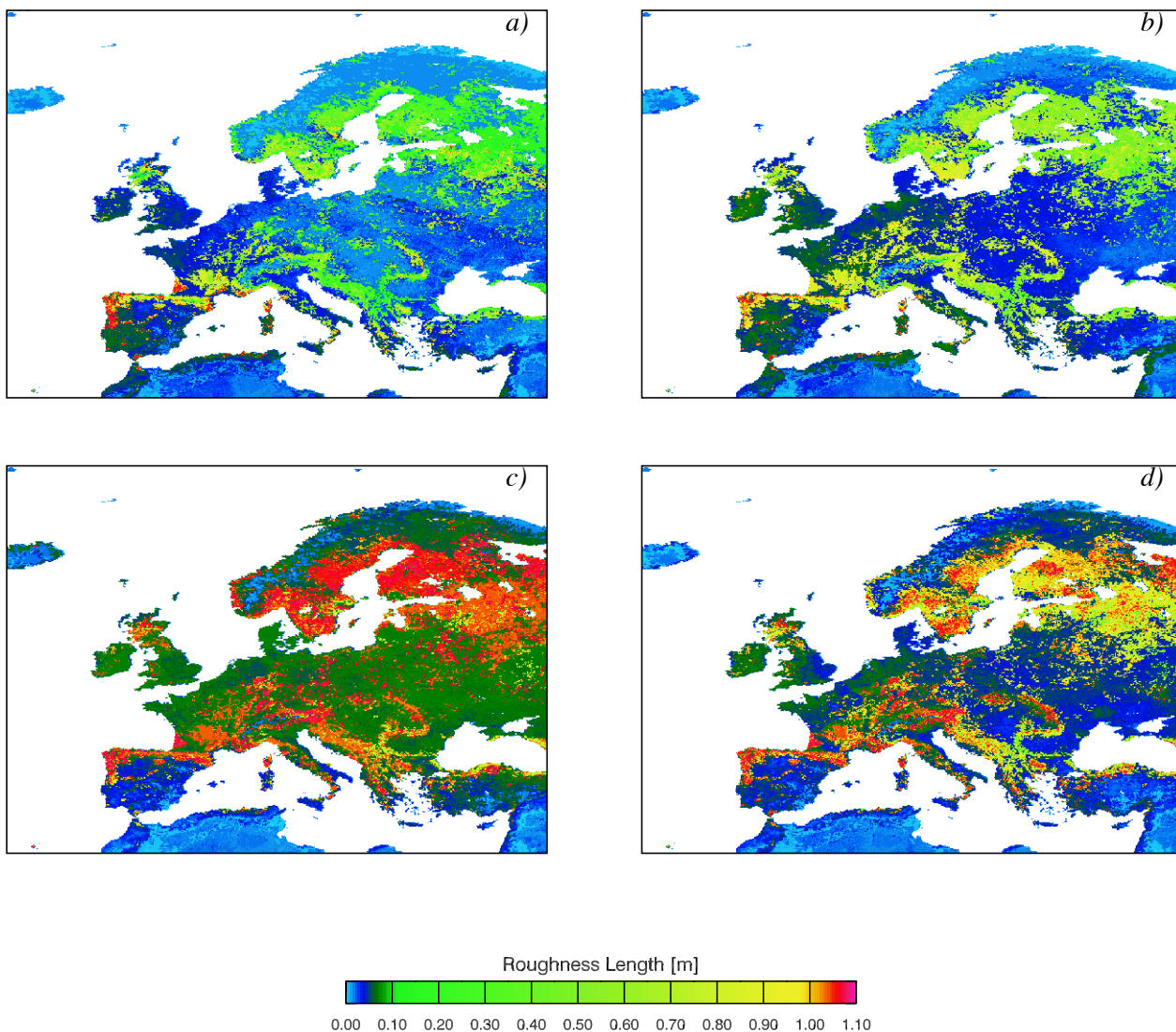


Figure C.10: *Roughness length, derived from the 1987 EFAI-NDVI by land cover class. a) January, b) April, c) July and d) October. The influence of snow is not considered at this point.*